








Evolutionary gain of oligosaccharide hydrolysis and sugar transport enhanced carbohydrate partitioning in sweet watermelon fruits

Yi Ren ^{1,2,*†}, Maoying Li ¹, Shaogui Guo ¹, Honghe Sun ¹, Jianyu Zhao ³, Jie Zhang ¹, Guangmin Liu ¹, Hongju He ¹, Shouwei Tian ¹, Yongtao Yu ¹, Guoyi Gong ¹, Haiying Zhang ¹, Xiaolan Zhang ³, Saleh Alseekh ^{4,5}, Alisdair R. Fernie ^{4,5}, Henrik V. Scheller ^{2,6} and Yong Xu ^{1,*†}

- 1 National Watermelon and Melon Improvement Center, Beijing Academy of Agriculture and Forestry Sciences, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (North China), Beijing Key Laboratory of Vegetable Germplasm Improvement, Beijing 100097, China
- 2 Joint BioEnergy Institute and Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA
- 3 Beijing Key Laboratory of Growth and Developmental Regulation for Protected Vegetable Crops, MOE Joint International Research Laboratory of Crop Molecular Breeding, China Agricultural University, Beijing 100193, China
- 4 Max-Planck-Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany
- 5 Center of Plant System Biology and Biotechnology, 4000 Plovdiv, Bulgaria
- 6 Department of Plant and Microbial Biology, University of California, Berkeley, California, USA

*Author for correspondence: xuyong@nercv.org (Y.X.), renyi@nercv.org (Y.R.)

†Senior authors.

The authors responsible for the distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plcell>) are Yong Xu (xuyong@nercv.org) and Yi Ren (renyi@nercv.org).

Y.R. conducted most experiments and wrote the manuscript. M.L. performed the phenotyping of transgenic plants. S.G. and H.S. performed most of the bioinformatics analysis. J. Zhao conducted RNA in situ hybridization. J. Zhang provided technical assistance. G.L. and H.H. contributed to the metabolic analysis. S.T. and Y.Y. participated in the phenotype analysis. G.G. planted the population and managed this procedure in the greenhouse. Z.H. performed the transient expression assays. X.Z. designed the RNA in situ hybridization. S.A. and A.R.F. conducted the metabolome study. H.V.S. contributed to the discussion of the experimental results and manuscript writing. Y.X. designed and managed the study and revised the manuscript.

Abstract

How raffinose (Raf) family oligosaccharides, the major translocated sugars in the vascular bundle in cucurbits, are hydrolyzed and subsequently partitioned has not been fully elucidated. By performing reciprocal grafting of watermelon (*Citrullus lanatus*) fruits to branch stems, we observed that Raf was hydrolyzed in the fruit of cultivar watermelons but was backlogged in the fruit of wild ancestor species. Through a genome-wide association study, the alkaline alpha-galactosidase *CIAGA2* was identified as the key factor controlling stachyose and Raf hydrolysis, and it was determined to be specifically expressed in the vascular bundle. Analysis of transgenic plants confirmed that *CIAGA2* controls fruit Raf hydrolysis and reduces sugar content in fruits. Two single-nucleotide polymorphisms (SNPs) within the *CIAGA2* promoter affect the recruitment of the transcription factor CINF-YC2 (nuclear transcription factor Y subunit C) to regulate *CIAGA2* expression. Moreover, this study demonstrates that *C. lanatus* Sugars Will Eventually Be Exported Transporter 3 (CISWEET3) and Tonoplast Sugar Transporter (CITST2) participate in plasma membrane sugar transport and sugar storage in fruit cell vacuoles, respectively. Knocking out *CIAGA2*, *CISWEET3*, and *CITST2* affected fruit sugar accumulation. Genomic signatures

indicate that the selection of *CIAGA2*, *CISWEET3*, and *CITST2* for carbohydrate partitioning led to the derivation of modern sweet watermelon from non-sweet ancestors during domestication.

Introduction

The concept of “source” organs that are net exporters of photosynthetic assimilates and “sink” organs that are net importers of photosynthate dates back over a century (Sonnewald and Fernie, 2018). In most plants, sucrose (Suc) is the predominant sugar transported in the vascular bundle, while raffinose (Raf) family oligosaccharides (RFOs) are the major sugars transported between the source and sink tissues in the Cucurbitaceae (Zhang et al., 2012), Lamiaceae (Büchi et al., 1998), Oleaceae (Davidson et al., 2011), and Scrophulariaceae (McCaskill and Turgeon, 2007), along with minor amounts of sucrose and hexose (Webb and Gorham, 1964; Schmitz et al., 1987; Zhang et al., 2012). This redistribution involves long-distance transport of Suc between source and sink organs and hydrolysis of RFOs in sieve elements and the companion cell complex (Bel, 1993). Suc is the predominant sugar transported in the phloem in most crops, while in watermelon and other Cucurbitaceae, RFOs are the principal sugars for translocation from the source to the sink (Webb and Gorham, 1964; Schmitz et al., 1987; Zhang et al., 2012). Metabolomic analyses demonstrated that in watermelon, only 6% stachyose (Sta) and 10% Raf are detected in the sink (apical region of the plant) compared to the contents in the central source region (Hu et al., 2016), suggesting that hydrolysis of RFOs may occur in sink tissues.

The RFOs are loaded in the phloem of leaves through a polymer trapping process (Rennie and Turgeon, 2009). After long-distance translocation in veins, RFOs are hydrolyzed into Suc and galactose (Gal) to be unloaded from the phloem in sink tissues (Chrost and Schmitz, 1997; Dai et al., 2011). Although it is hypothesized that this process occurs in the companion cells of fruit, experimental evidence is lacking (Ma et al., 2018). Thus, research-exploring genes contributing to RFO hydrolysis and subsequent Suc and hexose transport across the membrane in the phloem, cytomembrane, and tonoplast is warranted. In cucumber and watermelon, Suc could be unloaded from the phloem by sugar transporters through the apoplasmic pathway (Hu et al., 2011; Ren et al., 2020). High concentrations of reducing sugars in the cytosol may result in nonenzymatic glycosylation, damaging functional, and structural proteins (Yamauchi et al., 2002). Therefore, it is necessary for excess reducing sugars to be stored in vacuoles through tonoplast sugar transporter (TST) or tonoplast monosaccharide transporter (TMT), such as AtTMTs in Arabidopsis (Wingenter et al., 2010), BvTSTs in sugar beet (*Beta vulgaris*) (Jung et al., 2015), CITST2 in watermelon (Ren et al., 2018), and CmTST2 in melon (Cheng et al., 2018). Within the last decade, considerable research has been conducted to understand the

regulation of sugar partitioning by various sugar transporters. In Arabidopsis, AtSWEET9 plays a role in nectar secretion (Lin et al., 2014). ZmSWEET4c and OsSWEET4 take up hexose in maize and rice for seed filling (Sosso et al., 2015), respectively. Glucose (Glc) transporter CLOSED STOMATA 1 (ZmCST1) functions as a positive regulator of stomatal opening in maize, and its expression is induced by carbon starvation in maize (Wang et al., 2019). At the same time, the creation of broad-spectrum resistance to bacterial blight in rice through mutations in the promoters of three sucrose transporter genes, *SWEET11*, *SWEET13*, and *SWEET14*, was recently reported (Oliva et al., 2019). However, thorough functional studies of sugar transporters in fruit crops have not been conducted. To date, there is no clear understanding of the function and the evolution of genes controlling sugar import to the cytosol in fruits.

Crops in the Cucurbitaceae family are widely consumed as fruits and vegetables. The ancestral species of sweet watermelon (*Citrullus lanatus*, CL) and *C. colocynthis* (CC) originated in southern Africa and extended into Libya and Egypt during humid epochs of the Pleistocene and Holocene (Schulz, 1991; Chomicki and Renner, 2015), accumulating low levels of biomass, sugars and lycopene in a small, bitter fruit. The wild-type (WT) nonsweet watermelon *C. amarus* also originated in southern Africa, and the cultivated types are planted around the Mediterranean and used for jams and animal fodder and as sources of water (Renner et al., 2017). The semi-wild ancestor most closely related to modern cultivars, *C. mucospermus* (CM), contains both sweet and nonsweet fruits, while breeding selection has led to cultivars that store a large amount of biomass and sugars in large fruits and germinate seeds rapidly. The illustrations of watermelon served on a tray found in the tomb of Pharaoh Tutankhamun suggest that watermelon was cultivated as a dessert in ancient Egypt more than 4,000 years ago (Hepper, 1990). Currently, little is known regarding the proteins involved in vascular hydrolysis of RFOs and sugar transport in fruits. Moreover, whether the evolution and domestication of these traits occurred during natural and human breeding selection has not been determined.

