

Genomic and metabolomic insights into the selection and differentiation of bioactive compounds in citrus

Xiao Liang^{1,9}, Yue Wang^{1,9}, Wanxia Shen^{2,9}, Bin Liao¹, Xiaojuan Liu¹, Zimeng Yang¹, Jiebiao Chen¹, Chenning Zhao¹, Zhenkun Liao¹, Jinping Cao^{1,3}, Ping Wang⁴, Peng Wang⁴, Fuzhi Ke⁴, Jianguo Xu⁴, Qiong Lin⁵, Wanpeng Xi⁶, Lishu Wang⁷, Juan Xu⁸, Xiaochun Zhao² and Chongde Sun^{1,3,*}

¹Plant Growth, Development and Quality Improvement, Zhejiang Provincial Key Laboratory of Integrative Biology of Horticultural Plants, Zhejiang University, Hangzhou, China

²Citrus Research Institute, Southwest University/Chinese Academy of Agricultural Sciences, Chongqing, China

³Hainan Institute of Zhejiang University, Sanya, China

⁴Zhejiang Citrus Research Institute, Taizhou, China

⁵Key Laboratory of Agro-products Quality and Safety Control in Storage and Transport Process, Ministry of Agriculture and Rural Affairs/ Institute of Food Science and Technology, Chinese Academy of Agricultural Sciences, Beijing, China

⁶College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China

⁷Department of Hematology and Hematopoietic Cell Transplantation, Comprehensive Cancer Center, City of Hope National Medical Center, Duarte, CA, USA

⁸Key Laboratory of Horticultural Plant Biology (Ministry of Education), Huazhong Agricultural University, Wuhan, China

⁹These authors contributed equally

*Correspondence: Chongde Sun (adesun2006@zju.edu.cn)

<https://doi.org/10.1016/j.molp.2024.10.009>

ABSTRACT

Bioactive compounds play an increasingly prominent role in breeding functional and nutritive fruit crops such as citrus. However, the genomic and metabolic bases for the selection and differentiation underlying bioactive compound variations in citrus remain poorly understood. In this study, we constructed a species-level variation atlas of genomes and metabolomes using 299 citrus accessions. A total of 19 829 significant SNPs were targeted to 653 annotated metabolites, among which multiple significant signals were identified for secondary metabolites, especially flavonoids. Significant differential accumulation of bioactive compounds in the phenylpropane pathway, mainly flavonoids and coumarins, was unveiled across ancestral citrus species during differentiation, which is likely associated with the divergent haplotype distribution and/or expression profiles of relevant genes, including *p*-coumaroyl coenzyme A 2'-hydroxylases, flavone synthases, cytochrome P450 enzymes, prenyltransferases, and uridine diphosphate glycosyltransferases. Moreover, we systematically evaluated the beneficial bioactivities such as the antioxidant and anticancer capacities of 219 citrus varieties, and identified robust associations between distinct bioactivities and specific metabolites. Collectively, these findings provide citrus breeding options for enrichment of beneficial flavonoids and avoidance of potential risk of coumarins. Our study will accelerate the application of genomic and metabolic engineering strategies in developing modern healthy citrus cultivars.

Key words: citrus, metabolome, genome-wide association study, bioactive compounds

Liang X., Wang Y., Shen W., Liao B., Liu X., Yang Z., Chen J., Zhao C., Liao Z., Cao J., Wang P., Wang P., Ke F., Xu J., Lin Q., Xi W., Wang L., Xu J., Zhao X., and Sun C. (2024). Genomic and metabolomic insights into the selection and differentiation of bioactive compounds in citrus. *Mol. Plant.* **17**, 1753–1772.

INTRODUCTION

Horticultural crops are rich in bioactive compounds and are serving as good models to dissect the associations between genomics and metabolomics and their functional bioactivities. Plant bioactive compounds encompass a series of synthesized metabolites that have specific health-promoting effects on humans. Human intake of these diverse bioactive compounds occurs through consumption of plant-derived foods such as grains, vegetables, and fruits (Samtiya et al., 2021). The primary metabolites, comprising sugars, lipids, proteins, nucleic acids, and vitamins, provide essential nutrients and possess bioactive functions as well (Ramadan and Oraby, 2016; Patil et al., 2020; Li-Rong, 2021). Besides, the secondary metabolites, encompass phenolic acids, flavonoids, coumarins, alkaloids, terpenoids, etc., play crucial roles in plant defense and also constitute the majority of plant bioactive compounds (Jan et al., 2021). The secondary metabolites possess a wide range of biological properties, including anti-inflammatory, antioxidant, anticancer, anti-diabetic, and anti-obesity properties, and they are capable of effectively reducing the incidence of chronic disease (Kris-Etherton et al., 2002; Xiao and Bai, 2019; Samtiya et al., 2021). These findings are compelling to support the notion of “medicine food homology” and are propelling the exploration and development of plant-based bio-fortified foods, which are valued for their inherent natural, safe, and health-promoting attributes.

Citrus are some of the world’s most widely cultivated and consumed fruit crops and are valued for their nutritional and pharmacological significance (Lu et al., 2023). Citrus are considered “powerhouse” fruits for their strong association with reduced risk of chronic disease (Nanney et al., 2004). Many citrus products, such as ju pi (dried mandarin peel), dai dai hua (the immature flower of the sour orange), and fo shou (fingered citron), are used as important materials in traditional Chinese medicine owing to their beneficial impact on human health (Hou and Jiang, 2013). Flavonoids, among the most profound bioactive compounds in citrus, exhibit various bioactivities (Cirmi et al., 2016; Wang et al., 2017b; Zhao et al., 2020; Deng et al., 2022), and O-methylation and glycosylation of citrus flavonoids increase their stability and bioavailability (Bao et al., 2018; Chen et al., 2018a). Coumarins, types of allelopathic substances in citrus, are another essential class of bioactive compounds with various bioactive functions (Riveiro et al., 2010; Chen et al., 2014). However, the furanocoumarin abundant in citrus also have controversial effects on humans due to the phototoxic and interaction with drugs (Dugrand et al., 2013). In addition, citrus contain abundant phenolic acids, carotenoids, terpenes, alkaloids, dietary fiber, limonoids, and many other bioactive compounds (Liu et al., 2021; Saini et al., 2022; Kaur et al., 2023). The increasing recognition of the biological activity within citrus is accelerating the development and consumption of citrus-based functional foods and beverages.

Metabolite profiles in crops often undergo reshaping during natural selection and artificial domestication (Tieman et al., 2017; Yuan et al., 2023). Citrus, with their origins in the Himalayas, have undergone a journey of continuous migration, leading to the emergence of several ancestral species (Wu et al., 2018).

Multi-omics analyses on citrus bioactive compounds

The natural variation, along with artificial hybridization and selection, has yielded a large number of current citrus species, and this process is also accompanied by changes in metabolites (Wang et al., 2018). However, the variation and genomic basis of the bioactive compounds of citrus during natural evolution remain unexplored. Metabolite-based genome-wide association study (mGWAS) offers a powerful toolkit for dissecting the genomic basis of population metabolic diversity (Fang and Luo, 2019), which has been widely applied in crops (Peng et al., 2017; Tieman et al., 2017; Zhu et al., 2018; Zhang et al., 2022). There are many reported studies of genome sequences and evolution of citrus (Wu et al., 2014, 2018; Huang et al., 2023), and prior studies have reported the evolution-associated changes of bioactive metabolites and the key genes related to metabolite synthesis of pummelos through mGWAS (Shen et al., 2023; Zheng et al., 2023). However, the changes of bioactive metabolism of the entire citrus populations have not been studied. In addition, exploring citrus varieties with notable bioactivities using multi-omics tools and functional bioactivity evaluations would also be meaningful for establishing future citrus breeding goals and consumption guidance under the notion of “One Health”.

In this study, we performed genome resequencing and metabolomics of 299 representative accessions of extant citrus species, providing a global landscape of genetic variations and metabolic profiles of citrus populations. The results revealed the genetic basis of differential selection in the phenylpropane pathway during differentiation of citrus populations, which likely led to the featured synthesis of different bioactive compounds. Through mGWAS, we identified candidate genes related to the synthesis of flavonoids and coumarins, the important bioactive compounds in citrus. Furthermore, we validated the relationships between prominent biological activities and bioactive compounds using large-scale bioactivity evaluations and clarified the variation route of bioactivities alongside the evolution of citrus populations. These results will provide multi-omics insights into the enrichment and utilization of bioactive compounds in citrus and are helpful for the design breeding of citrus varieties with higher nutritional value.

RESULTS

Phylogenomic and metabolomic relationships among citrus species

A total of 299 citrus accessions originating from different countries and regions were collected (Supplemental Table 1) for genome resequencing: 104 mandarins (*Citrus reticulata*), 41 sweet oranges (*C. sinensis*), 45 pummelos (*C. grandis*), 6 citrons (*C. medica*), 7 sour oranges (*C. aurantium*), 7 grapefruits (*C. paradisi*), 7 lemons (*C. limon*), 3 limes (1 lime [*C. aurantiifolia*], 1 rangpur lime [*C. limonia*], and 1 rough lemon [*C. jambhiri*]), 8 yuzu (*C. junos*), 3 xiangyuan (*C. wilsonii*), 9 papedas (8 *C. ichangensis* and 1 *C. hystriculus*), 1 *C. mangshanensis*, 48 tangors/tangelos, 6 kumquats (*Fortunella*), and 4 *P. trifoliata*. One *Atalantia buxifolia* was used as an outgroup for phylogenetic analysis (Huang et al., 2023). A total of 1.64 Tb of clean bases were obtained with an average depth of 18.13x and a properly paired ratio of 94.53% using the reference genome Clementina v.1.0 (Wu et al., 2014). The analysis generated 654 929 high-quality SNPs for further analysis (Supplemental Table 2).

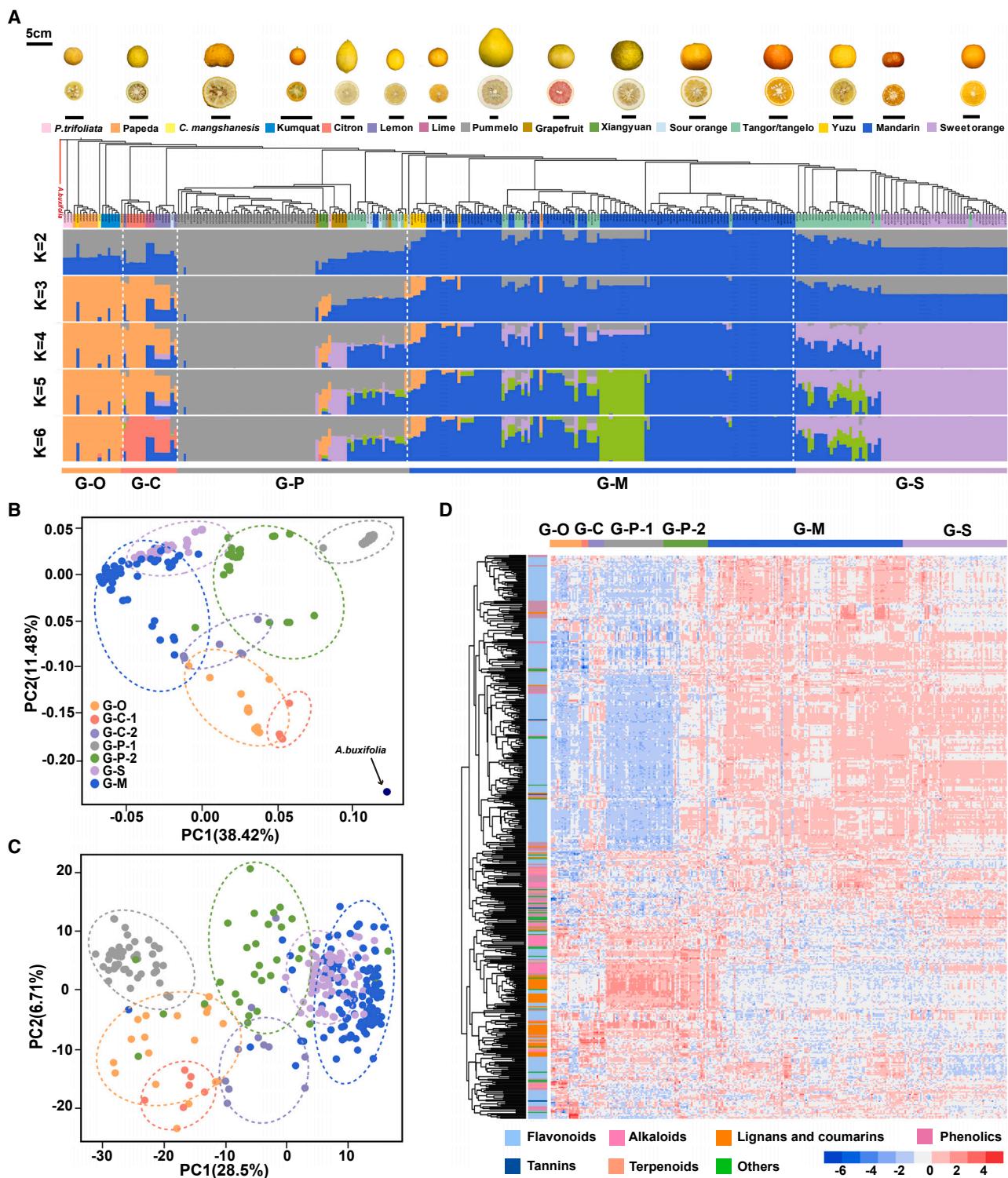


Figure 1. Phylogenetic relationships and secondary metabolic differentiations among 299 citrus accessions.

(A) Maximum-likelihood phylogenetic tree and population structure of 299 citrus accessions using an *A. buxifolia* accession as an outgroup. All accessions were separated into 5 groups: G-O, G-C, G-P, G-M, and G-S.

(B) PCA of different citrus groups based on SNPs.

(C) PCA of different citrus groups based on metabolites in flavedo. The group remarks are the same as those in (B).

(D) Heatmaps of secondary metabolite variation in citrus groups. The data were normalized horizontally, with red representing a higher relative content and blue representing a lower relative content.