In this study, we identified the protein responsible for RFO hydrolysis in the fruit vascular bundle. Moreover, we identified genes for sugar transporters that import sugars across the plasma membrane and tonoplasts in fruit storage cells. Furthermore, we demonstrated that these carbohydrate-partitioning genes have undergone evolution and domestication across wild, semi-wild, and cultivated watermelon species.

Results

Backlog of RFOs in nonsweet watermelon fruits

To elucidate the dynamic alterations in sugar content from stem flow to sink, we collected vascular exudates from stems and nearby carpodia in watermelon. Next, we measured the relative sugar contents in vascular exudates and determined that RFOs are the main sugars transported in both stem and carpodium “mobile” vascular exudates (Figure 1A), similar to the sugar component in vascular exudates of other *Cucurbitaceae*, indicating that RFOs are not hydrolyzed from stem to carpodium. However, watermelon cultivars primarily accumulate Suc, fructose (Fru), and Glc in their fruit (Dai et al., 2011; Ren et al., 2014), while RFOs are present at negligible levels in sink tissues (Figure 1B). It is reasonable to surmise that RFOs cannot be hydrolyzed in carpodia because their concentrations are almost the same in stem and carpodia vascular exudates (Figure 1A). Hence, it is highly possible that RFOs could be hydrolyzed in watermelon fruit, but there is no direct evidence to show that RFOs are hydrolyzed in sink tissues.

Meanwhile, using gas chromatography-mass spectrometry (GC-MS) and ion chromatography-based metabolite analysis, we observed that the Raf and Sta contents in sweet watermelon fruits were only negligible (0.26 and 0.12 mg g⁻¹ DW, respectively; Figure 1B; Supplemental Data Set 1). In contrast, approximately six-fold more Raf was detected in wild nonsweet watermelon accession PI 296341 fruit than in cultivated watermelon 97103 fruit (Figure 1B; Supplemental Data Set 1). However, the concentration of RFOs was 3.5-fold lower in the stem and carpodium exudate of the nonsweet watermelon PI 296341 compared with that of sweet cultivar 97103 (Figure 1A), supporting the previous finding that more biomass accumulated in cultivar 97103 than in nonsweet watermelon PI 296341 (Liu et al., 2013). Taken together, these results demonstrated that many more RFOs can be hydrolyzed in sweet watermelon fruit and were not entirely hydrolyzed in nonsweet watermelons. These findings indicated that the fruit of sweet watermelons possesses stronger RFO-hydrolyzing activity and considerably greater flow, and higher concentrations of carbohydrates are translocated from the source to the fruit sink.

To explore in which tissue RFOs were hydrolyzed, we performed reciprocal grafting of watermelon fruits. We grafted young fruits of the sweet cultivar 97103 to branch stems of the non-sweet wild accession PI 296341 (Figure 1C). The results indicated that sugar accumulation was not significantly different from that of the 97103 self-grafted controls at 34 days after pollination (DAP) (Figure 1D). Similarly, no significant difference in sugar content was detected between young PI 296341 fruits grafted to 97103 branch stems and self-grafted PI 296341 controls (Figure 1D). These results support that the fruit sink has the independent capacity for hydrolyzing RFOs and accumulating sugars without being influenced by leaf, stem, root, or rootstock tissues in watermelon. To further investigate whether differential activities for hydrolyzing RFOs are ubiquitous in different watermelon

fruits, we measured the Raf contents in fruits of a panel of 135 representative watermelon accessions collected worldwide. The results demonstrated that Raf hydrolysis was backlogged in non-sweet wild watermelon fruits, while Raf was hydrolyzed to a very low level in sweet cultivated watermelon fruit (Figure 1E; Supplemental Data Set 2).

CIAGA2 contributes to RFO hydrolysis

The cultivated watermelon fruit showed a lower Raf content than did wild ancestor watermelons (Figure 1E). Therefore, we performed a genome-wide association study (GWAS) for Raf content and identified significantly associated SNPs on chromosome 4, which were located at -159 (C-T) (CHR04_12223999) and -170 (C-T) (CHR04_12224010) of the promoter of *Cl97C04G070460*, the *C. lanatus alkaline alpha-galactosidase* gene (*CIAGA2*; EC3.2.1.22) (Figure 2A). RNA-Seq analysis demonstrated that the relative expression of *CIAGA2* in the fruit of cultivated 97103 increased sharply from 10 to 26 DAP (Guo et al., 2015, 2019) but remained at a low level in the mesocarp, stems, carpodia, leaves, and roots of 97103 and in the fruit of wild accession PI 296341 (Figure 2B). However, in 97103, the other three genes of the alkaline alpha-galactosidase family remained at low levels during fruit ripening (Figure 2B).

In cucurbit crops, the vascular bundles are bicollateral, with internal phloem (IP) located close to the fruit center and external phloem (EP) located beyond the fruit center, separated by xylem (Figure 2C; Hu et al., 2011; Ren et al., 2020). RNA in situ hybridization suggested localization in vascular bundles of *CIAGA2* in watermelon fruit, as expression was primarily detected in the IP, EP, and adjacent vascular parenchyma cells without obvious expression in surrounding fruit parenchymal (FP) cells (or fruit storage cells) (Figure 2D) compared to the negative control, which used a sense probe that could not complement *CIAGA2* mRNA. These combined results indicate that *CIAGA2* is likely a vascular bundle-localized protein for RFO hydrolysis in watermelon fruit.

Relative alpha-galactosidase activities in different tissues demonstrated that sweet watermelon fruit had relatively high activity, while it remained at considerably lower levels in other tissues, including stems and carpodia (Figure 2E), consistent with the transcript levels of *CIAGA2* in different tissues (Figure 2B). In vitro assays of the alkaline alpha-galactosidase activity of *CIAGA2* (Figure 2F) demonstrated that it possessed *K_m* values of 11 mM and 13.7 mM for hydrolyzing Raf and Sta, respectively (Figure 2G). Eight days after the injection of the alpha-galactosidase inhibitor 1-deoxy-galactonojirimycin (DGJ) (Platt et al., 2005; Blöchl et al., 2007), lighter fruit coloration (Figure 2H) and Suc, Fru, and Glc contents decreased by ~30% were observed; in contrast, the Raf content was six-fold higher than that of the negative control (Figure 2I). In keeping with this result, the activity of alpha-galactosidase in fruits was inhibited by DGJ (Figure 2J). These results strongly suggest that the *CIAGA2* enzyme is an essential factor mediating increased RFO hydrolysis and sugar accumulation in watermelon fruit.

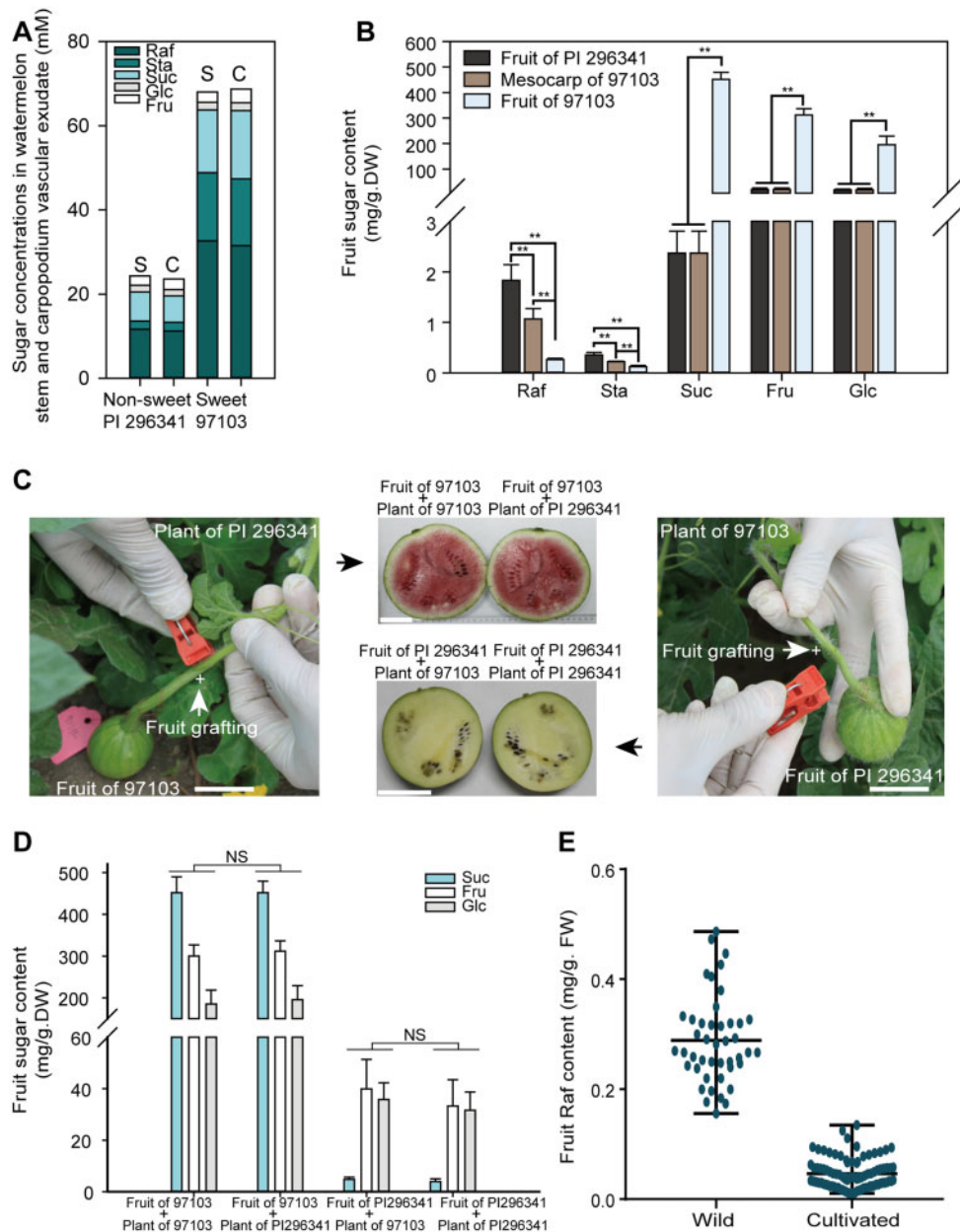


Figure 1 RFOs are hydrolyzed in fruits. **A**, Sugar content measured by ion chromatography in watermelon stem (S) and carpodium (C) vascular exudate. **B**, Fruit and mesocarp sugar content in wild and cultivated accessions. Samples collected from center fruit and mesocarp were measured by ion chromatography after freeze drying. Average values \pm standard deviation (SD) for $n = 3$ independent biological replicates are shown. ***t*-test significant at $P < 0.01$. **C**, Reciprocal grafting between wild and cultivated fruits (CVs). The grafting of cultivar 97103 fruit to stem of wild accession PI 296341 (left), and the grafting of wild accession PI 296341 fruit to stem of cultivar 97103 (right). The harvested fruits after grafting were shown in the middle. The grafting site between two accessions is shown as “+”. Bars = 5 cm. **D**, Sugar content in reciprocal grafted fruits. NS, no significant difference. **E**, Raf content in the non-sweet wild ancestors and cultivated sweet watermelons measured by FW. Each dot represents one accession, and the mean for $n = 3$ independent biological replicates for each accession is shown.

To further explore the function of *CIAGA2*, we used semi-wild PI 179878 plants as the WT for overexpressing (OE) *CIAGA2*. The fruit of *CIAGA2* overexpressors, containing one copy of the transgene, turned red earlier than that of the WT (Figure 3A), and the *CIAGA2* transcript abundance (Figure 3B) and the alpha-galactosidase activity (Figure 3C) were increased in OE plants compared to WT plants.

Correspondingly, the Raf content was decreased, and the Suc content increased significantly (Figure 3D). Next, using clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 gene editing (Doudna and Charpentier, 2014; Guo et al., 2019), *claga2* mutants from cultivar ZZJM were generated, and two mutant lines (T_2 generation) were analyzed (Figure 3E). We observed that the Raf content in