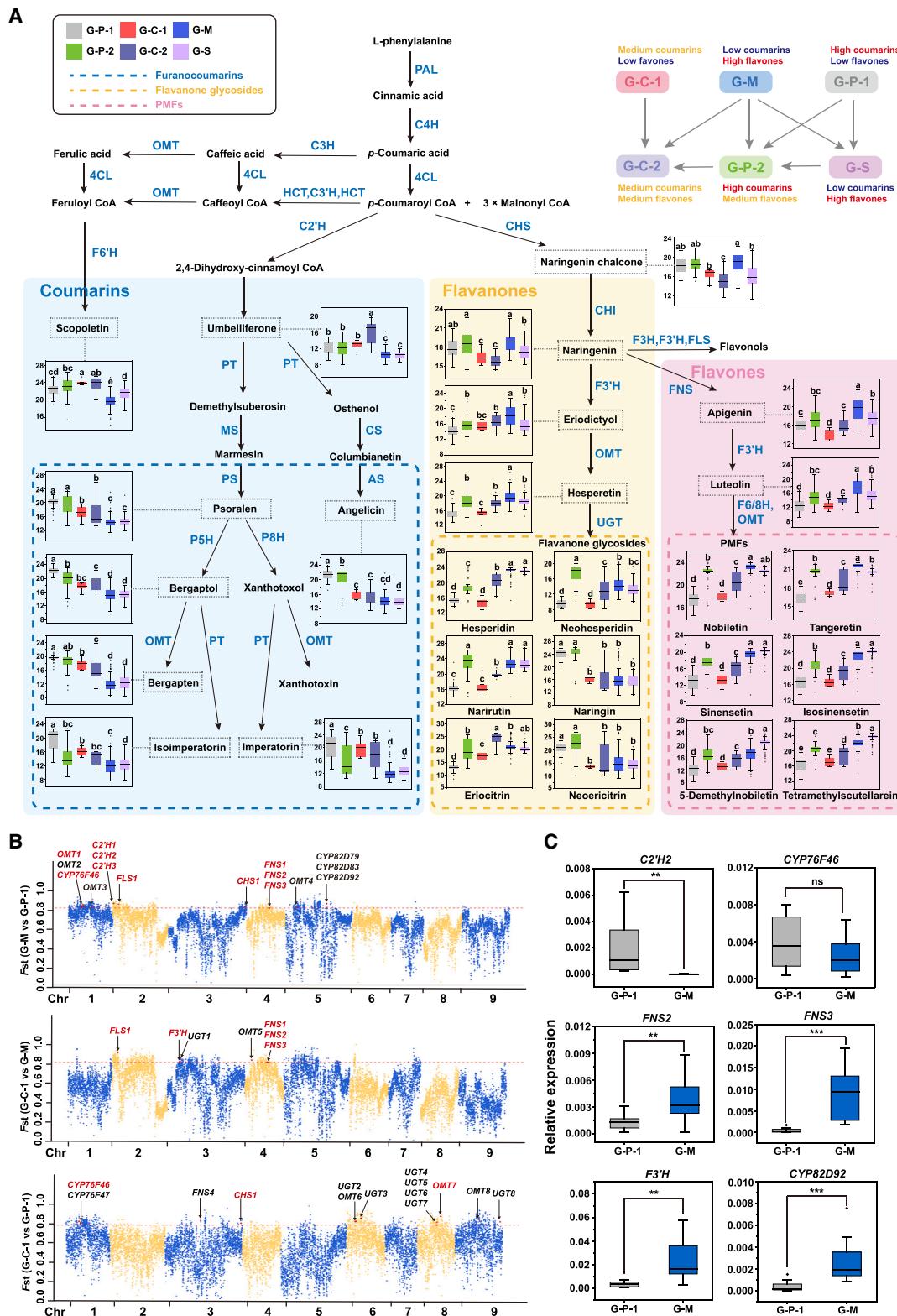


Figure 2. Population differentiation of metabolite content in the phenylpropane pathway and selective sweep analysis of related genes.

(A) Schematic of the phenylpropane metabolism pathway and content variation of important products in the pathway among six groups. For boxplots, different lowercase letters represent statistically significant differences ($p < 0.05$). The relative contents of metabolites were normalized by log2. The

(legend continued on next page)

A maximum-likelihood phylogenetic tree of all 299 accessions was constructed, and the population structure was analyzed (Figure 1A). When $K = 6$, the population can be broadly categorized into G-C (group of citron), G-P (group of pummelo), G-M (group of mandarin), G-S (group of sweet orange), and G-O (group of other citrus accessions) (Supplemental Table 3). Notable accumulation patterns were observed within these clusters. G-O includes *P. trifoliata*, papedas, and *C. mangshanesis*, which were regarded as early-diverging *Citrus*, along with kumquats, which exhibited a unique evolutionary position compared with other *Citrus* plants (Wang et al., 2017a; Huang et al., 2023). G-C was mainly composed of citrons and their hybrids, including lemons and limes. G-P contained the accessions that primarily showed the pummelo-related genetic background, including pure pummelos, and the hybrids, such as grapefruits, tangelos, sour oranges, and xiangyuan. G-M contained most of the mandarins and hybrids with more mandarin genetic background, including pure and admixture mandarins, and some of the tangors/tangelos. Accessions in G-S represented a more genetically sweet orange background, including most sweet oranges and tangor hybrids. The results of phylogenetic principal-component analyses (PCAs) based on SNPs showed that G-C and G-P can be both separated into two subgroups, in which G-C-1 contained pure citrons, G-C-2 contained lemons and limes, G-P-1 included pure pummelos, and G-P-2 covered pummelo-related hybrids such as grapefruits, sour oranges, tangelos, and xiangyuan (Figure 1B and Supplemental Figure 1). Citron, pummelo, and mandarin represented the three ancestral populations, which were clearly distinguished and located at three vertices of the graph, while the hybrids were clustered between the three ancestral groups. G-O stayed far away from cultivated accessions in PC2 and was close to G-C-1. All citrus populations could be well separated in both PC1 and PC2, and the results were consistent with those of phylogenetic tree and population structure analyses. In conclusion, our citrus populations have comprehensive clustering profiles with both clear differentiations and genetic structures that facilitate our subsequent analyses.

In citrus fruits, the peel, particularly the flavedo, is found to contain the richest concentration of functional compounds (Yang et al., 2017). We constructed a citrus-specific metabolome library using a broadly targeted metabolomic technology based on liquid chromatography–tandem mass spectrometry (LC–MS/MS), and the metabolites in the flavedo of all the citrus accessions were examined qualitatively and quantitatively. Integrating metabolite databases, 728 of 3858 metabolite data were validated and annotated, including flavonoids, lignans and coumarins, organic acids, lipids, amino acids and derivatives, nucleotides and derivatives, phenolic acids, alkaloids, terpenoids, and others (Supplemental Table 4). Among them, secondary metabolites account for the largest proportion (61% in total), especially flavonoids (38% in total) (Supplemental Figure 2). More than 95% of the metabolic traits showed a coefficient of

variation (CV) greater than 50%, and 40.36% of the metabolite traits had a CV greater than 200%, indicating significant differences in metabolite features of the citrus populations (Supplemental Figure 3; Supplemental Table 5).

Metabolome-based PCA revealed distinct groupings according to the metabolic profiles (Figure 1C and Supplemental Figure 4), mirroring the genetic divisions observed in the SNP-based PCA. This also indicated that the metabolic profiles of citrus flavedo can be used as characteristic biomarkers to distinguish diverse citrus populations. While primary metabolites showed no pronounced differences among groups (Supplemental Figure 5), secondary metabolites, especially flavonoids and coumarins derived from the phenylpropane metabolic pathway, exhibited distinct patterns across seven groups (Figure 1D). Most of the flavonoids, especially flavones, were notably enriched in G-S and G-M, whereas all furanocoumarins and most of the simple coumarins were specifically accumulated in G-P, especially G-P-1. In comparison, the amount of coumarin enriched in G-C is less than that in G-P, while G-M and G-S have rare coumarin accumulation (Supplemental Figure 6). Therefore, we hypothesized that ancestral groups, especially mandarins and pummelos, may have undergone selective pressures favoring divergent phenylpropane synthesis branches during their early differentiation that were passed on to their hybrid offspring, resulting in metabolic differences between groups.

Genomic basis of the differentiation of bioactive compounds in citrus

To explore the genetic determinants behind the distinct accumulation of flavonoids and coumarins among different citrus groups, we compared the contents of metabolites in the synthesis pathway of flavonoids and coumarins in all groups except G-O (Figure 2A; Supplemental Table 6). Under the catalysis of *p*-coumaroyl coenzyme A 2'-hydroxylase (C2'H), *p*-coumaroyl coenzyme A (CoA) was catalyzed to form umbelliferone, the basic structure initiating the coumarin pathway (Vialart et al., 2012). The contents of umbelliferone in G-Ps and G-Cs were significantly higher than that in G-M and G-S. Further along the pathway, G-P-1 showed the highest accumulation of furocoumarins, followed by G-P-2, while G-Cs were moderate. Another corridor of synthesis involves *p*-coumaroyl CoA combining with malonyl CoA under the influence of chalcone synthase to yield naringenin chalcone, the precursor of flavanones and flavones (Kreuzaler and Hahlbrock, 1972). After successive catalysis by chalcone isomerase and flavone synthase, G-M stood out with significantly higher levels of apigenin than other groups. Through continuous catalysis by hydroxylases and O-methoxytransferases, polymethoxylated flavones (PMFs) significantly accumulated in G-M and G-S compared with other groups, whereas G-Ps and G-Cs still maintained low content of flavones. For flavanone accumulation, naringenin, eriodictyol,

corresponding protein names for the abbreviations in this figure are listed in Supplemental Table 6. The chart at the top right shows parental relationship (illustrated by arrows) and coumarin/flavone content comparison among six groups.

(B) Distribution of *Fst* values from the comparison of G-M, G-P-1, and G-C-1. The top 5% region was selected as the candidate selective sweep. The genes in the overlapped *Fst* or XP-CLR differentiation regions of different pairwise groups are shown in red. Accession numbers of genes involved in the figures are listed in Supplemental Table 7.

(C) Relative expression levels of some structure genes in the differentiation sweep regions of G-M and G-P-1. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

and hesperetin accumulated more in G-M, followed by G-S and G-P-2, leading to downstream accumulation of narinutin, eriocitrin, hesperidin, and neohesperidin in these groups. However, the contents of neoeriocitrin and naringin were much higher in G-P-1 and G-P-2 compared to other flavanone glycosides, a difference we speculated might be related to the functional differentiation of uridine diphosphate (UDP) glycosyltransferases (UGTs) in different citrus species (Frydman et al., 2013).

In the flavedo, the six groups showed different metabolite accumulation characteristics, especially coumarins and flavones (Figure 2A), disclosing the preferences of the three ancestral species for different branches of the phenylpropane pathway. Among all groups, G-M and G-P-1 exhibited the most pronounced difference in coumarin/flavone content. We conducted population-selective sweep analysis between G-M, G-P-1, and G-C-1. A high fixation index (*Fst*) value (over 0.8) was found between three ancestor groups, suggesting a high degree of differentiation (Figure 2B). Numerous structural genes involved in flavonoids and coumarins synthesis were found in the differentiation sweep regions of G-M and G-P-1, including *C2'H2*, *FNSs*, *OMTs*, *CHSs*, *CHI*, *UGTs*, *PTs*, and *CYP450* members. Cross Population Composite Likelihood Ratio (XP-CLR) analysis uncovered a series of genes overlapped with *Fst* differentiation sweeps (Supplemental Figure 7). By contrast, in the selective sweeps of G-C-1 vs. G-M and G-C-1 vs. G-P-1, relatively fewer genes were observed.

A series of key structure genes in the differentiation sweep regions of G-M and G-P-1 exhibited different expression levels, such as *C2'H2*, *CYP76F46*, *FNS2*, *FNS3*, *F3'H*, and *CYP82D92* (Figure 2C). Two haplotypes in the promoters of *C2'H2*, *FNS2*, *FNS3*, and *CYP82D92* were observed, in which *C2'H2-Hap1* was correlated with a high gene expression level and high content of coumarins in citrus, while *FNS2-Hap1*, *FNS3-Hap1*, and *CYP82D92-Hap1* likely contributed to flavone accumulation (Supplemental Figure 8). This finding indicated that promoter region variation induced by differentiation may affect gene expression and thus determine the accumulation preference for metabolites in the phenylpropane pathway. Furthermore, G-S and G-P-2 are two hybrid offspring of G-M and G-P-1, and the accumulation preference of coumarins/flavones of G-S is similar to G-M, whereas G-P-2 is similar to G-P-1 (Figure 2A). We selected sweet oranges and grapefruits as representatives of G-S and G-P-2, and through calculation of SNP density preference, we identified the 44.62M region from mandarins and 1.18M region from pummelos in the genome of sweet oranges, the 3.95M region from mandarins, and the 74.83M region from pummelos in the genome of grapefruits (Supplemental Figure 9). In these regions, we also found many genes related to coumarin and flavonoid synthesis (Supplemental Table 8). Thus, the accumulation preference of coumarins/flavones in G-S and G-P-2 might somehow be related to the proportion of originated regions from corresponding parental ancestors.

Landscape of metabolite-associated loci determined by mGWAS

To explore the functional genes related to metabolic differences between citrus groups, a total of 279 citrus accessions (20 sweet oranges with high homogeneity were removed) were selected to

Multi-omics analyses on citrus bioactive compounds

perform mGWAS using a mixed linear model. We identified a total of 423 330 high-quality biallelic SNPs with minor-allele frequency ≥ 0.05 and missing call frequency ≤ 0.01 for further analysis. Of the 3858 metabolites, 3514 had at least 1 significant association site. This yielded a total of 58 293 non-repetitive significant SNP sites from 3514 metabolites ($p \leq 1.24e-06$) with an average of 16 significant sites per metabolite (Supplemental Table 9). Furthermore, a total of 19 829 non-repetitive significant SNP sites were obtained from 653 annotated metabolites. The identification and statistics of significant association sites of various metabolites provided a reference for screening potential candidate genes and exploring the genetic basis of natural variation of metabolites within citrus species.

Among all metabolites, flavonoids had the highest number of significant loci widely distributed on chromosomes. Besides, the co-localization of significant loci of multiple metabolites was mainly found in secondary metabolites, including phenolic acids, lignans and coumarins, and flavonoids and alkaloids (Figure 3, black triangles). For example, between Chr6:24.75M-Chr6:25.58M, the significant loci of five flavonoids (including vitexin 2''-O-rhamnoside, isovitexin 2''-O-arabinoside, apigenin 6-C-glucoside-7-O-glucoside, genistein 8-C-apiosyl (1 → 6) glucoside, and apigenin 6-C-(2''-xylosyl) glucoside) were significantly enriched, and they co-localized at eight significant loci. A distinct region from Chr1:25.76M-Chr1:28.93M showed significant enrichment of the significant loci of alkaloids, such as *N*-feruloyl spermidine, *N*-feruloylagmatine, *N*-feruloylputrescine, etc. The clustered distribution of significant loci within the same metabolite category suggests potential key genes influencing metabolite synthesis in these specific genomic regions during citrus species selection and differentiation.