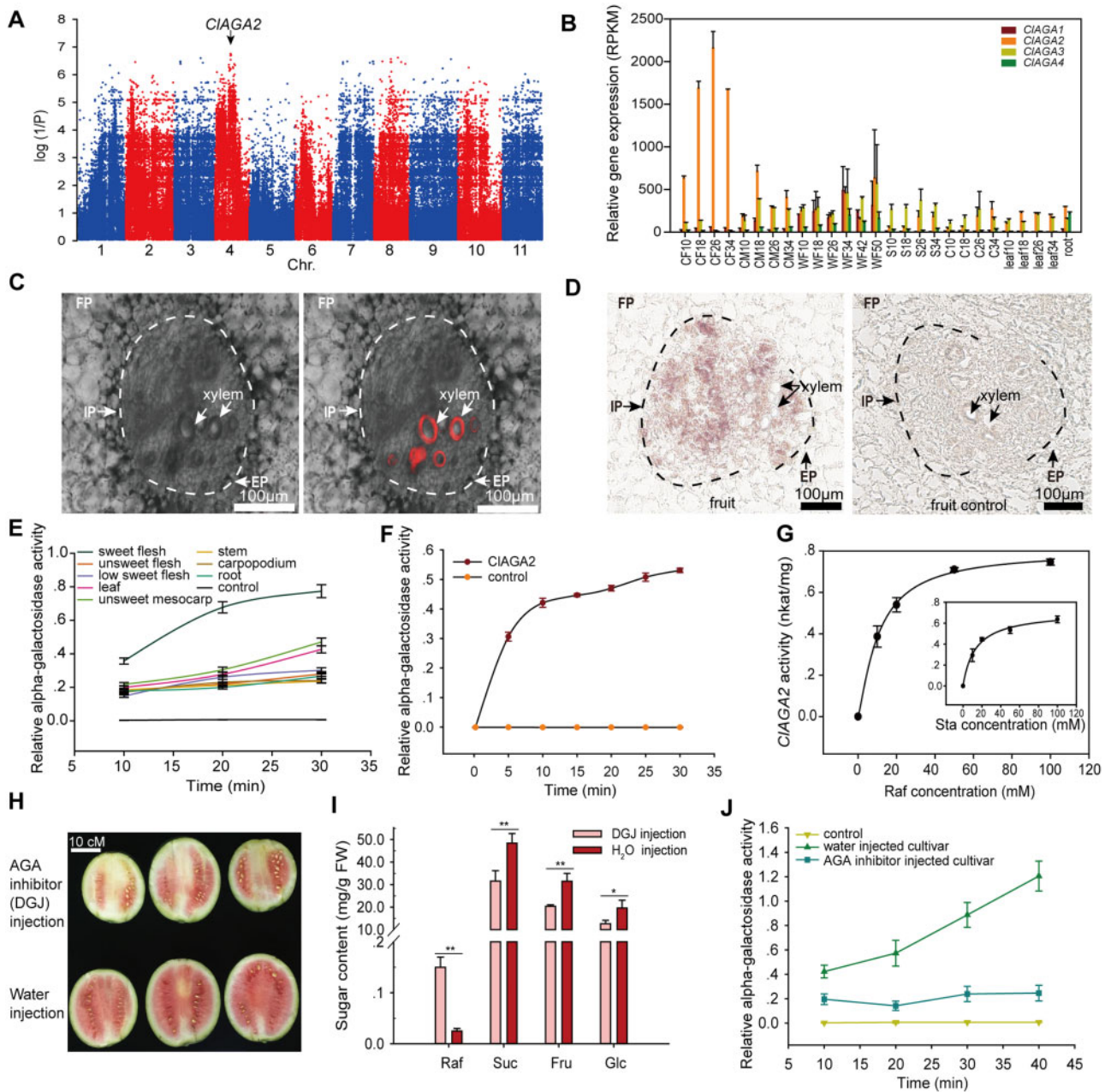


Figure 2 Mapping, expression pattern and functional analysis of *CIAGA2*. A, GWAS of the Raf content trait in 135 watermelon accessions. The significantly associated SNPs, marked with a black arrow, are located in the promoter of *CIAGA2*. B, RNA-Seq expression profiles of *CIAGA* family genes at different developmental stages in CF from 10 to 34 DAP, cultivated mesocarp (CM) from 10 to 34 DAP, nonsweet wild accession fruit from 10 to 50 DAP, stem (S) from 10 to 34 DAP, carpodium (C) from 10 to 34 DAP, leaf from 10 to 34 DAP, and in root. Average values \pm SD for $n = 2$ independent fruit replicates are shown, three technical replicates for each fruit were performed. C, Bright field (left) and RFP emission merged channels (right). The red areas indicate the xylem zone labeled by Texas red dextran in vascular bundles of watermelon fruit. D, In situ hybridization of *CIAGA2* expression in the watermelon fruit (left). Vascular bundles are labeled, with the IP and EP marked with black arrows. Vascular parenchyma cells are outlined by the dotted line; FP, fruit parenchyma cells. The sense probe was used for in situ hybridization in fruit as a negative control (right). E and F, Relative alpha-galactosidase activity in different watermelon tissues (E) and in vitro expressed *CIAGA2* (F). Average values \pm SD for $n = 3$ independent replicates of watermelon tissues or *CIAGA2* enzyme. G, Kinetics of *CIAGA2* enzyme activity with 50 mM Raf or Sta as substrates. Error bars, SD for $n = 3$ independent in vitro expressed *CIAGA2* enzyme. H and J, Fruit maturity (H), sugar content (I), and alpha-galactosidase activity (J) after injection of alpha-galactosidase inhibitor DGJ. Three independent fruits at 18 DAP for each treatment are shown. Bar = 10 cm. Average values \pm SD for $n = 3$ independent biological replicates are shown. **t*-test significant at $P < 0.05$, ***t*-test significant at $P < 0.01$.

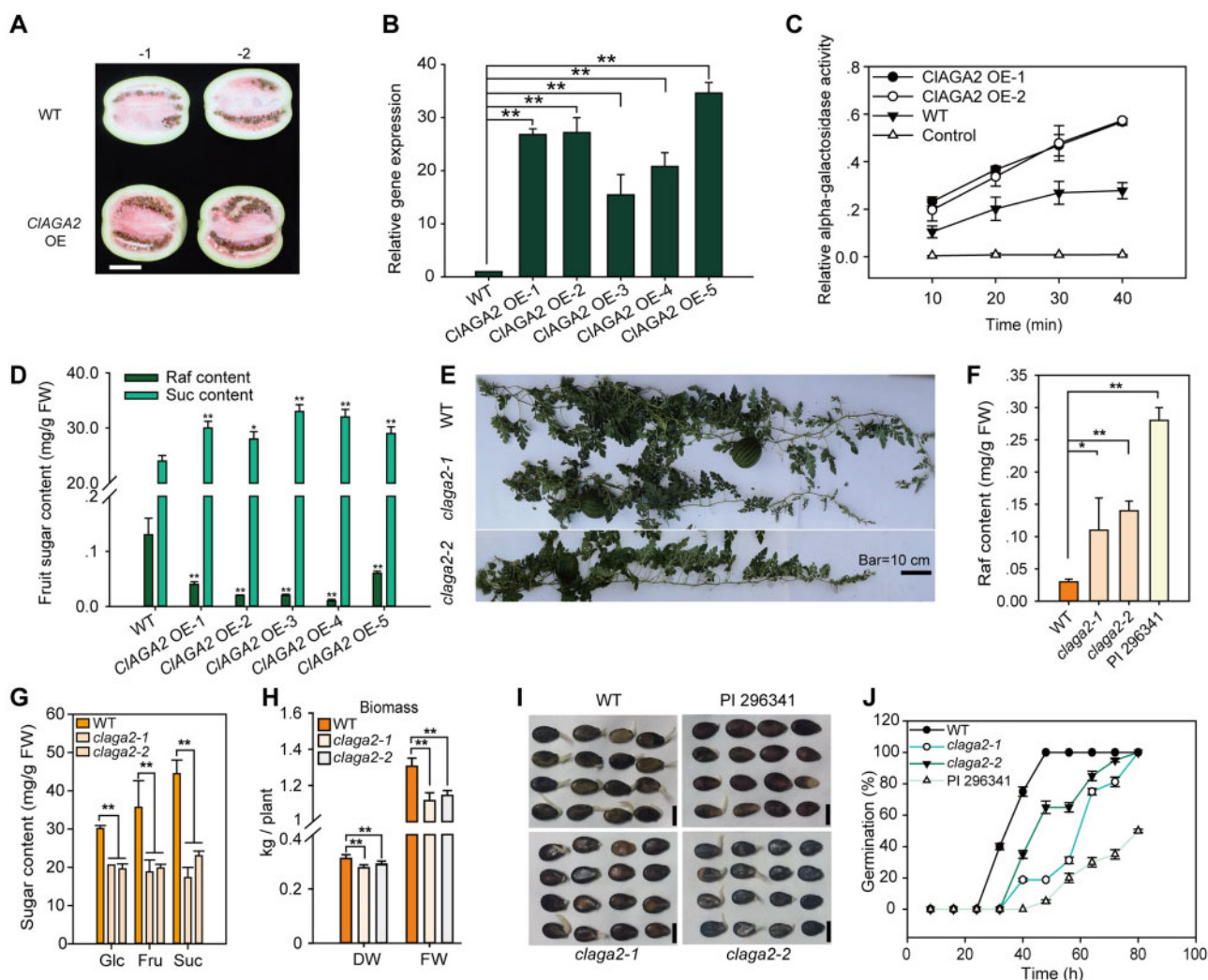


Figure 3 *CIAGA2* enhances sugar partition in watermelon. **A**, *CIAGA2* overexpression accelerates fruit coloration. Two independent fruits at 34 DAP for each are shown. Bar = 10 cm. **B**, Relative mRNA levels of *CIAGA2* in OE lines and WT at 34 DAP. Means \pm SD of $n = 3$ independent replicates are shown. ***t*-test significant at $P < 0.01$ for mRNA levels of *CIAGA2* OE lines compared to WT. **C** and **D**, *CIAGA2* overexpression increased alpha-galactosidase activity (**C**) and accelerated Raf hydrolyzation and sugar accumulation (**D**). Means \pm SD of $n = 3$ independent replicates are shown. **t*-test significant at $P < 0.05$ and ***t*-test significant at $P < 0.01$ for sugar content of *CIAGA2* OE lines compared to WT. **E**, Two lines of *claga2* mutant plants at 18 DAP. **F**, Raf content in the *claga2* mutants, WT and PI 296341. Means \pm SD of $n = 2$ independent replicates are shown. **t*-test significant at $P < 0.05$ and ***t*-test significant at $P < 0.01$. **G**, Fruit Glc, Fru, and Suc contents in *claga2* mutants at 30 DAP. Average values \pm SD for $n = 3$ independent biological replicates are shown. ***t*-test significant at $P < 0.01$. **H**, Biomass measured on FW and DW of the *claga2* mutants and WT. **t*-test significant at $P < 0.05$, ***t*-test significant at $P < 0.01$. Means \pm SD of $n = 3$ independent replicates are shown. *P*-values were calculated by two-tailed, unpaired *t*-tests. **I**, Germination of WT, *claga2* mutant, and wild species PI 296341 seeds incubated under dark conditions at 28°C for 80 h. Bars = 2 cm. **J**, Seed germination rates of *claga2* mutants and WT. Means \pm SD of $n = 3$ independent replicates are shown.

claga2 fruits increased almost to the level observed in wild PI 296341 (Figure 3F). Approximately 35%–50% decreases in Glc, Fru, and Suc contents were observed in *claga2* Cas9-free homozygous plants (Figure 3G). The plant biomass of *claga2* was reduced by 15% (Figure 3H). The *claga2* mutants exhibited delayed seed germination (Figure 3I) and a slower germination rate than did the WT plants (Figure 3J).

CINP-YC2 regulates the expression of *CIAGA2*

CIAGA2 exhibited different expression patterns in different tissues and during different stages of fruit development

(Figure 2B). We observed that the *CIAGA2* promoter (–1,512) has five SNPs between cultivated and wild watermelon accessions (Supplemental Figure 1). To identify the regulatory transcription factor controlling *CIAGA2* expression, we screened a watermelon fruit cDNA library using yeast one-hybrid (Y1H) assays and determined that the regulator transcription factor CINP-YC2 can activate the *CIAGA2* promoter (Supplemental Data Set 3). Furthermore, we determined that only the NF-YC family was predicted to bind the 18-bp flanking sequences of SNPs –159 (C-T) and –170 (C-T), while the significant SNPs –159 (C-T) and

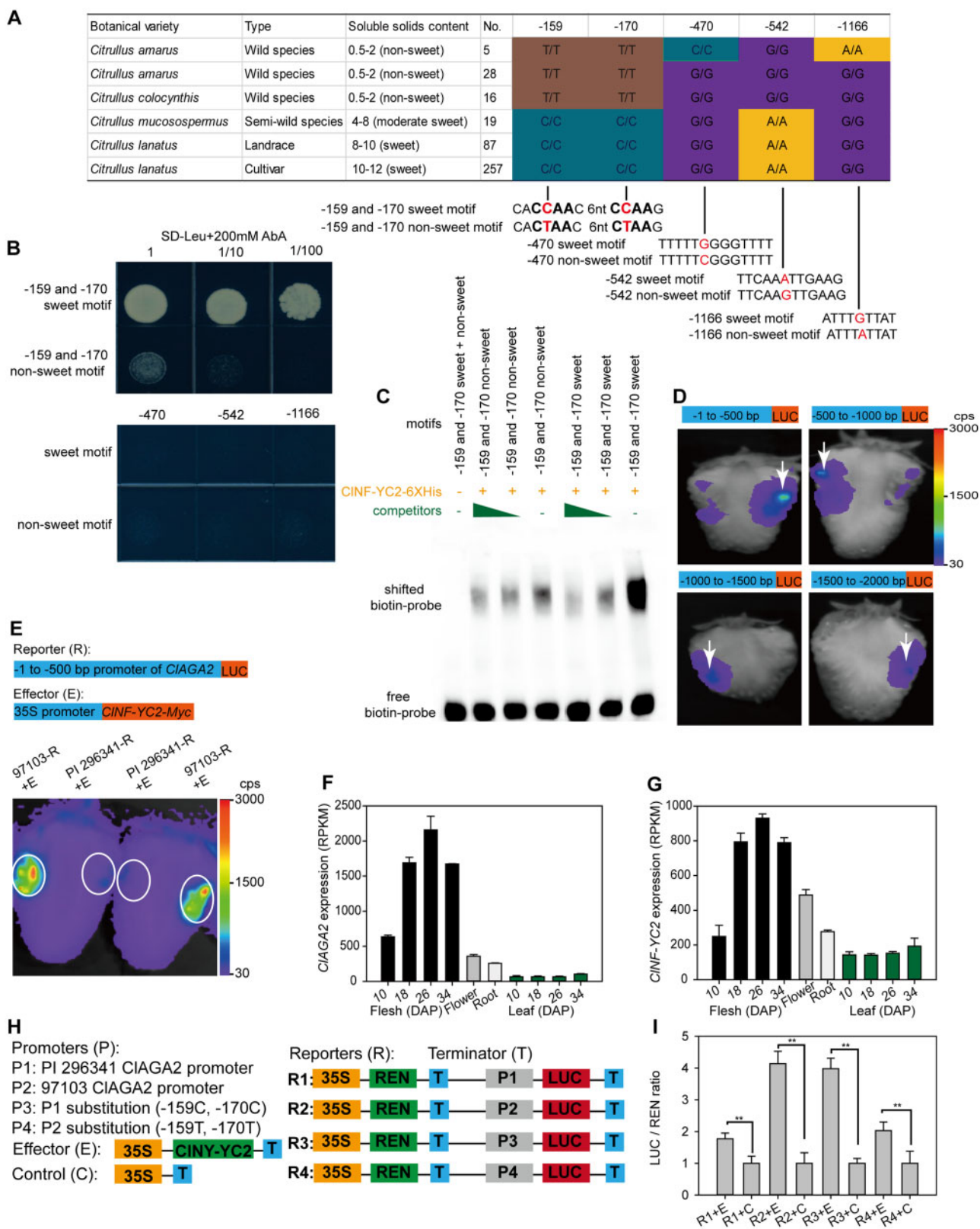


Figure 4 *CIN-YC2* regulates *CIAGA2*. **A**, Haplotype of *CIAGA2* gene promoter. Homozygous genotype “A,” “T,” “C,” and “G” was shown as yellow, brown, cyan, and purple. **B**, Y1H experiment coexpressing transcription factor *CIN-YC2* with sweet or nonsweet motifs in *CIAGA2* promoter. **C**, EMSAs of *CIN-YC2* fusion protein binding to the cis-element in *CIAGA2* from sweet 97103 and nonsweet PI 296341. **D**, LUC activity in strawberry fruits expressing the four 5'-truncated fragments of the *CIAGA2* gene promoter reporter constructs. The white arrows indicate injection sites of *Agrobacterium*. Color in scale represents the fluorescence intensity in counts per second (cps). **E**, LUC assay of strawberry fruits cotransformed with *CIN-YC2* as the effector and the -500 bp promoter region of *CIAGA2*. Strong LUC activity was detected with the promoter of the sweet

–170 (C-T) in the promoter (Figure 4A) led to two changes in the predicted NF-YC2 recognition core site (Shi et al., 2014), while no transcription factor binding site was predicted for the other three SNPs, that is, –470 (G-C), –542 (A-G), and –1,166 (G-A). To determine whether CIN-FYC2 can bind to regulatory elements of the *CIAGA2* promoter and whether SNPs –159 (C-T) and –170 (C-T) affect the binding activity, we performed a Y1H experiment that measures protein–DNA interactions. For the Y1H assay, CIN-FYC2 was fused with the AUR1-C transcriptional activation domain to serve as prey. The –159 and –170 sweet or nonsweet motif *cis*-elements, in which the 18-bp flanking sequences of SNPs –159 (C-T) and –170 (C-T) obtained from sweet and nonsweet watermelon accessions were cloned into pAbAi-bait (Figure 4A). Notably, after cotransformation of yeast with prey and individual baits, the CIN-FYC2 protein activated AUR1-C expression only when the –159 and –170 sweet motifs from cultivar 97103 were expressed in the presence of 200 mM aureobasidin A (AbA) in Y1H cells, while no activation was observed for the –470, –542, and –1,166 SNP-containing motifs in the *CIAGA2* promoter (Figure 4B).

Additionally, we produced a His-CIN-FYC2 recombinant protein in *Escherichia coli* that was used as a probe in an electrophoretic mobility shift assay (EMSA). Bands with slowed migration were observed when both –159 and –170 sweet or nonsweet motifs were incubated with the His-CIN-FYC2 protein. The bands became significantly weaker in the presence of competitive probes, and no shifted probe band was detected when only the motifs were present (Figure 4C). As shown by EMSA, the band binding by the cultivar 97103 sweet motif (–159 and –170) *cis*-element was significantly enhanced compared to that of PI 296341 nonsweet motif containing SNPs –159 (C-T) and –170 (C-T) (Figure 4C).

To functionally determine the major expression-regulating regions, four 5′-truncated fragments of different lengths from the 5′-end of the *CIAGA2* promoter region were fused to the luciferase (LUC) protein gene and transiently expressed in strawberry fruits. LUC activity in strawberry fruits expressing the four 5′-truncated fragments of the *CIAGA2* promoter was measured. Notable LUC activity was detected in strawberry fruits when using the upstream 500-bp promoter region. Therefore, the DNA sequence elements regulating the expression of *CIAGA2* are likely located between –1 and –500 bp (Figure 4D). Notably, the two significantly associated SNPs (–159 and –170) were also located in this region. Next, to determine whether CIN-FYC2 is able to regulate the expression

of *CIAGA2* in plants, we performed cotransformation of strawberry fruits with *p35S::CIN-FYC2::tNOS* and with a –500-bp promoter region harboring the SNP –159 and –170 motifs of the sweet and unsweet lines. The LUC activity showed that the CIN-FYC2 protein could activate the upstream 500-bp promoter of cultivar 97103 *CIAGA2*, rather than the wild PI 296341 promoter (Figure 4E). The expression levels of *CIAGA2* were relatively higher in sink tissues than in sugar loading source tissue during plant development (Figure 4F). Consistent with this result, CIN-FYC2 showed a similar expression pattern to that of *CIAGA2* in sink and source tissues at different stages (Figure 4G). Furthermore, promoter substitution assays using a dual-LUC reporter system in watermelon fruit protoplasts demonstrated the essential role played by SNPs –159 (C-T) and –170 (C-T) in the promoter in orchestrating the transcriptional regulation of *CIAGA2* by CIN-FYC2 in cultivar and ancestor watermelons (Figure 4, H and I). In contrast to the wild watermelons, the cultivated watermelons that were examined all possessed a genotype for CIN-FYC2 binding, indicating the selection of the *CIAGA2* promoter during watermelon evolution. Taken together, these results indicated that significantly associated loci (SNPs –159 and –170) in motifs within the *CIAGA2* promoter were associated with differences in binding activity in cultivated sweet versus wild nonsweet watermelon accessions.