Candidate genes associated with coumarin accumulation in citrus

We searched the mGWAS signals of all coumarins and found a gene cluster consisting of tandem repeats of four *C2'H* genes (*Ciclev10015715m*, *Ciclev10018343m*, *Ciclev10017914m*, and *Ciclev10015700m*) in the 50 to 100-kb region from the lead SNP (chr2: 442 595) of the strong peak associated with psoralen, a typical furanocoumarin (Figure 4A). *C2'H* belongs to the Fe(II)- and 2-oxoglutarate-dependent dioxygenase family, which also act as feruloyl CoA 6'-hydroxylase in some species (Kai et al., 2008). We collected the sequences of functional 2-oxoglutarate-dependent dioxygenases reported in different plant species and constructed a phylogenetic tree together with four candidate *C2'H* genes, and they clustered in the same branch with the reported functional *C2'H*/feruloyl CoA 6'-hydroxylase (Supplemental Figure 10). The SNPs located in the upstream or coding sequence (CDS) region of *Ciclev10018343m*, *Ciclev10017914m*, and *Ciclev10015700m* were significantly associated with the content of psoralen (Figure 4B and Supplemental Figure 11A; Supplemental Table 10). These three *C2'H*s were also located in the genetic differentiation and selective sweep regions of G-M and G-P-1 (*Ciclev10018343m*, *Ciclev10017914m*, and *Ciclev10015700m* were equivalent to *C2'H1*, *C2'H2*, and *C2'H3*, respectively) (Figure 2B), suggesting a potentially important role in differential coumarin accumulation. Different nucleotide diversity indicated that the gene and promoter regions of the three *C2'H*s had undergone selection in

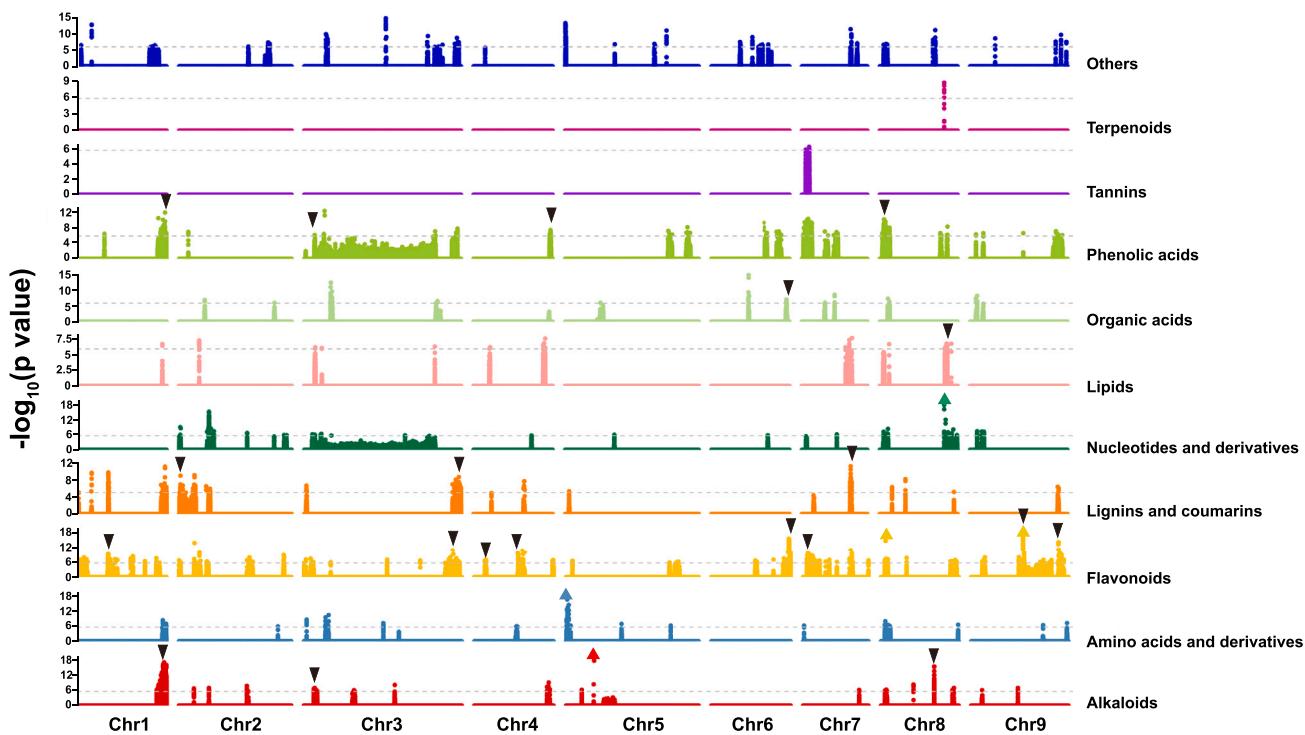


Figure 3. The distribution of significant mGWAS signals of annotated metabolites on the chromosomes.

Different colors refer to different classifications of metabolites. Black triangles represent a co-localization region of significant loci of multiple (more than three) metabolites in the same classification. The gray dashed lines represented the threshold of mGWAS and the arrows represented p values below $1E-20$.

G-M compared with G-P-1 (Supplemental Figure 11B). All accessions in G-S and G-M showed haplotypes with low psoralen content, while most accessions in G-P-1 and G-C-1 contained high-psoralen-content haplotypes (Figure 4B). Among three C2'Ns, only *Ciclev10017914m* was expressed higher in citrus accessions with high content of psoralen than those with low content of psoralen (Figure 4C).

We further carried out *in vitro* enzymatic assays of the recombinant proteins encoded by *Ciclev10018343m*, *Ciclev10017914m*, and *Ciclev10015700m* and demonstrated that, when *p*-coumaroyl CoA and feruloyl CoA were used as substrates, the three recombinant proteins could catalyze the production of umbelliferone and scopoletin, respectively (Figure 4D). Moreover, even when a nonsynonymous SNP was found in the *Ciclev10018343m* CDS, recombinant proteins of both alleles could catalyze the substrate to produce corresponding products with no difference in content (Supplemental Figure 12). Subsequently, we overexpressed these 3 C2'Ns in the peel of lemons using *Agrobacterium*-mediated transient overexpression. The results suggested that the overexpression of *Ciclev10017914m* and *Ciclev10015700m* in lemon peel significantly increased the content of limettin, which is the downstream product of umbelliferone and the main coumarin with the highest content in lemon peel (Zhao et al., 2022) (Supplemental Figure 13). We further assessed the fluctuations of coumarin content alongside the transcript levels of C2'N genes during citrus development using *C. grandis* 'Jiangxi Zaoyou' as material, discovering that the coumarin content in the flavedo gradually decreased during the maturation of

C. grandis 'Jiangxi Zaoyou', and *Ciclev10017914m* exhibited a high transcript level in the early developmental stage, showing a significant positive correlation with coumarin accumulation (Supplemental Figure 14).

In the 15.73 Mb region of Chr7, we observed robust association signals of the content of four furanocoumarins: isooxypeucedaneine, imperatorin, isoimperatorin, and cinidicin (Figure 4E), leading us to identify two candidate genes within the region: *Ciclev10027371m* and *Ciclev10025349m*. *Ciclev10027371m* belongs to the UbiA family and acts as a prenyltransferase (PT) in the synthesis of furanocoumarins (Munakata et al., 2021). The SNP7:15699890 at 5' upstream of *Ciclev10027371m* was significantly associated with the content of three furanocoumarins (Supplemental Figures 15A and 16; Supplemental Table 11). A phylogenetic tree of reported PTs showed that *Ciclev10027371m* was clustered in the same evolutionary branch with the functional PTs in citrus (Supplemental Figure 17). *Ciclev10025349m* belongs to the cytochrome P450 (CYP450) gene family (Liu et al., 2023) and plays a pivotal role in the synthesis and hydroxylation of furanocoumarins (Kruse et al., 2008; Larbat et al., 2009). The non-synonymous SNP and 5' upstream SNP of *Ciclev10025349m* showed significant correlations with the content of four furanocoumarins (Supplemental Figures 15A and 16; Supplemental Table 12), and the gene region of *Ciclev10025349m* has undergone selection in G-P-1 compared with G-M (Supplemental Figure 15B). Moreover, a positive correlation between the expression levels of *Ciclev10027371m* and *Ciclev10025349m* and the contents of four

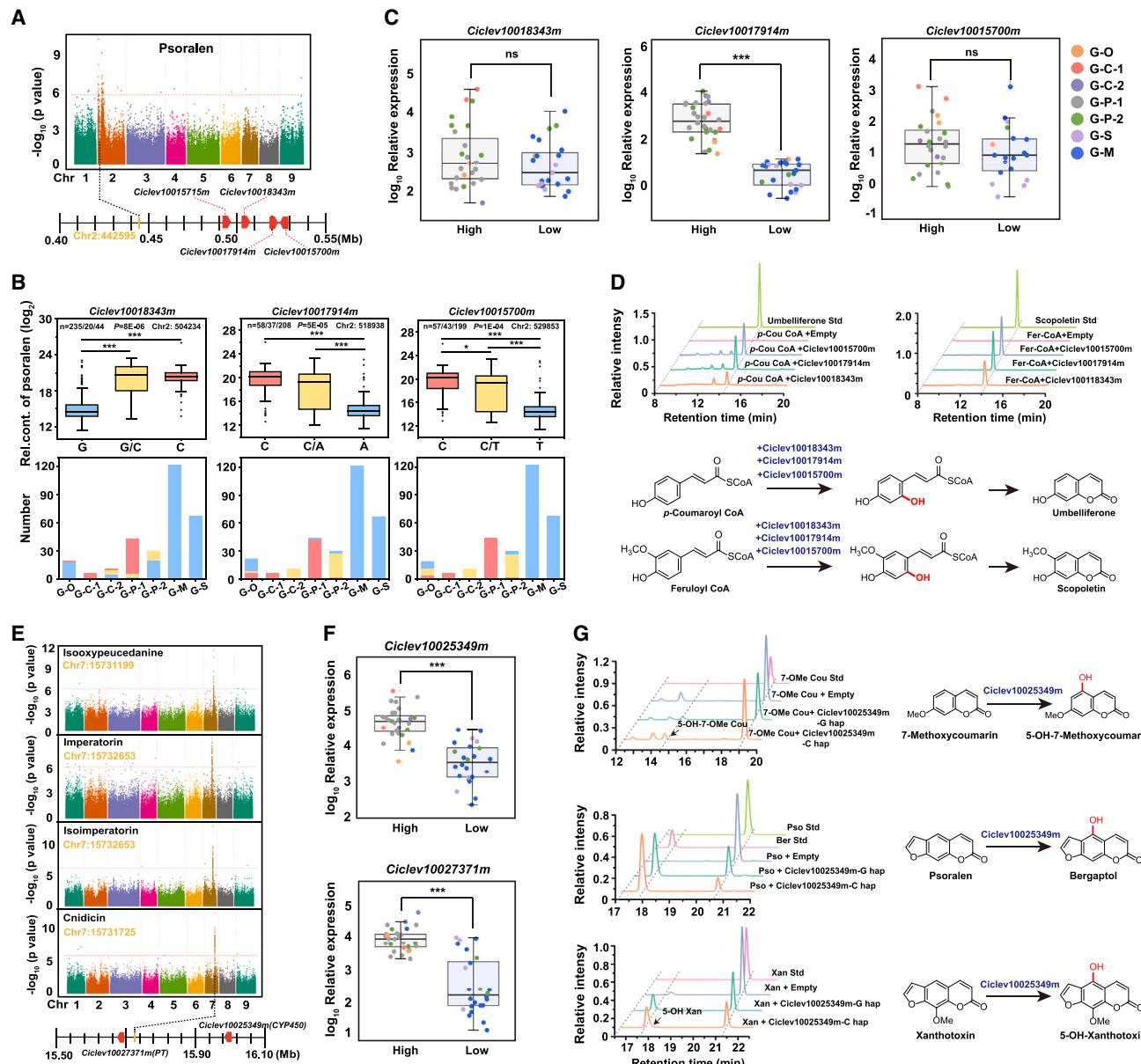


Figure 4. Identification and functional validation of candidate genes for coumarin accumulation.

(A) Identification of candidate *C2'hs* from the Manhattan plot of psoralen.

(B) Relationship between significant SNP types and psoralen content as well as the number of samples with different haplotypes in different groups. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(C) The expression levels of three candidate *C2'hs* in citrus accessions with high/low content of psoralen. Different colored dots represent accessions from different groups. The relative expression levels were normalized. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(D) *In vitro* enzyme assay of *Ciclev10018343m*, *Ciclev10017914m*, and *Ciclev10015700m*. *p*-Cou-CoA, *p*-coumaroyl-coenzyme A; Fer-CoA, feruloyl-coenzyme A.

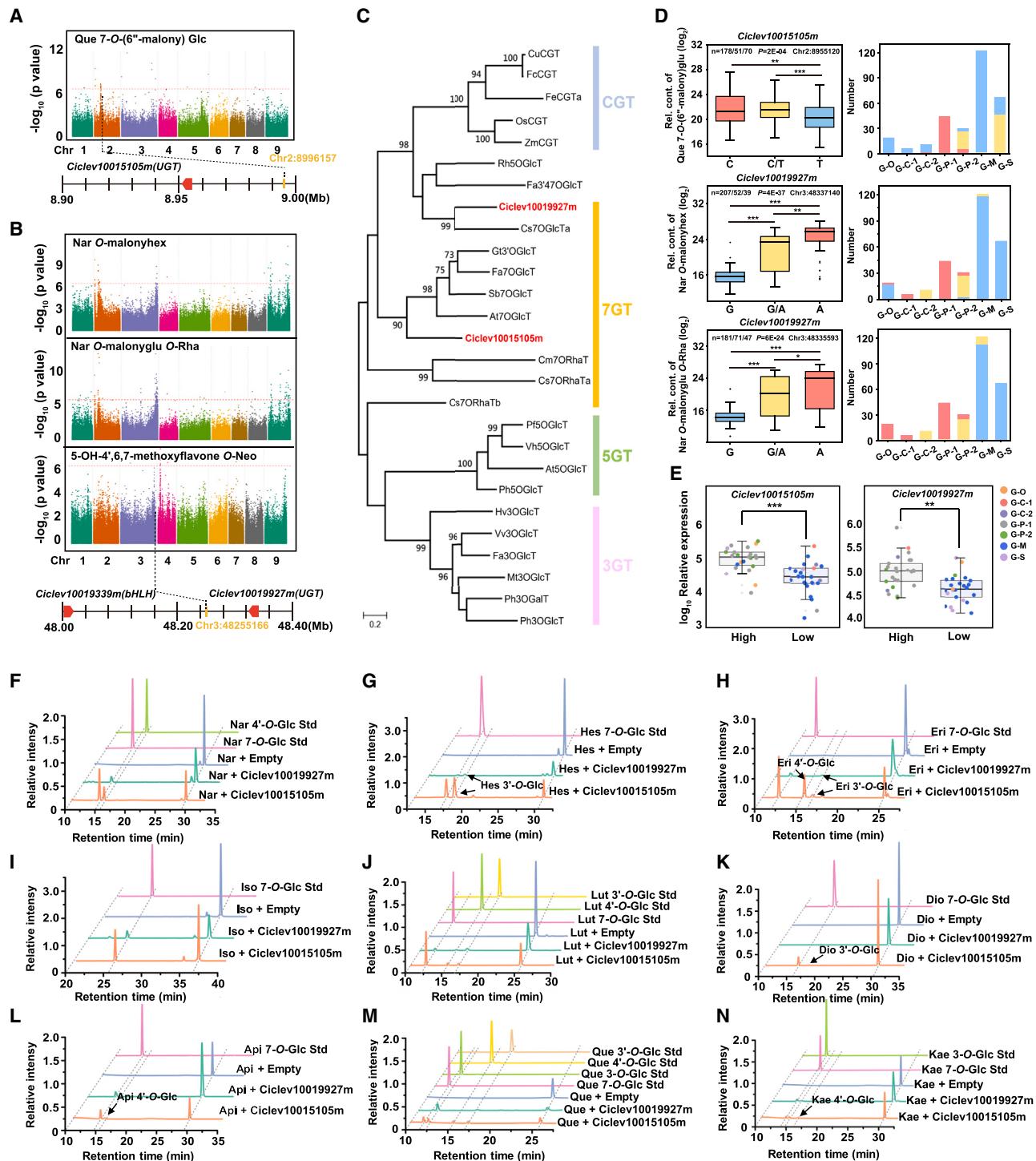
(E) Identification of candidate PT and CYP450 member from the Manhattan plots of isoxyppucedanine, imperatorin, isoimperatorin, and cnidicin.