CISWEET3 imports hexose across the plasma membrane in FP cells

Mutation of the *CIAGA2* gene resulted in a backlog of Raf hydrolysis in fruit (Figure 3F), leading to increased Raf content in *claga2* fruits, almost reaching the content of the wild *C. amarus*. Metabolomic comparison between *claga2* and WT fruits demonstrated that starch and Suc metabolism-related sugars, such as Suc, Glc, UDP-Glc, and trehalose-6-phosphate, were significantly downregulated in *claga2* fruits (Supplemental Data Set 4). To identify the possible genes causing reduced metabolites in the Suc metabolism pathway, we subsequently compared gene expression profiles between the fruits of *claga2* mutants and WT plants using the transcriptome and identified 177 upregulated genes and 295 downregulated genes in *claga2* mutants (Supplemental Data Set 5). The transcriptome data indicated that genes encoding key enzymes in Suc metabolism pathways, such as hexokinase (*Clag97C03G063270*), trehalose (Tre) phosphate synthase (*Clag97C11G223240*), and glycoside hydrolase (*Clag97C08G145190*), were downregulated in the fruit of *claga2* mutants (Supplemental Figure 2). The sugar

Figure 4 (Continued)

(97103), while no LUC activity was detected with PI 296341-*pCIAGA2::LUC* construction. Two representative results of five independent experiments are shown. The white circles indicate sites for *Agrobacterium*-mediated infiltration. Scale shows the fluorescence intensity from 30 to 3000 cps. F, Relative expression levels of *CIAGA2* in sugar loading tissue (leaf) and unloading tissues (fruit, flower, and root) during plant development. Error bars, SD values for $n = 3$ independent biological replicates. G, RNA-seq analysis of expression levels of CIN-FYC2 in sugar loading and unloading tissues at different stages. H, Diagram of the various constructs with the –1,512 bp promoter used for dual-LUC reporter assays. I, The reporter and effector vectors were cotransformed into watermelon fruit protoplasts. LUC and REN (Renilla) LUC activities were assayed, and the activation of CIN-FYC2 to the reporters was shown by the ratios of LUC to REN. The LUC/REN ratios of the empty vector cotransformed with reporters were set as 1 for controls. Each value represents the mean \pm SD of five biological replicates. ***t*-test significant at $P < 0.01$.

transporter *C. lanatus* Sugars Will Eventually Be Exported Transporter 3 (*CISWEET3*; *Cl97C01G000640*) and vacuolar sugar transporter (*CIVST1*; *Cl97C02G031010*) were downregulated in the fruit of *claga2* mutants, while trehalases (*Cl97C01G001930* and *Cl97C01G001950*) were upregulated (Supplemental Data Set 5, Figure 5A).

Analysis of downregulated genes associated with carbohydrate partitioning in the *claga2* mutants initially implicated sugar transporter gene *CISWEET3*, which could be involved in the sugar partitioning process (Figure 5A). To further identify genes responsible for sugar partitioning in fruit cells, we evaluated the transcript levels of orthologs to known sugar transporters in watermelon during fruit development. RNA-Seq data showed that *CISWEET3* was the most highly induced sugar transporter in the fruit during sweet watermelon fruit development (Supplemental Figure 3), with mRNA levels increasing notably from 10 to 34 DAP during sugar content increase (Figure 5B), indicating that the transcript level of *CISWEET3* was positively correlated with sugar content. Moreover, in the 100 core germplasms selected from 1,197 accessions (Zhang et al., 2016), the *CISWEET3* mRNA levels were correlated with sugar content (Figure 5C). To elucidate the cellular expression pattern, we collected fruit phloem and surrounding FP cell tissues using laser capture microdissection (LCM). The results of real time quantitative polymerase chain reaction (RT-qPCR) showed that *CISWEET3* was expressed in FP cells (fruit storage cells) in watermelon fruit, and its expression was not detected in the phloem (Figure 5D). Furthermore, we determined the plasma membrane localization of *CISWEET3* by fusing *CISWEET3* with the enhanced yellow fluorescent protein (EYFP) coding sequence and overexpressing the fused protein in watermelon fruit cells (Figure 5E). The hexose transport activity of *CISWEET3* was initially obtained by monitoring the time-dependent uptake of [¹⁴C]-Glc and [¹⁴C]-Fru in *Xenopus* oocytes (Figure 5F). Moreover, *CISWEET3* enabled the yeast mutant EBY4000 lacking all 18 hexose transporters (Wieczorke et al., 1999) to uptake hexose and grow on Glc and Fru selection plates (Figure 5G). Taken together, these results indicate that *CISWEET3* is a plasma membrane-localized hexose transporter in watermelon FP cells.

Our results suggest that increased expression of *CISWEET3* may be the reason for sugar accumulation. To validate this possibility, we overexpressed *CISWEET3* from cultivar 97103 in white-fleshed semi-wild PI 179878. A total of 21 overexpression (OE) transgenic T₀ plants were harvested, and two of the T₂ lines (*CISWEET3* OE-1 and *CISWEET3* OE-2) with increased *CISWEET3* mRNA levels and one copy of the transgene were selected for further analysis. Transgenic plants that did not show Basta resistance during the seeding period in the same T₂ segregation population were utilized as negative WT controls. In two representative overexpressors of *CISWEET3*, the fruit turned red (Figure 5H), and sugar contents were increased in watermelon fruit (Figure 5I) compared to the semi-wild PI 179878 WT at 30

DAP. Meanwhile, we generated two CRISPR-*clsweet3* mutants, with one harboring a “T” insertion between two protospacer-adjacent motif sites and the other containing a “G” insertion in the first exon of the coding region (Figure 5J), leading to a frameshift to generate a null mutant allele. Shorter and fewer branches were observed in *clsweet3* (Figure 5J) with homozygous frameshift mutants in cultivar zhong zi jing mu (ZZJM). The plant biomass of dry weight (DW) and fresh weight (FW) at 2 weeks of age in the two lines of *clsweet3* mutant plants showed a reduction of 14% and 10% at 30 DAP, respectively (Figure 5K). Approximately 10%–15% decreases in Suc, Glc, and Fru contents were observed in *clsweet3* Cas9-free homozygous plants, while no significant difference was observed for Raf content (Figure 5L). Furthermore, to investigate whether *CISWEET3* is a selected gene during domestication, we analyzed both cross-population composite likelihood ratio (XP-CLR) and π values to demonstrate that *CISWEET3* was located in a domesticated region between the closest ancestor semi-wild and cultivated watermelon groups (Supplemental Figure 4).

CIAGA2, CISWEET3, and CITST2 exhibit different effects on sugar accumulation

Our results suggest that AGA2 determines RFO hydrolysis and that *CISWEET3* uptakes hexose into fruit cells from the intercellular space. Finally, sugars are stored in the fruit vacuoles by the vacuolar membrane-localized CITST2 (Tonoplast Sugar Transporter), as we reported previously through genetic mapping, functional study, and overexpression experiments (Ren et al., 2018). Moreover, in this study, to further characterize the role played by CITST2 and compare it with the activities of CIAGA2 and *CISWEET3*, we generated two *cltst2* mutants with 5-nt and 1-nt deletions, which resulted in frameshifted *CITST2*, using CRISPR/Cas9 in cultivar ZZJM. Fewer branches with poor growth vigor were also observed in *cltst2* mutants (Figure 6A). The two lines of *cltst2* mutants displayed delayed fruit coloration (Figure 6B) and decreased sugar content by ~35% compared to WT at 30 DAP (Figure 6C). The results of RT-qPCR sampled from the LCM-collected phloem and FP tissues showed *CITST2* was dominantly expressed in FP cells in watermelon fruit, exhibiting ~20-fold less expression in the phloem (Figure 6D). These results strongly suggest that CITST2 is a key factor leading to sugar accumulation in vacuoles of watermelon fruit.