(F) The expression levels of *Ciclev10027371m* and *Ciclev10025349m* in citrus accessions with high/low content of furanocoumarins. Different colored dots represent accessions from different groups, the same as in (C). The relative expression levels were normalized. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(G) Yeast heterologous expression of *Ciclev10025349m*. 7-O-Me Cou, 7-methoxycoumarin; Pso, psoralen; Ber, bergaptol; Xan, xanthotoxin.

furanocoumarins was observed (Figure 4F and Supplemental Figure 18). We applied yeast heterologous expression of both *Ciclev10025349m* alleles, and when taking xanthotoxin and 7-methoxycoumarin as substrates, both proteins exhibited catalytic activity to produce corresponding 5-OH products, as evidenced by the reported CYP82D64 (Limones-Mendez et al., 2020).

Besides, psoralen was also found as a new substrate of *Ciclev10025349m* to produce bergaptol (Figure 4G). The decreased expression of the two candidate genes during the developing stage correlated positively with coumarin accumulation (Supplemental Figure 14), underscoring their potential roles in orchestrating furanocoumarin synthesis.

**Figure 5. Identification and functional verification of candidate UGT genes.**

- (A) Identification of *Ciclev10015105m* from the Manhattan plot of quercetin 7-O-(6"-malonyl) glucoside (Que 7-O-(6"-malonyl) Glc).
- (B) Identification of *Ciclev10019927m* from the Manhattan plots of naringenin O-malonyhexoside (Nar O-malonyhex), naringenin O-malonylglycoside O-rhamnoside (Nar O-malonyglu O-Rha), and 5-hydroxy-4',6,7-trimethoxyflavone O-Neocarrabioside (5-OH-4',6,7-methoxyflavone O-Neo).
- (C) Phylogenetic tree for candidate UGT genes.
- (D) Relationship between significant SNP types and flavonoid glycoside contents as well as the number of samples with different haplotypes in different groups. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
- (E) Expression levels of candidate UGTs in accessions with high and low metabolite contents. For *Ciclev10015105m*, High/Low refers to the content of quercetin 7-O-(6"-malonyl) glucoside. For *Ciclev10019927m*, High/Low refers to the contents of naringenin O-malonyhexoside and naringenin

(legend continued on next page)

Identification and functional characterization of UGTs for flavonoid synthesis

Glycosylation can increase the solubility and chemical stability of flavonoids, which is crucial for their storage and transport *in vivo*, thus influencing their bioactivities (Plaza et al., 2014). Based on the significant signal of quercetin 7-O-(6"-malonyl)glucoside content on Chr2 and the co-located significant loci associated with the content of naringenin O-malonyhexoside, naringenin O-malonyglucoside O-rhamnoside, and 5-hydroxy-4',6,7-trimethoxyflavone O-neocarrabioside on Chr3, we identified two candidate flavonoid *UGTs*, *Ciclev10015105m* and *Ciclev10019927m* (Figures 5A and 5B). The previous phylogenetic tree classifies glycosyltransferases (GTs) into CGTs, 7GTs, 5GTs, and 3GTs based on their specific catalytic functions (Peng et al., 2017); *Ciclev10015105m* and *Ciclev10019927m* are located in the clade of 7GTs (Figure 5C). Significant SNPs located in the upstream or exon of *Ciclev10015105m* and *Ciclev10019927m* were significantly correlated with the content of corresponding flavonoid glycosides in different citrus groups, in which G-P-1 mostly contains haplotypes with high metabolite contents (Figure 5D and Supplemental Figure 19A; Supplemental Table 13). Nucleotide diversity analysis showed that the gene and promoter region of *Ciclev10015105m* had undergone selection in G-P-1 compared with G-M (Supplemental Figure 19B). Correspondingly, the two candidate *UGTs* exhibited significantly difference expression levels among citrus groups with different flavonoid glycoside contents. (Figure 5E). The expression levels of *Ciclev10015105m* were positively correlated with the content of quercetin 7-O-(6"-malonyl) glucoside, while the expression levels of *Ciclev10019927m* were positively correlated with the content of naringenin O-malonyhexoside and naringenin O-malonyglucoside O-rhamnoside (Supplemental Figure 20).

There is a basic-helix-loop-helix (bHLH) transcription factor (*Ciclev10019339m*) located in 270 kb upstream of the candidate *UGT* (*Ciclev10019927m*) (Figure 5B), which is close to the reported bHLHs involved in flavonoid synthesis regulation in the phylogenetic tree (Supplemental Figure 21A). Different alleles of *Ciclev10019339m* are significantly associated with the content of two flavonoid glycosides (Supplemental Figure 21B). There is a positive correlation between the expression level of *Ciclev10019339m* and flavonoid glycoside content (Supplemental Figure 21C), and co-expression is also observed in *Ciclev10019339m* and *Ciclev10019927m* among citrus varieties (Supplemental Figure 21D). Transient overexpression of *Ciclev10019339m* in the peel of *C. sinensis* 'Bingtangcheng' significantly increased total flavonoid glycoside content, indicating that it may be a potential bHLH transcription factor associated with the regulation of flavonoid synthesis in citrus (Supplemental Figure 21E).

UDP-glucose was used as a donor, and nine typical flavonoid aglycones—four flavanones (naringenin, hesperetin, eriodictyol,

Multi-omics analyses on citrus bioactive compounds

and isosakuranetin), three flavones (apigenin, luteolin, and diosmetin) and two flavonols (quercetin and kaempferol)—were tested as potential substrates for enzymatic reaction. The results showed that the recombinant protein of *Ciclev10015105m* catalyzed 7-O-glucosylation of all nine substrates, while 4'-O-glucosylation and 3'-O-glucosylation could also be performed with six and five flavonoids as substrates, respectively, and 3-O-glucosylation could be performed with two flavonol substrates (Figure 5F–5N; Supplemental Table 14). Similarly, the recombinant protein of *Ciclev10019927m* could carry out 7-O-glucosylation and 3'-O-glucosylation of eight and four flavonoid substrates, respectively, as well as 3'-O-glucosylation of two flavonol substrates, but it could only perform 4'-O-glucosylation on luteolin (Figure 5F–5N; Supplemental Table 14). The two candidate genes displayed multi-functional UGTs due to their diverse substrates and catalytic sites, and with the same substrate concentration, the peaks of 7-O-glycosylation products produced by the two enzymes were higher than those of other products, showcasing their dominant role in 7-O-glycosylation catalytic activity among other possible activities. We overexpressed the two UGTs in the peel of *C. sinensis* 'Bingtangcheng', and it significantly increased the total content of flavonoid glycosides (Supplemental Figure 22A), while the overexpression of *Ciclev10015105m* also increased the content of quercetin 7-O-(6"-malonyl) glucoside (Supplemental Figure 22B).

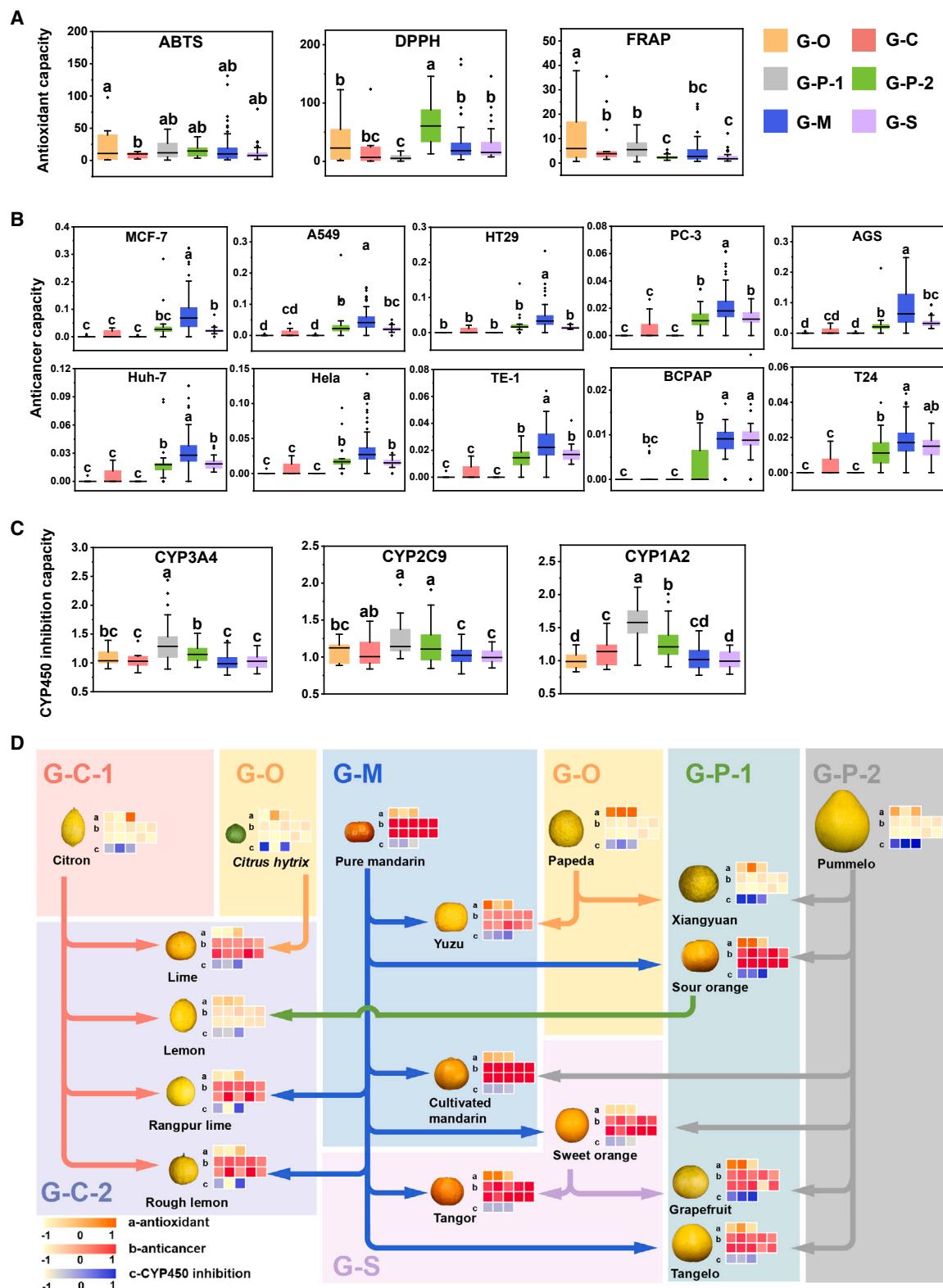
Metabolomic basis of bioactive functions of citrus

The composition of metabolites in plants is the basis of their biological activities. For different citrus varieties, the variation in the accumulation of metabolites often leads to a difference in biological activity (Wang et al., 2017b; Cao and Pan, 2022). We embarked on an exploration of the bioactive functionalities exhibited by a diverse array of citrus varieties and used extracts from the flavedo of 219 accessions to evaluate their bioactivities, including antioxidant activity, anticancer activity, and CYP450 enzyme inhibition (Supplemental Table 15). Three chemical assays, the 2,2'-azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) radical scavenging assay, 2,2-diphenyl-1-picrylhydrazyl (DPPH) free radical scavenging assay, and ferric ion-reducing antioxidant power (FRAP) assay, were applied to evaluate the antioxidant capacity of citrus extracts. However, no group with outstanding comprehensive antioxidant capacity was found (Figure 6A). We sorted the antioxidant capacity value of citrus accessions evaluated by different chemical assays and selected the top 20 varieties with the strongest antioxidant capacity (Supplemental Table 16), and according to intersection of Venn plot, four varieties (TZ198, CQGJ11, NEW7, and NEW5) ranked as the top 20 under the evaluation of all three chemical assays, which can be considered as valuable germplasm resources with a strong antioxidant capacity overall (Supplemental Table 17; Supplemental Figure 23A).

Anticancer capacity of citrus extracts was evaluated by testing the inhibition level against 10 cancer cell lines that correspond

O-malonyglucoside O-rhamnoside. Different colored dots represent accessions from different groups. The relative expression levels were normalized.
p* < 0.05, *p* < 0.01, ****p* < 0.001.

(F–N) *In vitro* enzyme assay of *Ciclev10015105m* and *Ciclev10019927m* on different substrates. The detailed substrates and products are listed in Supplemental Table 14. Nar, naringenin; Hes, hesperetin; Eri, eriodictyol; Iso, isosakuranetin; Api, apigenin; Lut, luteolin; Dio, diosmetin; Que, quercetin; Kae, kaempferol; Glc, glucoside.

**Figure 6. Bioactivity evaluation of citrus populations.**(A–C) Boxplots of bioactivity evaluations of 219 citrus accessions. Different lowercase letters represent statistically significant differences ($p < 0.05$).

(A) Comparison of antioxidant capacity (represented by mg Trolox/g fresh weight) of citrus extracts from six groups.

(B) Comparison of anticancer capacity (represented by reciprocal IC_{50} of 10 cell lines) of citrus extracts from six groups.

(legend continued on next page)

to the 10 most prevalent cancer types based on the 2018 Global Burden of Cancer data released by the International Agency for Research on Cancer (Bray et al., 2020) (Supplemental Table 18). The reciprocal of the inhibitory concentration (IC_{50}) was used as the evaluation indicator for inhibitory effect; the higher the value, the stronger the inhibitory capacity. Significant differences in the anticancer abilities were observed within different groups as follows: G-M > G-S ≈ G-P-2 > G-P-1 = G-C = G-O (due to the small amount of accessions, we used G-C for bioactivity evaluation rather than G-C-1 and G-C-2) (Figure 6B). A total of 43 accessions, primarily from G-M, reached the top 20 in their inhibitory ability against at least 2 cancer cell lines (Supplemental Table 19). Notably, various red tangerine (called hongju in Chinese) accessions, such as CQGJ104, TZ50, CQGJ68, CQGJ32, TZ48, and TZ56, exhibited a strong and extensive inhibitory effect against more than five cancer cell lines (Supplemental Table 19).

In addition to the beneficial bioactivities, such as antioxidant and anticancer activities, there are potential adverse effects of citrus bioactive compounds, notably furanocoumarins' interaction with specific drugs due to their inhibition of the CYP450 enzymes in the liver (Bailey et al., 2013; Petric et al., 2020). We evaluated the inhibitory ability of citrus extracts on three significant CYP450 enzymes in the human body (CYP3A4, CYP2C9, and CYP1A2), using the reciprocal of relative expression level of corresponding genes as the evaluation indicator; the higher the value, the stronger the inhibitory ability. Notably, compared with other groups, G-P-1 and G-P-2 exhibited significant CYP450 enzyme-inhibitory ability (Figure 6C).