In further comparing the roles played by CITST2, CIAGA2, and *CISWEET3* and assessing the extent to which the RFO hydrolysis (CIAGA2) and sugar transport (*CISWEET3* and CITST2) processes contribute to fruit sugar accumulation, we observed that the ratio of fruit biomass to the whole plant was significantly reduced in *claga2*, *cltst2*, and *clsweet3* mutants compared to WT; however, the ratio was notably decreased in two lines of *claga2* and *cltst2* mutants compared to *clsweet3* mutants (Figure 6E). Both the Suc and hexose contents of *claga2* and *cltst2* were similarly lower than those of the *clsweet3* mutants (Figures 3, H, 5, L, and 6,

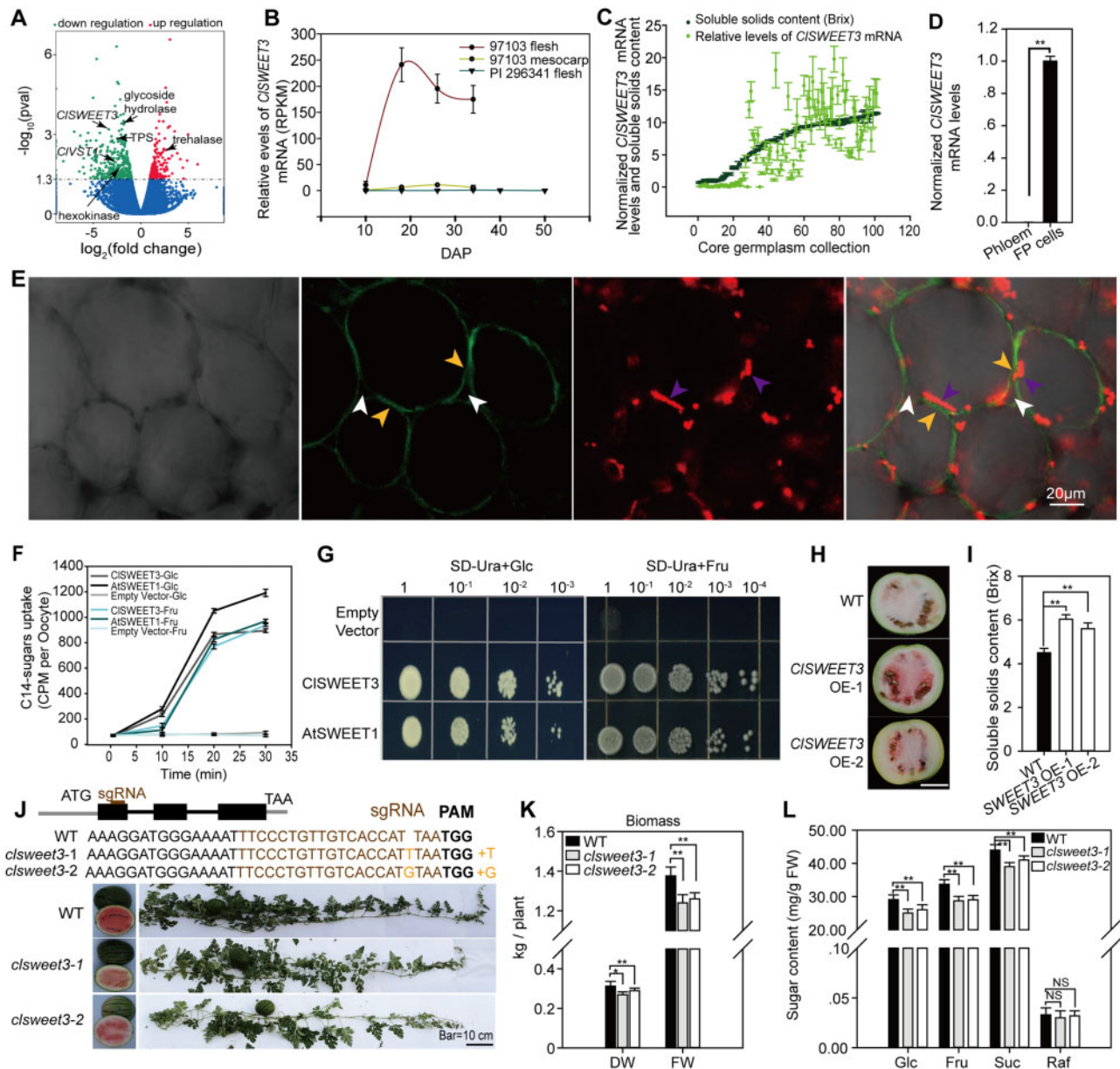


Figure 5 *CISWEET3* accelerates sugar accumulation. **A**, Transcriptome profiling showing upregulated and downregulated genes in *claga2* knockout fruit compared to WT. Differently expressed genes in Suc metabolism pathway, Suc transporter *CIVST1* and hexose transporter *CISWEET3* are indicated by black arrows. **B**, Expression profiles of *CISWEET3* in fruit and mesocarp of sweet watermelon 97103 and fruit of non-sweet PI 296341. DAP, days after pollination. Error bars, SD for $n = 3$ independent biological replicates of RNA-seq data. **C**, *CISWEET3* mRNA levels and soluble solids content levels in 100 core germplasms. Error bars, SD for $n = 3$ fruit samples. **D**, *CISWEET3* mRNA levels in FP and phloem cells. Means $\pm \text{SD}$ of $n = 3$ LCM collected tissue samples from fruit are shown. P -values were calculated by two-tailed, unpaired t -tests. **E**, Plasma membrane localization of *CISWEET3*-EYFP in watermelon fruit cells. Bright light, EYFP, RFP emission, and merged channels in *CISWEET3*-EYFP transgenic overexpressing fruit cells after plasmolysis. The white arrows point to the cell wall, the purple arrows point to the red fluorescence of the chromoplastid, indicating the position of the cytosol, and the yellow fluorescent arrows point to the *CISWEET3*-EYFP position on the plasma membrane separated from the cell wall. **F**, Time-point hexose uptake assay in *Xenopus* oocyte cells with [¹⁴C]-labeled Glc and Fru. Means of C¹⁴ intensity $\pm \text{SD}$ for $n = 5$ independent oocyte cells are shown. **G**, Complementation assay of yeast EB4000 lacking 18 hexose transporter genes with *CISWEET3*, *AtSWEET1*, and an empty vector. **H**, Two representative watermelon fruit of *CISWEET3* overexpressors. Bar = 10 cm. **I**, Soluble solids content in fruit of *CISWEET3* overexpressors and WT. Means $\pm \text{SD}$ for $n = 3$ fruits are shown. ** t -test significant at $P < 0.01$. **J**, CRISPR/Cas9-generated *clsweet3* alleles. The schematic diagram illustrates sgRNA targeting the first exon of *CISWEET3*, which results in a null mutation of *CISWEET3*. **K**, Biomass measured on the basis of FW and DW in the two lines of *clsweet3* mutants. Error bars, SD for $n = 3$ independent plants. * t -test significant at $P < 0.05$, ** t -test significant at $P < 0.01$. **L**, Fruit Suc, hexose and Raf contents in *clsweet3* mutants. Average values $\pm \text{SD}$ for $n = 3$ independent biological replicates are shown. ** t -test significant at $P < 0.01$.