The differences in bioactivities between groups and individuals are often closely related to the variations in metabolites. We conducted a correlation analysis between metabolite content and bioactivities and selected the top 100 annotated metabolites with the highest correlation coefficient (Pearson's r) with different bioactivities under the evaluation of each chemical assay/cell line/CYP450 enzyme. For antioxidant capability, most of the top 100 metabolites were secondary metabolites, of which flavonoids, especially flavonoid glycosides, dominated the top-ranked metabolites (Supplemental Figure 24A; Supplemental Table 20), and 18 common metabolites were found in the top 100 metabolites in the evaluation by three chemical assays, of which 13 were flavonoid glycosides (Supplemental Figure 23B; Supplemental Table 21). Narirutin exhibited the highest correlation with antioxidant ability (r average = 0.40) (Table 1). For anticancer capacity, the majority of the top 100 metabolites were secondary metabolites, with flavonoids occupying the highest proportion (Supplemental Figure 24B; Supplemental Table 22). A total of 51 common metabolites were found in the top 100 metabolites under the evaluation by 10 cell lines, among which 43 were flavonoids, including 20 PMFs and 17 flavonoid glycosides (Supplemental Figure 23C; Supplemental Table 23). There were 9 PMFs among the top 10 metabolites with the

Multi-omics analyses on citrus bioactive compounds

highest correlation with anticancer capacity, and nobiletin showed the highest correlation (r average = 0.76) (Table 1). For CYP450 enzyme inhibition capability, lignans and coumarins showed the largest proportion among the top 100 metabolites (Supplemental Figure 24C; Supplemental Table 24). A total of 53 common metabolites were found in the top 100 metabolites under the evaluation by 3 CYP450 enzymes, of which 24 were lignans and coumarins (Supplemental Figure 23D; Supplemental Table 25). There were 8 coumarins in the top 10 metabolites with the highest correlation with CYP450 enzyme inhibition capability, and bergaptol, a typical furanocoumarin in citrus, displayed the highest correlation (r average = 0.56) (Table 1).

We integrated the results of bioactivity evaluation in heatmaps into the evolutionary roadmap of citrus populations (Figure 6D), and the indicators of three bioactivities (antioxidant activity, anticancer activity, and CYP450 enzyme inhibition) were represented by a, b, and c, respectively. Our analysis highlighted that pure mandarin, cultivated mandarin, tangors in G-M, and sweet oranges in G-S exhibited robust antioxidant and anticancer abilities without a significant inhibitory effect on CYP450 enzymes, making them the first choice for the advancement of functional medicinal citrus through selective breeding. Some descendants of mandarins, such as yuzu, rangpur lime, and rough lemon, also had excellent anticancer properties. Although grapefruits and sour oranges demonstrated strong antioxidant and anticancer properties due to the genetic contribution from pummelo, they also have a strong inhibitory effect on CYP450 enzymes. Thus, their utilization as sources for functional foods or medicinal applications warrants careful consideration. Furthermore, the strong antioxidant capacity of papaya may focus our attention on the utilization of this old early-diverging citrus resource.

DISCUSSION

Traditional crop breeding is often centered on improving yield and flavor qualities. In recent years, due to the increasing awareness of human nutrition and health, plant-derived bioactive compounds have attracted attention, and more research on the mining of genetic loci related to bioactive functions and bioactive compounds is emerging (Li et al., 2022; Shen et al., 2023; Zhao et al., 2024). The release of the citrus genome and its origin and evolution accelerates the genomic breeding of this economically important fruit crop worldwide. Wu et al. (2018) analyzed the genomes of 60 citrus and elucidated their evolutionary trajectory, and Huang et al. (2023) elucidated the systematic evolution of the *Citrus* genus and its related genera through pangenome analysis of *Citrus* subfamilies. Expanding upon this foundation, we examined two papaya hybrids, xiangyuan (*C. wilsonii*) and yuzu (*C. junos*) (Shimizu et al., 2016; Yun et al., 2024), which are widely used as medicines and drinks in Asia (Jeong Ah et al., 2019; Yan et al., 2021). In the phylogenetic tree, they were located in the same clade as their pummelo or mandarin parents, respectively, but were distant

(C) Comparison of CYP450 enzyme inhibition capacity (represented by reciprocal relative gene expression levels of three CYP450 enzymes) of citrus extracts from six groups.

(D) Bioactivity variation map of the citrus evolution. The orange (a), red (b), and blue (c) blocks represent the strength of antioxidant activity (evaluated by ABTS, DPPH, and FRAP assay), anticancer activity (evaluated by 10 cancer cell lines; first row, MCF-7, A549, HT29, PC-3, and AGS; second row, Huh-7, HeLa, TE-1, BCPAP, and T24), and CYP450 enzyme inhibition (evaluated by CYP3A4, CYP2C9, and CYP1A2), respectively. The darker shade indicates heightened bioactivity levels (only for comparison between different citrus groups). Arrows with the same color suggest the same genetic origin.

Bioactivity	ID	Order	Compound	Class	Pearson's r average
Antioxidant	TGR0006	1	naringenin 7-O-rutinoside (narirutin)	flavonoids	0.40
	TGR2187	2	naringenin 7-O-glucoside (prunin)	flavonoids	0.40
	TGR2099	3	limocitrin	flavonoids	0.39
	TGR2176	4	8-hydroxy-3,5,6,7,3',4'-hexamethoxyflavone	flavonoids	0.39
	TGR0238	5	syringetin	flavonoids	0.39
	TGR2100	6	spinacetin	flavonoids	0.38
	TGR2179	7	5-hydroxy-3,6,7,8,3',4'-hexamethoxyflavone	flavonoids	0.37
	TGR2178	8	monohydroxy-hexamethoxyflavone	flavonoids	0.36
	TGR2177	9	8-hydroxy-5,6,7,3',4',5'-hexamethoxyflavone	flavonoids	0.34
	TGR2225	10	isorhamnetin 5-O-hexoside	flavonoids	0.34
Anticancer	TGR2170	1	nobiletin (5,6,7,8,3',4'-hexamethoxyflavone)	flavonoids	0.76
	TGR2167	2	3,5,7,8,3',4'-hexamethoxyflavone	flavonoids	0.75
	TGR2145	3	tangeretin	flavonoids	0.74
	TGR2144	4	3,5,6,7,4'-pentamethoxyflavone	flavonoids	0.74
	TGR2153	5	5,8-dihydroxy-3,3',4',7-tetramethoxyflavone	flavonoids	0.74
	TGR2161	6	4'-hydroxy-3',5,6,7,8-pentamethoxyflavone	flavonoids	0.71
	TGR2173	7	5,4'-dihydroxy-3,6,7,8,3'-pentamethoxyflavone	flavonoids	0.71
	TGR2146	8	3',4',5',7,8-pentamethoxyflavone	flavonoids	0.71
	TGR2091	9	5,7,8,4'-tetramethoxyflavone	flavonoids	0.70
	TGR0218	10	pterostilbene	others	0.70
CYP450 inhibition	TGR0086	1	bergaptol	lignans and coumarins	0.56
	TGR1994	2	meranzin	lignans and coumarins	0.51
	TGR2014	3	crucigasterin E	others	0.49
	TGR1989	4	7-methoxy-5-prenyloxycoumarin	lignans and coumarins	0.47
	TGR1993	5	isomeranzin	lignans and coumarins	0.47
	TGR1981	6	suberosin	lignans and coumarins	0.46
	TGR1990	7	auraptene	lignans and coumarins	0.45
	TGR1995	8	3-(1,1-dimethylallyl) scopoletin	lignans and coumarins	0.45
	TGR2747	9	epoxybergamottin	lignans and coumarins	0.44
	TGR1991	10	3,4-diallyloxy cinnamic acid	others	0.43

Table 1. The top10 metabolites showing the highest correlation coefficient (Pearson's r average) with the capacity of antioxidant/anticancer/CYP450 inhibition.

Statistical significance for the correlation coefficient is provided in [Supplemental Table 26](#).

from papeda ([Figure 1A](#)). We clarified the ambiguous positions of these two hybrids in phylogenomic relationships. In addition, we provided new insights into the origin of some modern hybrids. The origin of Huyou (*C. changshanensis*), a special landrace in the Zhejiang Province of China, has long been disputed ([Xu et al., 2006](#)). Our results show that the Huyou accessions (TZ71, TZ100, and TZ101) are not clustered together with typical grapefruit accessions but close to sour oranges, and there is little sweet orange genetic composition shown in the population structure ([Figure 1A](#)). Therefore, Huyou should be considered a hybrid admixture of pummelo and mandarin approximate to tangelo or sour orange rather than grapefruit.

Differentiation and selection often reshape the metabolome in crops, resulting in changes in related metabolites such as flavor, resistance, and biological activity ([Zeng et al., 2020](#); [Yu et al.,](#)

[2021; Cao et al., 2022](#)). Previous studies showed that PMFs in citrus mainly come from the ancestors of mandarin ([Peng et al., 2021](#)), while coumarins originate from pummelo and citron ([Dugrand-Judek et al., 2015](#)). In this study, our results pinpoint that a large number of core structural genes and downstream modification genes appeared in the differentiation sweep regions of three ancestral species, indicating that metabolic profile variation may be influenced by multiple genes in the pathway ([Figure 2B](#) and [Supplemental Figure 7](#)). [Zhang et al. \(2024\)](#) found that a taxon-specific distribution of allelic variations of the *CitCYP97B* promoter may cause the differential accumulation of β-cryptoxanthin in citrus, and *Fst* at the *CitCYP97B* locus is high in different citrus groups, indicating that species divergence may contribute to shaping β-cryptoxanthin diversity. Similarly, we also found two main haplotypes in the promoters of several structural genes under differentiation, such

as FNSs, C2'H2, and CYP82D92, which are closely related to differential gene expression level ([Supplemental Figure 8](#)). We deduce that it is the probable genetic basis for the remarkable variation of flavones/coumarins accumulation during citrus differentiation.

Notably, the presence of C2'H genes in the differentiation sweeps was observed ([Figure 2B](#) and [Supplemental Figure 7](#)), and further identification of these C2'HS was accomplished through mGWAS ([Figure 4A](#)). Initially, researchers attempted to verify the catalytic function of citrus C2'H but failed to find the resulting product ([Vialart et al., 2012](#)). Fortunately, we extended the reaction time and successfully obtained the expected products, which indicates that the catalytic efficiency of citrus C2'HS may not be high enough. Consequently, we speculate that the tandem duplication of C2'H genes in citrus might contribute to sufficient metabolite accumulation. In addition, we consider *Ciclev10017914m* to be an important C2'H responsible for coumarin content diversity in citrus because it is specifically highly expressed in accessions with high coumarin accumulation ([Figure 4C](#)), and the expression levels during citrus development stages are significantly correlated with coumarin content ([Supplemental Figure 14](#)). *Ciclev10018343m* and *Ciclev10015700m*, even if they possess similar C2'H catalytic activity, may not be the main effector genes in citrus based on their expression characteristics. In our study, we also found a *PT* gene and a *CYP450* member that potentially contribute to the synthesis of furanocoumarins. Their haplotypes and their expression levels aligned well with furanocoumarin contents ([Figure 4F](#) and [Supplemental Figures 16 and 18](#)). Therefore, the specific accumulation of coumarins may indeed be an outcome of the combined action of multiple genes in the pathway. In addition, we noted that non-synonymous SNPs in the coding region of candidate genes do not seem to affect the activity of the protein ([Figure 4G](#) and [Supplemental Figure 12](#)), so we suspect that polymorphisms in the promoter region and differences in transcription levels are more likely to be responsible for the differential accumulation of coumarin among citrus populations.

Flavonoid glycosides are prevalent among various plant species. Research indicates that, during the domestication and improvement of crops such as rice ([Peng et al., 2017](#)), barley ([Zeng et al., 2020](#)), Tartary buckwheat ([Zhao et al., 2023](#)), and others, the selection of glycosyltransferase related to flavonoid synthesis has led to the accumulation of flavonoid glycosides. This has significantly improved the biological activity and adaptability of these crops to the environment. In citrus peel, flavonoids are predominantly present in glycoside form ([Tripoli et al., 2007](#)). In this study, we identified two genes belonging to the UGT family based on the signal of several flavonoid glycosides and subsequently confirmed their catalytic functionality ([Figure 5](#)). At present, UGTs with 7-O-glycosylation, 3-O-glycosylation, and C-glycosylation capabilities have been reported in citrus ([Owens and McIntosh, 2009](#); [Frydman et al., 2013](#); [Ito et al., 2017](#)). In this study, we identified two multifunctional UGT genes that conduct O-glycosylation at positions 3, 7, 3', and 4' on flavanones, flavones and flavonols, which expanded the functional diversity of citrus UGTs. A study by [Cao et al. \(2021\)](#) suggested that there was a strong correlation between various flavonoid glycosides and the

Multi-omics analyses on citrus bioactive compounds

antioxidant capacity of peaches. Similarly, through correlation analysis, we established a close association between flavonoid glycosides and the antioxidant capacity of citrus. Therefore, we speculate that these two multifunctional candidate UGTs might serve as potential target genes for significantly enhancing the functional compounds through precise breeding of citrus crops. This approach could assist in augmenting flavonoid glycosides, thereby fully capitalizing on the antioxidant potential of citrus.

Traditionally, citrus breeding has predominantly focused on enhancing superficial attributes such as appearance, flavor, and resistance ([Gmitter et al., 2014](#); [Sun et al., 2019](#)). However, current functional and medicinal citrus varieties mostly result from empirical breeding, relying on non-targeted selection strategies. Research has begun to focus on the medicinal functions of citrus and their genetic basis. [Zheng et al. \(2023\)](#) analyzed the anti-inflammatory function of two flavonoid glycosides in pummelos and identified *CmtMYB108*, a transcription factor that regulates flavonoid synthesis, by multi-omics methods. [Shen et al. \(2023\)](#) clarified the biosynthesis pathway of melitidin, an anti-cholesterol statin drug candidate in pummelo through mGWAS, providing support for the breeding and improvement of medicinal pummelo cultivars ([Shen et al., 2023](#)). On this basis, we extend the research to the varieties of the *Citrus* genus and focus on more kinds of bioactivities, conducting a largest-scale evaluation and quantification of the bioactivities of citrus population ([Figures 6A–6C](#)). Our aim was to identify the dominant citrus varieties and metabolites that possess exceptional bioactive potential. In addition to antioxidant capacity, the anticancer efficacy of citrus has generated significant attention. Several constituents within citrus have been ascribed anticancer attributes ([Ejaz et al., 2006](#); [Gansukh et al., 2019](#); [Goh et al., 2019](#); [Li et al., 2009](#)). Our correlation analyses underscored the pivotal role of PMFs, emphasizing their paramount importance in exerting anticancer effects. Notably, the red tangerines, widely cultivated in China as local varieties, can be considered compelling candidates for potential exploitation in anticancer research and therapeutic applications.

From the perspective of food safety, we also explore the negative biological activity of citrus metabolites. Our findings support the inhibitory influence of coumarins, particularly furanocoumarins, on CYP450 enzymes, aligning with previous studies ([Ho et al., 2001](#); [Petric et al., 2020](#)). Although we used citrus peel (flavedo) for metabolome analysis and bioactive evaluation, the accumulation characteristics of furanocoumarins in the peel and pulp of different citrus groups were quite similar. [Dugrand-Judek et al. \(2015\)](#) compared the coumarin distribution in different citrus species and found that pummelos and some of their descendants also synthesize coumarins and furanocoumarins rich in the peel and pulp, whereas mandarin, mandarin hybrids, and sweet oranges all display low amounts of coumarins in both peel and pulp. Therefore, for citrus species with a high content of furanocoumarins in the peel, the consumption of their flesh is associated with the risk of drug interaction. Therefore, it becomes imperative to reduce furanocoumarin accumulation in pummelo or pummelo hybrids via omics-aided breeding approaches in modern citrus improvement. Our findings suggest the potentiality of G-M and G-S populations as genetic reservoirs for breeding functional healthy

cultivars owing to their desirable alleles for high flavone and low coumarin accumulation. In contrast, G-P-1 and G-P-2 should be approached cautiously as potential genetic donors due to their less favorable attributes. Taken together, we suggest a possible bioactive compounds profile for nutritious and functional citrus varieties with optimal accumulation of flavonoids and coumarins, and provide a breeding pathway through pyramiding beneficial haplotypes of flavonoids and coumarin synthesis-related genes.

In conclusion, this study unveiled the genomic and metabolomic foundations responsible for the differential accumulation of bioactive compounds during the genetic differentiation in citrus as well as clarified the positive role of flavonoid glycosides and PMFs in antioxidant and anticancer activity, respectively, and the negative inhibitory function of coumarin on CYP450 enzymes. These findings offer a genomic and metabolomic framework for strategically enhancing beneficial bioactive compounds in the pursuit of breeding modern citrus with improved nutritional ingredients by stacking the favorable haplotypes related to flavonoid and coumarin synthesis. This advancement in genetic improvement of health-promoting fruit crops in citrus will contribute to the practice of the notion of One Health.

METHODS

Plant materials

All citrus accessions were collected from the Citrus Research Institute of Zhejiang Province (Taizhou, China) and Citrus Research Institute of Chinese Academy of Agricultural Sciences (Chongqing, China) in 2019 and 2020 for metabolome determination. All samples had reached commercial maturity (about 220–250 days after flowering), and healthy fruits of similar size were randomly picked from different locations of the tree. The detailed information of all cultivars is listed in *Supplemental Table 1*. *C. grandis* ‘Jiangxi Zaoyou’ was collected from the Citrus Research Institute of Chinese Academy of Agricultural Sciences at seven developmental stages: S1–S7. The fruits of each accession were separated into two parts, flavedo and albedo, and flavedo parts were frozen immediately in liquid nitrogen. The samples were stored at –80°C for subsequent analysis.

Whole-genome resequencing and SNP identification

Young leaves of citrus accessions were collected for genomic DNA extraction using the EZ-10 Spin Column Plant Genomic DNA Purification Kit (Sangon Biotech, Shanghai, China). A small-fragment DNA library of ~200–300 bp was constructed for each cultivar and sequenced using the BGISEQ-500/MGISEQ-2000 platform. The raw reads were filtered using SOAPnuke 1.5.6 (Chen et al., 2018b) to obtain clean reads, and the clean reads were then mapped to the reference genome *Citrus clementina* v.1.0 (<https://www.ncbi.nlm.nih.gov/genome/?term=Citrus+clementina>) using BWA (v. 0.7.17-r1188) (Li and Durbin, 2009, 2010). The mapping results were sorted and filtered using Samtools (v.1.9), and reads with mapping quality (MQ) > 30 were used for later analysis (Li and Durbin, 2009). Qualimap2 (v.2.2.2-dev) was used for quality control statistics (Okonechnikov et al., 2016). Furthermore, the UnifiedGenotyper function of GATK was used for SNP calling of all samples. Then the results were further filtered, and the parameters are as follows: SNP: QD < 2.0||FS>60.0||MQ<40.0||MQRankSum<−12.5||ReadPosRankSum<−8.0. Finally, the annotation of all identified SNPs was performed using ANNOVAR.

Population genetic analysis and detection of differentiation sweeps

PCA of whole-genome SNPs was performed by TASSEL (5.2.50) with default parameters and visualized using the R package ggplot2. The

phylogenetic tree was constructed by SNPhylo (v.20180901) software and then displayed and annotated by iTOL. Population structure of the accessions was performed using ADMIXTURE (v.1.23). In order to avoid the impacts caused by admixed individuals, the accessions with over 60% genetic components (Q value in the population structure) from G-M, G-P-1, or G-C-1 were selected to perform an analysis of differentiation sweeps (Kang et al., 2021). The *Fst* and nucleotide diversity (π) values were calculated using VCFtools software (v. 0.1.12b) in 40-kb windows with a step size of 20-kb, and windows with *Fst* values exceeding top 5% were regarded as highly differentiated regions. The XP-CLR value was calculated by XP-CLR (v.1.0) software in 40-kb windows with a step size of 20 kb, and the top 5% regions with the highest values were defined as candidate differentiation sweeps.

Genome origin analysis of sweet oranges and grapefruits

There were 30 mandarins, 30 pummelos, 30 sweet oranges, and 4 grapefruits chosen to perform the origin analysis, and the accessions in each population maintained a consistent genetic background (the influence of admixture individuals was removed). Homozygous SNPs shared by over 70% of individuals in each population were selected as the characteristic SNPs of the population. SNP density preferences for sweet oranges and grapefruits were determined by sliding window method according to Xu et al. (2013): counting the number of SNPs from mandarin or pummelo in a 100-kb window with a step size of 10 kb (windows with more than 30 SNPs were used for analysis) and calculating the ratio of mandarin SNPs/pummelo SNPs. If the ratio is ≥ 4 , then the region is thought to originate from mandarin; if it is ≤ 0.25 , then the region is thought to originate from pummelo. In other cases, the origin cannot be determined.

Metabolite profiling

The citrus flavedo samples were vacuum freeze dried and then ground (30 Hz, 1 min) to powder using a mixer mill (MM 400, Retsch). 60–80 mg powder was weighed and dissolved in 70% methanol and then vortexed violently for ~5–6 times to mix it well. After being stored at 4°C overnight, the extract was ultracentrifuged at 12 000 rpm for 10 min, and the supernatants were pooled and filtered through a 0.22-μm membrane prior for LC-MS/MS analysis. The extracts were analyzed using an LC–electrospray ionization (ESI)–MS/MS system (high performance liquid chromatography [HPLC], Shim-pack ultra-fast liquid chromatography [UFLC] Shimadzu CBM30A system; MS, Applied Biosystems 6500 plus Q TRAP). Samples were separated by Waters ACQUITY UPLC HSS T3 C18 column (1.8 μm, 2.1 mm × 100 mm). The mobile phase was water (A) and acetonitrile (B), both containing 0.04% (v/v) acetic acid. The linear gradient program was as follows at a flow rate of 0.35 ml min^{−1}: 0 min, 95% of A; ~0–11 min, ~95%–5% of A; 12 min, 5% of A; 12.1 min, 95% of A; ~12–15 min, 95% of A. The injection volume was 2 μl, and the column temperature was set at 40°C.

Linear Ion Trap (LIT) and triple quadrupole (QQQ) scans were acquired on a QQQ-linear ion trap mass spectrometer (Q TRAP) API 6500 Q TRAP LC/MS/MS system equipped with an ESI turbo ion-spray interface operating in positive ion mode and controlled by Analyst 1.6 software (AB Sciex, USA). The ESI source operation parameters were as follows: ion source, turbo spray; source temperature, 500°C; ion spray voltage, (+) 5500 V and (−) 4500 V; ion source gas I, gas II, and curtain gas were set at 55, 60, and 35.0 psi, respectively; the collision gas was medium. Instrument tuning and mass calibration were performed with 10 and 100 μmol/l polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired as multiple reaction monitoring (MRM) experiments with collision gas (nitrogen) set to 5 psi. Declustering potential (DP) and collision energy (CE) for individual MRM transitions were done with further DP and CE optimization (Chen et al., 2013). A specific set of MRM transitions was monitored for each period according to the metabolites eluted within this period.

The signal peaks detected in different samples of each metabolite were corrected by MultiQuant software according to the information on the retention time and peak shape of the metabolites (Fraga et al., 2010). After peak shape correction and metabolite filtering, the peak area of the metabolite quantification was taken as the relative content of the substance.

Statistical analysis of metabolomics data

The CV was used to analyze the variation of metabolite content in the citrus population. The value of CV for each metabolite was calculated independently with the formula σ/μ , where σ and μ represent the standard deviation and mean, respectively. The heritability (H^2) of all the metabolites was estimated using GEMMA, and the calculation formula is as follows: $H^2 = V_G/(V_G + V_E)$, where V_G and V_E represent the genetic and environmental variances, respectively. PCA of metabolomics data was performed by TASSEL (5.2.50) with default parameter and visualized using the R package ggplot2. Clustering heatmaps were performed using the Metware Cloud, a free online platform for data analysis (<https://cloud.metware.cn>).

mGWAS

A total of 423 330 high-quality biallelic SNPs with minor-allele frequency ≥ 0.05 and missing call frequency ≤ 0.01 were used to perform mGWAS analysis. In this study, we chose a mixed linear model with population structure and kinship as fixed effects and inter-individual kinship as random effects. The population structure matrix was generated by ADMIXTURE, and the kinship matrix was constructed using GEMMA. The association analysis was performed by GEMMA. A total of 57 309 SNPs were finally obtained after removing those that were in linkage disequilibrium using PLINK. The genome-wide significance threshold ($1.24e-06$) was determined after Bonferroni correction. Annotation of significant SNPs after deduplication was performed by ANNOVAR.

Phylogenetic analysis

The amino acid sequences of genes involved in phylogenetic analysis were downloaded from NCBI (<http://www.ncbi.nlm.nih.gov>) according to accession numbers. MEGA7 equipped with ClusterW was used to perform sequence comparison, and the neighbor-joining method was adopted to build the phylogenetic tree with 1000 bootstrap replicates.

Gene cloning and recombinant protein purification

Total RNA was extracted from the flavedo of citrus fruit by the method of Kiefer et al. (2000). Then the RNA was reverse transcribed by a PrimeScript RT Reagent Kit (Takara, Dalian, China) to construct a cDNA library. Primers were designed to amplify the full CDS of candidate genes from the cDNA obtained in the previous step. The CDSs were subcloned into the pET32a vector (without a terminator), and afterward, the recombinant plasmids were transformed into *E. coli* strain BL21(DE3) for expression. The bacteria were cultured in Luria-Bertani (LB) medium (containing 0.1 g/l ampicillin) at 37°C until the optical density 600 (OD₆₀₀) reached 0.8, then the recombinant protein synthesis was induced by 1 mM isopropyl-β-D-thiogalactopyranoside for 24 h at 16°C. The bacterial cells were collected by centrifugation (4000 g, 4°C for 15 min) and resuspended in 1× PBS. After freezing at -80°C for more than 24 h, the cells were disrupted with a sonicator, and the supernatant was collected by centrifugation (10 000 g, 4°C for 30 min). The recombinant proteins were purified using a HisTALON Gravity Column Purification Kit (Takara) and eluted by a storage buffer (0.1 M Tris-HCl, pH 6.5, containing 2 mM DTT and 10% glycerol) through a PD-10 column. The purified recombinant protein was detected by SDS-PAGE and stored at -80°C for further analysis.

Enzyme assays

For the *in vitro* validation of C2'hs, the enzymatic reaction system includes 20 μl recombinant protein, 0.1 mM Tris-HCl (pH 6.5), 5 mM α-ketoglutaric acid disodium salt, 10 mM L-ascorbic acid sodium salt, 0.2 mM FeSO₄,

Multi-omics analyses on citrus bioactive compounds

and 200 μM substrate (including *p*-coumaroyl-CoA and feruloyl-CoA), and the total volume of the system was 200 μl. The reaction was performed overnight (about 12 h) at 30°C. For the *in vitro* validation of UGTs, the enzymatic reaction system includes 20 μl recombinant protein, 0.1 mM Tris-HCl (pH 7.5), 1 mM UDP-glucoside, and 1 mM substrate. The reaction was performed for 2 h at 30°C. Finally, the reaction was terminated, and the product was extracted twice by adding an equal volume of ethyl acetate. The upper organic phase was dried and re-dissolved in 100 μl chromatographic methanol. After centrifuging for 10 min at 12 000 rpm, the supernatant was analyzed by HPLC equipped with a diode array detector (DAD) on a Cosmosil 5C18-AR-II column.

Yeast heterologous expression

The CDS of *Ciclev10025349m* was subcloned into the pYES2/NT C vector (without a terminator), and the recombinant plasmids were transformed into WAT11-competent cells for expression. Positive monoclonal cultures with a diameter of about 2 mm were selected and cultured in 10 ml synthetic dextrose minimal medium without uracil (SD-Ura)+glucose liquid medium (30°C, 12–16 h) until OD₆₀₀ reached 0.6–1.2. The yeast solution was centrifuged at 4000 g for 5 min, and then the supernatant was discarded, and the yeast cells were re-suspended with 10 ml SD-Ura+galactose liquid medium. After induction culture at 30°C for 10 h, the substrate was added with a final concentration of 50 μM. After 6–8 h of enzyme catalysis, the reactions were terminated, and the product was extracted twice by adding an equal volume of ethyl acetate. The upper organic phase was dried and re-dissolved in 100 μl chromatographic methanol. After centrifuging for 10 min at 12 000 rpm, the supernatant was analyzed by HPLC-DAD on a Cosmosil 5C18-AR-II column.

Transient overexpression in citrus fruit

The CDSs of candidate genes were inserted into the destination vector pBI121, and the constructs were transferred into *Agrobacterium tumefaciens* strain EHA105 for overexpression. Eight or ten mature fruits were used for the study of each candidate gene. *Agrobacterium* suspensions containing empty vector and recombinant vector were injected on both sides of the equatorial plane of each fruit. After injection, the fruits were incubated in the dark at 28°C overnight to allow better penetration of *Agrobacterium*. Then, the fruits were left at room temperature for 5 days, and the peel of the injection area was sampled and snap frozen with liquid nitrogen. Contents of metabolites and relative expression levels of each candidate gene were detected in the sampling area.

Extraction and HPLC/MS analysis of flavonoids and coumarins

The samples of citrus peel were frozen in liquid nitrogen and crushed immediately. The compounds of 0.1 g fresh sample were extracted in 1 ml of 80% ethanol with sonication (ultrasound frequency, 60 kHz; power, 30 W) for 30 min. After repeating the extraction once, the supernatants were combined and concentrated on the Eppendorf Concentrator Plus (Germany), after which the residue was re-dissolved with 1 ml of chromatographic methanol. The extract was ultracentrifuged at 12 000 rpm for 15 min, and the supernatant was taken for HPLC detection.

HPLC analysis was performed according to the method described by Liu et al. (2020) with some modifications. All chromatographic experiments were carried out with a Waters 2695 liquid phase system (2998 DAD detector) coupled with an ODS C18 column (SunFire 5 μm, 4.6 × 250 mm, Waters, USA). The mobile phase was acetonitrile (A) and water containing 0.1% (v/v) formic acid (B), and the injection volume was 10 μl. The linear gradient program was as follows at a flow rate of 1 mL min⁻¹: 0–3 min, 20% of A; ~3–5 min, 20%–30% of A; ~5–10 min, 30% of A; ~10–15 min, 30%–40% of A; ~15–20 min, ~40%–60% of A; ~20–23 min, ~60%–80% of A; ~23–25 min, ~80%–100% of A; ~25–27 min, ~100%–20% of A; ~27–30 min, 20% of A. Different coumarins were detected by UV at 300–350 nm, and flavonoids were detected by UV at 280 nm. The column temperature was set at 25°C.

Multi-omics analyses on citrus bioactive compounds

Enzyme activity products were identified by MS using an AB Triple TOF 5600plus System (AB Sciex). The process was performed by total ion chromatography and ESI. The ion scan range (*m/z*) was 100–1500 Da, and the analysis was performed in negative ion mode or positive ion mode. The data were analyzed using PeakView1.2 software (AB Sciex).

RNA extraction and real-time qPCR

Total RNA was extracted using the cetyl trimethyl ammonium bromide (CTAB) method according to Kiefer et al. (2000), and three biological replicates were set up for each sample. A total of 1 µg RNA of each sample was synthesized to cDNA by PrimerScript RT Reagent Kit with gDNA Eraser (Takara), and the cDNA was diluted 10-fold as an RT-qPCR reaction template. Gene-specific primers were designed using the online program Primer BLAST (<https://www.ncbi.nlm.nih.gov/tools/primerblast/>) and verified by product sequencing. Citrus β-actin was used as an internal reference gene (Pillitteri et al., 2004), and $2^{-\Delta\Delta Ct}$ was used to analyze the expression levels of candidate genes.

Antioxidant capacity assay

The ABTS, DPPH and FRAP assays were employed to determine the antioxidant capacity of citrus peel extracts.

For the ABTS assay, to prepare the ABTS working solution, 50 mL of ABTS solution (7 mM) and 50 ml of potassium persulfate solution (2.6 mM) were mixed together. The mixture was allowed to react in the dark for 12 h at room temperature. After the reaction, the mixture was diluted approximately 28 times to achieve an OD at 734 nm of approximately 0.63. This diluted solution was used as the ABTS working solution. For the evaluation method, a mixture of 10 µl of appropriately diluted citrus extract solution and 200 µl of the ABTS working solution was prepared. The mixture was then reacted for 5 min at room temperature without light. After the reaction, the absorbance at 734 nm was measured using a microplate reader (Synergy H1, BioTek, VT, USA).

For the DPPH assay, to prepare the DPPH working solution, 2.4 mg DPPH powder was dissolved in 100 ml methanol for a final concentration of 60 µM. Then, 2 µl of appropriately diluted citrus extract solution was added to 198 µl of freshly prepared DPPH solution. The mixture was then allowed to react for 2 h in the dark at 25°C. After the reaction, the absorbance at 517 nm was measured using a microplate reader.

For the FRAP assay, to prepare the FRAP working solution, 100 ml of sodium acetate buffer (300 mM, pH 3.6), 10 ml of tripyridyltriazine solution (10 mM), and 10 ml of ferric chloride solution (20 mM) were combined and thoroughly mixed. For the evaluation method, 20 µl of appropriately diluted citrus extract solution was mixed with 180 µl of the FRAP working solution. The mixture was allowed to react for 5 min at room temperature in the dark. After the reaction, the absorbance at 593 nm was measured using a microplate reader.

Trolox was used as a control in the assays. The experiments were repeated three times, and the results were expressed as milligrams of Trolox equivalent per gram of fresh weight.

Cell culture assay

Cells were cultured in different media according to the protocol provided by the supplier. The culture medium for MCF-7 cells consisted of 90% Dulbecco's modified eagle medium (DMEM) supplemented with 0.01 mg/ml insulin and 10% fetal bovine serum (FBS). The culture medium for A549, PC-3, and AGS cells comprised 90% F-12K medium and 10% FBS. HepG2 and Huh-7 cells were cultured in a medium consisting of 90% DMEM and 10% FBS. HeLa cells were maintained in a medium containing 89% minimum essential medium (MEM), 10% FBS, and 1% non-essential amino acids (NEAA). TE-1 cells were cultured in a medium composed of 90% RPMI 1640 medium and 10% FBS. T24 cells were grown in a medium consisting of 90% McCoy's 5a medium and 10%

Molecular Plant

FBS. BCPAP cells were cultured in a medium consisting of 89% RPMI 1640 medium, 10% FBS, and 1% NEAA. All of these cells were cultured in an incubator at 37°C and 5% CO₂.

Cell viability assay

Cell lines in logarithmic phase growth were seeded into 96-well plates at different densities with three replicate wells of treatment and control groups: for MCF-7, HepG2, A549, and PC-3 cells, 1 × 10⁴ cells per well; for HT29, HeLa, TE-1, and T24 cells, 5 × 10³ cells per well; for AGS, Huh-7, and BCPAP cells, 2 × 10³ cells per well. After culturing overnight, the original medium was removed, and the corresponding serum-containing medium for the cells was added, and then citrus extracts were added at a final concentration of 12.5, 25, 50, 100, and 200 µg/ml to incubate with the cells for 48 h. Cell viability assays were performed by Cell Counting Kit-8 assay. The cell culture medium was removed, and the cells were washed twice with PBS. Next, a solution of Cell Counting Kit-8 diluted in serum-free medium was added to the wells. After 1 h of incubation, the absorbance was measured at 620 nm and 450 nm using a microplate reader. Each experiment was performed in triplicate and repeated three times independently. The percentage of cell viability for each treatment group relative to the control group was calculated. The results were plotted as a dose-response curve to determine the IC₅₀.

CYP450s inhibition assay and real-time qPCR

HepG2 cells in logarithmic phase growth were seeded into 6-well plates at a density of 1 × 10⁶ cells per well and cultured overnight. Subsequently, the original medium was removed, and serum-containing DMEM was added. To investigate the effects of citrus extracts, the cells were treated with a final concentration of 10 µg/mL of the extracts and incubated for 48 h. Total RNA was extracted from the HepG2 cells using the Trizol method (Thermo Fisher Scientific, 15596-018), following the manufacturer's protocol. To ensure the removal of genomic RNA contamination, cDNA was synthesized using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara, RR047A). GAPDH was used as an internal control for normalization. Each treatment was performed in triplicate and repeated independently at least three times to ensure the reliability of the results. The relative gene expression was determined using the $2^{-\Delta\Delta Ct}$ method, which compares the target gene expression to that of the control group.

Statistical analysis

The statistical significance analysis of pairwise comparison was performed by Student's *t*-test using SPSS v.19.0 (IBM, Armonk, NY, USA), in which the analysis involving gene transient overexpression (Supplemental Figures 13, 21E, and 22) was performed by paired *t*-test (**p* < 0.05, ***p* < 0.01, ****p* < 0.001). The statistical significance analysis in Figures 2A and 6A–6C was performed by ANOVA with Duncan's correction for multiple comparisons using SPSS v.19.0. Statistical significance was set at *p* < 0.05.

DATA AND CODE AVAILABILITY

The sequence data of 50 accessions have been published in The Citrus Genomic Variation Database v.1.0 (<http://citgvd.cric.cn/home/index>). The sequence data of the remaining 249 accessions were deposited in NCBI with the BioProject accession number PRJNA993172.

FUNDING

This research was supported by the National Key Research and Development Program of China (2023YFD2300604), the National Natural Science Foundation of China (32072132, 32101932), the Hangzhou Joint Fund of the Zhejiang Provincial Natural Science Foundation of China (LHZSD24C150001), the Fundamental Research Funds for the Zhejiang Provincial Universities (226-2024-00211), and the Fundamental Research Funds for the Central Universities (226-2024-00235).

ACKNOWLEDGMENTS

We would like to thank the Citrus Research Institute of Zhejiang Province (Taizhou, China) and the Citrus Research Institute of Chinese Academy of Agricultural Sciences (Chongqing, China) for providing the citrus materials and resequenced data. We would like to appreciate Mr. Dengliang Wang from the Citrus Research Institute, Quzhou Academy of Agricultural Sciences for providing citrus materials for performing transient gene overexpression. We would like to express our gratitude to Prof. Xingxing Shen from Zhejiang University for his refinement of the language and images in the article. We thank the staff of Wuhan Metware Biotechnology Co., Ltd. (Wuhan, China) for their support during the metabolite detection and data analysis. No conflict of interest is declared.

AUTHOR CONTRIBUTIONS

C.S. contributed the design of the project. X. Liang performed the material preparation, completed most of the experiments and data analyses, and dominated the writing of this manuscript. Y.W. performed the experiments and data analysis of the bioactivity evaluation part and participated in the writing of this manuscript. W.S. dominated the variety collection and resequencing data collation. B.L., X. Liu, Z.Y., C.Z., and Z.L. participated in the experiments of gene function verification. J. Chen helped to draw the figures. J. Cao provided guidance on manuscript revision. Peng Wang, Ping Wang, F.K., J.X., Q.L., W.X., J.X., and X.Z. provided advice on the population structure and evolution of citrus. W.L. provided advice on biological activity evaluation and helped to revise the manuscript.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Molecular Plant Online*.

Received: January 4, 2024

Revised: June 30, 2024

Accepted: October 22, 2024

Published: October 23, 2024

REFERENCES

- Bailey, D.G., Dresser, G., and Arnold, J.M.O. (2013). Grapefruit-medication interactions: Forbidden fruit or avoidable consequences? *CMAJ*. **185**:309–316.
- Bao, Y., Huiling, L., Jiali, Y., Kumar Gupta, V., and Yueming, J. (2018). New insights on bioactivities and biosynthesis of flavonoid glycosides. *Trends Food Sci. Technol.* **79**:116–124.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2020). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries (vol 68, pg 394, 2018). *CA A Cancer J. Clin.* **70**:313.
- Cao, X., and Pan, S. (2022). Progress in research on secondary metabolites and biological activity of medicinal and edible Citrus plants. *Food Sci. (N. Y.)* **43**:305–315.
- Chen, X., Sun, W., Gianaris, N.G., Riley, A.M., Cummins, T.R., Fehrenbacher, J.C., and Obukhov, A.G. (2014). Furanocoumarins are a novel class of modulators for the transient receptor potential vanilloid type 1 (TRPV1) channel. *J. Biol. Chem.* **289**:9600–9610.
- Chen, L., Teng, H., Xie, Z., Cao, H., Cheang, W.S., Skalicka-Woniak, K., Georgiev, M.I., and Xiao, J. (2018a). Modifications of dietary flavonoids towards improved bioactivity: An update on structure-activity relationship. *Crit. Rev. Food Sci. Nutr.* **58**:513–527.
- Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., Yu, S., Xiong, L., and Luo, J. (2013). A Novel Integrated Method for Large-Scale Detection, Identification, and Quantification of Widely Targeted Metabolites: Application in the Study of Rice Metabolomics. *Mol. Plant* **6**:1769–1780.
- Cao, K., Wang, B., Fang, W., Zhu, G., Chen, C., Wang, X., Li, Y., Wu, J., Tang, T., Fei, Z., et al. (2022). Combined nature and human selections reshaped peach fruit metabolome. *Genome Biol.* **23**:146.
- Cao, K., Yang, X., Li, Y., Zhu, G., Fang, W., Chen, C., Wang, X., Wu, J., and Wang, L. (2021). New high-quality peach (*Prunus persica* L. Batsch) genome assembly to analyze the molecular evolutionary mechanism of volatile compounds in peach fruits. *Plant J.* **108**:281–295.
- Chen, Y., Chen, Y., Shi, C., Huang, Z., Zhang, Y., Li, S., Li, Y., Ye, J., Yu, C., Li, Z., et al. (2018b). SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience* **7**:1–6.
- Cirmi, S., Ferlazzo, N., Lombardo, G.E., Ventura-Spagnolo, E., Gangemi, S., Calapai, G., and Navarra, M. (2016). Neurodegenerative Diseases: Might Citrus Flavonoids Play a Protective Role? *Molecules* **21**:1312.
- Deng, Y., Tu, Y., Lao, S., Wu, M., Yin, H., Wang, L., and Liao, W. (2022). The role and mechanism of citrus flavonoids in cardiovascular diseases prevention and treatment. *Crit. Rev. Food Sci. Nutr.* **62**:7591–7614.
- Dugrand, A., Olry, A., Duval, T., Hehn, A., Froelicher, Y., and Bourgaud, F. (2013). Coumarin and furanocoumarin quantitation in citrus peel via ultraperformance liquid chromatography coupled with mass spectrometry (UPLC-MS). *J. Agric. Food Chem.* **61**:10677–10684.
- Dugrand-Judek, A., Olry, A., Hehn, A., Costantino, G., Ollitrault, P., Froelicher, Y., and Bourgaud, F. (2015). The Distribution of Coumarins and Furanocoumarins in Citrus Species Closely Matches Citrus Phylogeny and Reflects the Organization of Biosynthetic Pathways. *PLoS One* **10**:e0142757.
- Ejaz, S., Ejaz, A., Matsuda, K., and Lim, C.W. (2006). Limonoids as cancer chemopreventive agents. *J. Sci. Food Agric.* **86**:339–345.
- Fang, C., and Luo, J. (2019). Metabolic GWAS-based dissection of genetic bases underlying the diversity of plant metabolism. *Plant J.* **97**:91–100.
- Fraga, C.G., Clowers, B.H., Moore, R.J., and Zink, E.M. (2010). Signature-discovery approach for sample matching of a nerve-agent precursor using liquid chromatography-mass spectrometry, XCMS, and chemometrics. *Anal. Chem.* **82**:4165–4173.
- Frydman, A., Liberman, R., Huhman, D.V., Carmeli-Weissberg, M., Sapir-Mir, M., Ophir, R., W Sumner, L., and Eyal, Y. (2013). The molecular and enzymatic basis of bitter/non-bitter flavor of citrus fruit: evolution of branch-forming rhamnosyltransferases under domestication. *Plant J.* **73**:166–178.
- Gansukh, E., Nile, A., Sivanesan, I., Rengasamy, K.R.R., Kim, D.H., Keum, Y.S., and Saini, R.K. (2019). Chemopreventive Effect of beta-Cryptoxanthin on Human Cervical Carcinoma (HeLa) Cells Is Modulated through Oxidative Stress-Induced Apoptosis. *Antioxidants* **9**:28.
- Gmitter, F.G., Chen, C., Wei, X., Yu, Y., and Yu, Q. (2014). New genetic tools to improve citrus fruit quality and drive consumer demand. In 29th International Horticultural Congress on Horticulture - Sustaining Lives, Livelihoods and Landscapes (IHC)/International Symposium on Plant Breeding in Horticulture Brisbane, pp. 199–202.
- Goh, J.X.H., Tan, L.T.H., Goh, J.K., Chan, K.G., Pusparajah, P., Lee, L.H., and Goh, B.H. (2019). Nobletin and derivatives: functional compounds from citrus fruit peel for colon cancer chemoprevention. *Cancers* **11**:867.
- Ho, P.C., Saville, D.J., and Wanwimolruk, S. (2001). Inhibition of human CYP3A4 activity by grapefruit flavonoids, furanocoumarins and related compounds. *J. Pharm. Pharmaceut. Sci.* **4**:217–227.
- Hou, Y., and Jiang, J.G. (2013). Origin and concept of medicine food homology and its application in modern functional foods. *Food Funct.* **4**:1727–1741.

Multi-omics analyses on citrus bioactive compounds

- Huang, Y., He, J., Xu, Y., Zheng, W., Wang, S., Chen, P., Zeng, B., Yang, S., Jiang, X., Liu, Z., et al. (2023). Pangenome analysis provides insight into the evolution of the orange subfamily and a key gene for citric acid accumulation in citrus fruits. *Nat. Genet.* **55**:1964–1975.
- Ito, T., Fujimoto, S., Suito, F., Shimosaka, M., and Taguchi, G. (2017). C-Glycosyltransferases catalyzing the formation of di-C-glucosyl flavonoids in citrus plants. *Plant J.* **91**:187–198.
- Jan, R., Asaf, S., Numan, M., Kim, K.M., and Kim, K.-M. (2021). Plant secondary metabolite biosynthesis and transcriptional regulation in response to biotic and abiotic stress conditions. *Agronomy* **11**:968.
- Jeong Ah, P., Ji Eun, O., and Mi Sook, C. (2019). Development of yuja (*Citrus junos*) beverage based on antioxidant properties and sensory attributes using response surface methodology. *J. Food Sci. Technol.* **56**:1854–1863.
- Kai, K., Mizutani, M., Kawamura, N., Yamamoto, R., Tamai, M., Yamaguchi, H., Sakata, K., and Shimizu, B.-i. (2008). Scopoletin is biosynthesized viaortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase inArabidopsis thaliana. *Plant J.* **55**:989–999.
- Kang, L., Qian, L., Zheng, M., Chen, L., Chen, H., Yang, L., You, L., Yang, B., Yan, M., Gu, Y., et al. (2021). Genomic insights into the origin, domestication and diversification of Brassica juncea. *Nat. Genet.* **53**:1392–1402.
- Kaur, S., Panesar, P.S., and Chopra, H.K. (2023). Citrus processing by-products: an overlooked repository of bioactive compounds. *Crit. Rev. Food Sci. Nutr.* **63**:67–86.
- Kiefer, E., Heller, W., and Ernst, D. (2000). A simple and efficient protocol for isolation of functional RNA from plant tissues rich in secondary metabolites. *Plant Mol. Biol. Rep.* **18**:33–39.
- Kreuzaler, F., and Hahlbrock, K. (1972). Enzymatic synthesis of aromatic compounds in higher plants: Formation of naringenin (5,7,4'-trihydroxyflavanone) from p-coumaroyl coenzyme A and malonyl coenzyme A. *FEBS Lett.* **28**:69–72.
- Kris-Etherton, P.M., Hecker, K.D., Bonanome, A., Coval, S.M., Binkoski, A.E., Hilpert, K.F., Griel, A.E., and Etherton, T.D. (2002). Bioactive compounds in foods: Their role in the prevention of cardiovascular disease and cancer. *Am. J. Med.* **113**:71S–88S.
- Kruse, T., Ho, K., Yoo, H.D., Johnson, T., Hippely, M., Park, J.H., Flavell, R., and Bobzin, S. (2008). In planta biocatalysis screen of P450s identifies 8-methoxysoralen as a substrate for the CYP82C subfamily, yielding original chemical structures. *Chem. Biol.* **15**:149–156.
- Larbat, R., Hehn, A., Hans, J., Schneider, S., Jugdé, H., Schneider, B., Matern, U., and Bourgaud, F. (2009). Isolation and functional characterization of CYP71AJ4 encoding for the first P450 monooxygenase of angular furanocoumarin biosynthesis. *J. Biol. Chem.* **284**:4776–4785.
- Li-Rong, K. (2021). Research progress on the effects of bioactive polysaccharides in natural products on intestinal disease. *Journal of Food Safety and Quality* **12**:5507–5512.
- 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. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**:589–595.
- Li, S., Pan, M.-H., Lo, C.-Y., Tan, D., Wang, Y., Shahidi, F., and Ho, C.-T. (2009). Chemistry and health effects of polymethoxyflavones and hydroxylated polymethoxyflavones. *J. Funct. Foods* **1**:2–12.
- Li, X., Gao, J., Song, J., Guo, K., Hou, S., Wang, X., He, Q., Zhang, Y., Zhang, Y., Yang, Y., et al. (2022). Multi-omics analyses of 398 foxtail millet accessions reveal genomic regions associated with domestication,metabolite traits, and antiinflammatory effects. *Mol. Plant* **15**:1367–1383.
- Limones-Mendez, M., Dugrand-Judek, A., Villard, C., Coqueret, V., Froelicher, Y., Bourgaud, F., Olry, A., and Hehn, A. (2020). Convergent evolution leading to the appearance of furanocoumarins in citrus plants. *Plant Sci.* **292**:110392.
- Liu, N., Li, X., Zhao, P., Zhang, X., Qiao, O., Huang, L., Guo, L., and Gao, W. (2021). A review of chemical constituents and health-promoting effects of citrus peels. *Food Chem.* **365**:130585.
- Liu, X., Gong, Q., Zhao, C., Wang, D., Ye, X., Zheng, G., Wang, Y., Cao, J., and Sun, C. (2023). Genome-wide analysis of cytochrome P450 genes in Citrus clementina and characterization of a CYP gene encoding flavonoid 3'-hydroxylase. *Hortic. Res.* **10**:uhac283.
- Liu, X., Zhao, C., Gong, Q., Wang, Y., Cao, J., Li, X., Grierson, D., and Sun, C. (2020). Characterization of a caffeoyl-CoA O-methyltransferase-like enzyme involved in biosynthesis of polymethoxylated flavones in Citrus reticulata. *J. Exp. Bot.* **71**:3066–3079.
- Lu, X., Zhao, C., Shi, H., Liao, Y., Xu, F., Du, H., Xiao, H., and Zheng, J. (2023). Nutrients and bioactives in citrus fruits: Different citrus varieties, fruit parts, and growth stages. *Crit. Rev. Food Sci. Nutr.* **63**:2018–2041.
- Munakata, R., Olry, A., Takemura, T., Tatsumi, K., Ichino, T., Villard, C., Kageyama, J., Kurata, T., Nakayasu, M., Jacob, F., et al. (2021). Parallel evolution of UbiA superfamily proteins into aromatic O-prenyltransferases in plants. *Proc. Natl. Acad. Sci. USA* **118**:e2022294118.
- Nanney, M.S., Haire-Joshu, D., Hessler, K., and Brownson, R.C. (2004). Rationale for a consistent “powerhouse” approach to vegetable and fruit messages. *J. Am. Diet Assoc.* **104**:352–356.
- Okonechnikov, K., Conesa, A., and García-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**:292–294.
- Owens, D.K., and McIntosh, C.A. (2009). Identification, recombinant expression, and biochemical characterization of a flavonol 3-O-glucosyltransferase clone from *Citrus paradisi*. *Phytochemistry* **70**:1382–1391.
- Patil, S.P., Goswami, A., Kalia, K., and Kate, A.S. (2020). Plant-derived bioactive peptides: a treatment to cure diabetes. *Int. J. Pept. Res. Therapeut.* **26**:955–968.
- Peng, M., Shahzad, R., Gul, A., Subthain, H., Shen, S., Lei, L., Zheng, Z., Zhou, J., Lu, D., Wang, S., et al. (2017). Differentially evolved glucosyltransferases determine natural variation of rice flavone accumulation and UV-tolerance. *Nat. Commun.* **8**:1975.
- Peng, Z., Zhang, H., Li, W., Yuan, Z., Xie, Z., Zhang, H., Cheng, Y., Chen, J., and Xu, J. (2021). Comparative profiling and natural variation of polymethoxylated flavones in various citrus germplasms. *Food Chem.* **354**:129499.
- Petric, Z., Žuntar, I., Putnik, P., and Bursać Kovacević, D. (2020). Food-Drug Interactions with Fruit Juices. *Foods* **10**:33.
- Pillitteri, L.J., Lovatt, C.J., and Walling, L.L. (2004). Isolation and characterization of *LEAFY* and *APETALA1* homologues from *Citrus sinensis* L. Osbeck ‘Washington’. *J. Am. Soc. Hortic. Sci.* **129**:846–856.
- Plaza, M., Pozzo, T., Liu, J., Gulshan Ara, K.Z., Turner, C., and Nordberg Karlsson, E. (2014). Substituent effects on in vitro antioxidant properties, stability, and solubility in flavonoids. *J. Agric. Food Chem.* **62**:3321–3333.
- Ramadan, M.F., and Oraby, H.F. (2016). Fatty Acids and Bioactive Lipids of Potato Cultivars: An Overview. *J. Oleo Sci.* **65**:459–470.
- Riveiro, M.E., De Kimpe, N., Moglioni, A., Vázquez, R., Monczor, F., Shayo, C., and Davio, C. (2010). Coumarins: Old Compounds with Novel Promising Therapeutic Perspectives. *Curr. Med. Chem.* **17**:1325–1338.

Molecular Plant

- Saini, R.K., Ranjit, A., Sharma, K., Prasad, P., Shang, X., Gowda, K.G.M., and Keum, Y.S. (2022). Bioactive Compounds of Citrus Fruits: A Review of Composition and Health Benefits of Carotenoids, Flavonoids, Limonoids, and Terpenes. *Antioxidants* **11**:239.
- Samtiya, M., Aluko, R.E., Dhewa, T., and Moreno-Rojas, J.M. (2021). Potential Health Benefits of Plant Food-Derived Bioactive Components: An Overview. *Foods* **10**:839.
- Shen, S., Wang, S., Yang, C., Wang, C., Zhou, Q., Zhou, S., Zhang, R., Li, Y., Wang, Z., Dai, L., et al. (2023). Elucidation of the melittidin biosynthesis pathway in pummelo. *J. Integr. Plant Biol.* **65**:2505–2518.
- Shimizu, T., Kitajima, A., Nonaka, K., Yoshioka, T., Ohta, S., Goto, S., Toyoda, A., Fujiyama, A., Mochizuki, T., Nagasaki, H., et al. (2016). Hybrid Origins of Citrus Varieties Inferred from DNA Marker Analysis of Nuclear and Organelle Genomes. *PLoS One* **11**:e0166969.
- Sun, L., Xu, J., Nasrullah, K., Nie, Z., and Wang, P. (2019). Citrus Genetic Engineering for Disease Resistance: Past, Present and Future. *Int. J. Mol. Sci.* **20**:5256.
- Tieman, D., Zhu, G., Resende, M.F.R., Lin, T., Nguyen, C., Bies, D., Rambla, J.L., Beltran, K.S.O., Taylor, M., Zhang, B., et al. (2017). A chemical genetic roadmap to improved tomato flavor. *Science* **355**:391–394.
- Tripoli, E., Guardia, M.L., Giannamico, S., Majo, D.D., and Giannamico, M. (2007). Citrus flavonoids: Molecular structure, biological activity and nutritional properties: A review. *Food Chem.* **104**:466–479.
- Vialart, G., Hehn, A., Olry, A., Ito, K., Krieger, C., Larbat, R., Paris, C., Shimizu, B.I., Sugimoto, Y., Mizutani, M., and Bourgaud, F. (2012). A 2-oxoglutarate-dependent dioxygenase from *Ruta graveolens* L. exhibits p-coumaroyl CoA 2'-hydroxylase activity (C2'H): a missing step in the synthesis of umbelliferone in plants. *Plant J.* **70**:460–470.
- 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, X., Xu, Y., Zhang, S., Cao, L., Huang, Y., Cheng, J., Wu, G., Tian, S., Chen, C., Liu, Y., et al. (2017a). Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nat. Genet.* **49**:765–772.
- Wang, Y., Qian, J., Cao, J., Wang, D., Liu, C., Yang, R., Li, X., and Sun, C. (2017b). Antioxidant Capacity, Anticancer Ability and Flavonoids Composition of 35 Citrus (*Citrus reticulata* Blanco) Varieties. *Molecules* **22**:1114.
- 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., Ibáñez, 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.
- Xiao, J., and Bai, W. (2019). Bioactive phytochemicals. *Crit. Rev. Food Sci. Nutr.* **59**:827–829.
- Xu, C.J., Bao, L., Zhang, B., Bei, Z.M., Ye, X.Y., Zhang, S.L., and Chen, K.S. (2006). Parentage analysis of huyou (*Citrus changshanensis*) based on internal transcribed spacer sequences. *Plant Breed.* **125**:519–522.
- 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.
- Multi-omics analyses on citrus bioactive compounds
- Yan, H., Pu, Z.J., Zhang, Z.Y., Zhou, G.S., Zou, D.Q., Guo, S., Li, C., Zhan, Z.L., and Duan, J.A. (2021). Research on Biomarkers of Different Growth Periods and Different Drying Processes of *Citrus wilsonii* Tanaka Based on Plant Metabolomics. *Front. Plant Sci.* **12**:700367.
- Yang, Y., Wang, X., Zhao, C., Tian, G., Zhang, H., Xiao, H., He, L., and Zheng, J. (2017). Chemical Mapping of Essential Oils, Flavonoids and Carotenoids in Citrus Peels by Raman Microscopy. *J. Food Sci.* **82**:2840–2846.
- Yu, Y., Guan, J., Xu, Y., Ren, F., Zhang, Z., Yan, J., Fu, J., Guo, J., Shen, Z., Zhao, J., et al. (2021). Population-scale peach genome analyses unravel selection patterns and biochemical basis underlying fruit flavor. *Nat. Commun.* **12**:3604.
- Yuan, P., Xu, C., He, N., Lu, X., Zhang, X., Shang, J., Zhu, H., Gong, C., Kuang, H., Tang, T., et al. (2023). Watermelon domestication was shaped by stepwise selection and regulation of the metabolome. *Sci. China Life Sci.* **66**:579–594.
- Yun, Y.B., Jeong, H., Cho, Y., and Kim, S. (2024). Analysis of genetic diversity of yuzu (*Citrus junos* Sieb. Ex Tanaka) using single nucleotide polymorphisms identified through RNA-seq and whole-genome resequencing analyses. *J. Hortic. Sci. Biotechnol.* **99**:298–310.
- Zeng, X., Yuan, H., Dong, X., Peng, M., Jing, X., Xu, Q., Tang, T., Wang, Y., Zha, S., Gao, M., et al. (2020). Genome-wide Dissection of Co-selected UV-B Responsive Pathways in the UV-B Adaptation of Qingke. *Mol. Plant* **13**:112–127.
- Zhang, Y., Jin, J., Wang, N., Sun, Q., Feng, D., Zhu, S., Wang, Z., Li, S., Ye, J., Chai, L., et al. (2024). Cytochrome P450 CYP97B modulates carotenoid accumulation diversity by hydroxylating beta-cryptoxanthin in Citrus. *Plant Commun.* **5**:100847.
- Zhang, Z., Shi, Q., Wang, B., Ma, A., Wang, Y., Xue, Q., Shen, B., Hamaila, H., Tang, T., Qi, X., et al. (2022). Jujube metabolome selection determined the edible properties acquired during domestication. *Plant J.* **109**:1116–1133.
- Zhao, C., Wang, F., Lian, Y., Xiao, H., and Zheng, J. (2020). Biosynthesis of citrus flavonoids and their health effects. *Crit. Rev. Food Sci. Nutr.* **60**:566–583.
- Zhao, H., He, Y., Zhang, K., Li, S., Chen, Y., He, M., He, F., Gao, B., Yang, D., Fan, Y., et al. (2023). Rewiring of the seed metabolome during Tartary buckwheat domestication. *Plant Biotechnol. J.* **21**:150–164.
- Zhao, J., Xu, Y., Li, H., An, W., Yin, Y., Wang, B., Wang, L., Wang, B., Duan, L., Ren, X., et al. (2024). Metabolite-based genome-wide association studies enable the dissection of the genetic bases of flavonoids, betaine and spermidine in wolfberry (*Lycium*). *Plant Biotechnol. J.* **22**:1435–1452.
- Zhao, X., Pang, W., Tan, T., Zhang, Y., and Jiao, B. (2022). Simultaneous determination of 26 bioactive components in lemon fruits based on UPLC-Q-TOF-HRMS and SWATH acquisition mode. *Food Ferment. Ind.* **48**:306–314.
- Zheng, W., Zhang, W., Liu, D., Yin, M., Wang, X., Wang, S., Shen, S., Liu, S., Huang, Y., Li, X., et al. (2023). Evolution-guided multiomics provide insights into the strengthening of bioactive flavone biosynthesis in medicinal pummelo. *Plant Biotechnol. J.* **21**:1577–1589.
- Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., Lin, T., Qin, M., Peng, M., Yang, C., et al. (2018). Rewiring of the Fruit Metabolome in Tomato Breeding. *Cell* **172**:249–261.e12.