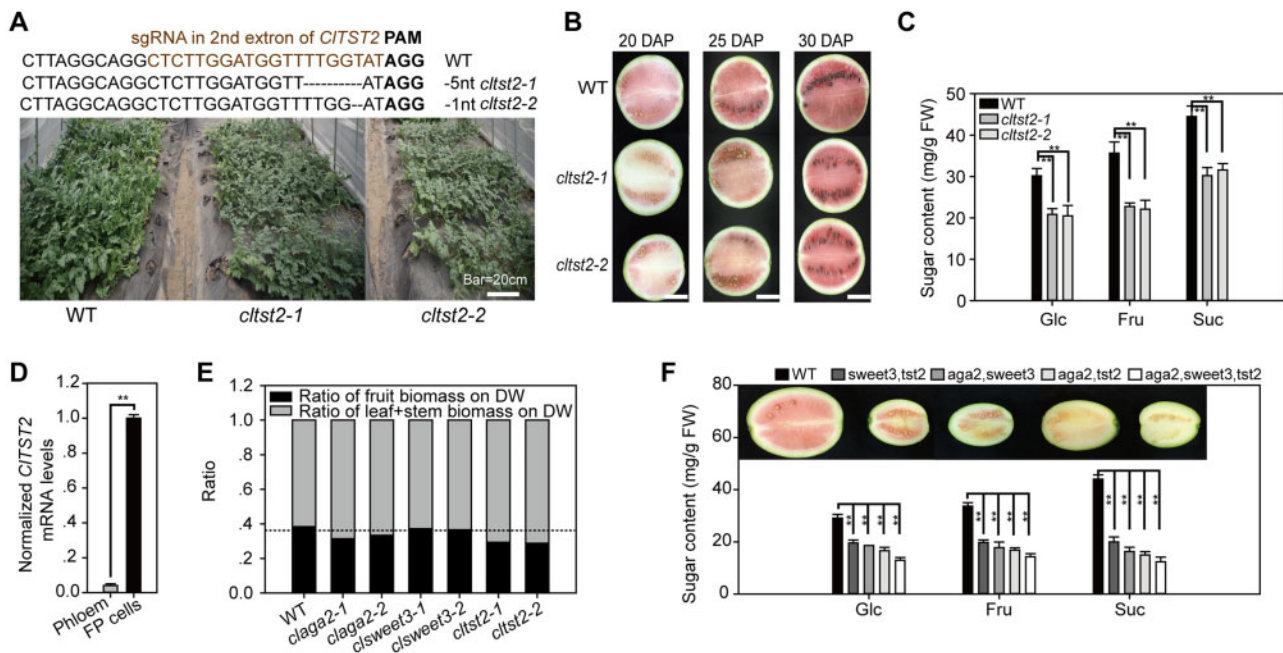


Figure 6 Different effects of *CIAGA2*, *CISWEET3*, and *CITST2*. **A**, CRISPR/Cas9-generated *cltst2* alleles and plants compared to the WT. **B**, Fruit development of *cltst2* single mutation lines and WT at 20, 25, and 30 DAP. Bar = 10 cm. **C**, Suc, Glc, and Fru content in *cltst2* mutants and WT at 30 DAP. Means \pm SD for $n = 3$ independent fruits are shown. ***t*-test significant at $P < 0.01$. **D**, *CITST2* mRNA levels in FP and phloem cells in watermelon fruit. Means \pm SD of $n = 3$ LCM collected tissue samples are shown. **Significant at $P < 0.01$ calculated by two-tailed, unpaired *t*-tests. **E**, Ratio of fruit and source fixed biomass. The horizontal dashed line indicates the fruit biomass ratio of *clsweet3*. **F**, Phenotype of double and triple mutants of *aga2*, *sweet3*, and *tst2*. Average values \pm SD for $n = 3$ independent biological replicates are shown in (C), (D), and (F). ***t*-test significant at $P < 0.01$.

C). Taken together, these results indicate that *CIAGA2* and *CITST2* play essential roles in the first step of RFO hydrolysis in the vascular bundle and the final step for sugar storage in vacuoles for sugar repartitioning between the source and sink. The weaker phenotype exhibited by *clsweet3* mutants may be due to a lack of Suc transport activity because watermelon primarily accumulates Suc in fruit. Furthermore, we generated double and triple mutants of *aga2*, *sweet3*, and *tst2* through crossing and segregation. As expected, the different double-mutant combinations showed a vigorous phenotype compared to each single mutant, exhibiting smaller fruit and lower sugar content (Figure 6F). The triple mutant displayed a more severe sugar content-related phenotype considerably smaller fruit size and less coloration than did the double mutants (Figure 6F). Therefore, further research is warranted to study double and triple mutants to determine the synergistic effect of these genes.

Discussion

RFO translocation, hydrolysis, and regulation

In watermelon and other Cucurbitaceae, unlike other major crops, RFOs are the primary translocated sugars in the vascular bundle (Webb and Gorham, 1964; Schmitz et al., 1987; Zhang et al., 2012); however, tissue RFOs are hydrolyzed, and subsequent sugar partitioning has not been fully characterized. Moreover, it has not been elucidated how RFO-translocated plants control RFO flux, utilization, and effects

in sink tissues. The *Sta* and *Raf* synthase genes *Cla017113* and *Cla015152* are highly expressed in watermelon leaves (Guo et al., 2015; <http://cucurbitgenomics.org/>), indicating that the synthesis of RFOs occurred in leaves. Next, RFOs are loaded into the phloem of source tissues through a polymer-trapping mechanism (Rennie and Turgeon, 2009). Therefore, in principle, the enzyme that hydrolyzes RFOs to Suc and Gal could play an essential role in the phloem of sink tissues for utilizing carbohydrates.

In this study, we designed an innovative reciprocal grafting of watermelon fruits to branch stems to determine that *Raf* hydrolysis is backlogged in nonsweet wild watermelons but is entirely hydrolyzed in sweet watermelons. Backlogging raffinose could decrease sink strength in nonsweet wild ancestors, while cultivars can increase sink strength through rapid hydrolysis of RFOs and transport Suc/hexose into vacuoles for storage and sequestration. RFOs are broken down into Suc and Gal by *CIAGA2* in the fruit vascular bundle. Sweet watermelons translocating RFOs in stems can transport more photoassimilates from source to sink for the rapid growth rate of fruits.

What is the effect of RFO translocation during fruit development? Most likely, the rapid increase in sugar content during the late stage of fruit development could explain this effect. In our previous study, the Suc content in watermelon fruit increased sharply over a short period, that is, from 18 to 26 DAP, and accumulated approximately half of the total sugar in this short period; considerably more biomass

accumulated in the fruit of cultivated sweet watermelon than in nonsweet wild accessions (Liu et al., 2013). Moreover, watermelon fruit weight increases ~2,500-fold from 2 g on the flowering day to ~5,000 g within 20 days, while cucumber fruit weight increases ~100-fold from 2.5 g to ~250 g, which only requires 8–10 days (Ma et al., 2018). Therefore, it is feasible to increase the quality and yield of fruit crops via the regulation of vascular RFO transport and hydrolysis.

In vascular plants, particularly in RFO-translocating plants, photosynthate redistribution is a complex process influenced by a large number of tissue-specific genes whose encoded proteins are located at points extending from the sieve element and companion cell to the vacuole of FP cells. To date, most studies related to carbohydrate partitioning have been limited to a single gene (Julius et al., 2017), rather than multiple genes as constituents within the entire system. One systematic study showed that the active microbiome regulates the gene expression of nine genes in the galactosyl oligosaccharide metabolism pathway in ripe watermelon fruits, and it is hypothesized that new phenotypes could be generated without modifying the watermelon genome (Saminathan et al., 2018). However, functional evidence for these genes is absent. The fact that the nonsweet ancestral watermelons cultivated in Africa for thousands of years have been selected to yield today's high-quality sweet watermelons motivated us to use forward population genetics to study domestication, as well as to combine CRISPR/Cas9 gene-editing technology and multi-omic datasets to assess carbohydrate partitioning in a similar manner as a recent study characterizing the domestication of fruit metabolism in tomato (Zhu et al., 2018).

RFO utilization plays an essential role in sugar and biomass accumulation in RFO-translocated crops. Therefore, it is necessary to elucidate the transcriptional regulation of RFOs. In our study, the two significant SNPs in the *CIAGA2* promoter were associated with changes in the NF-YC2 recognition core site, and NF-YC2 could activate the expression of *CIAGA2*. Previous reports have shown that NF-YC plays multiple essential roles in plant vegetative organ development, stress responses, and growth (Shi et al., 2014). The regulation of CIN-FYC2 and *CIAGA2* is complex, and further research should be conducted.

Proposed model for oligosaccharide hydrolysis and sugar transport

A large number of sugar transporters and metabolic enzymes are involved in sugar accumulation and partitioning (Julius et al., 2017). A model for sugar transport and accumulation from source to sink in sucrose-translocated plants was proposed (Lemoine et al., 2013). We integrated the results of this study with the findings of our previous studies and proposed a model for oligosaccharide hydrolysis and sugar transport in the watermelon fruit sink (Figure 7). First, Sta and Raf, the principal sugars transported in the stem vein (Figure 7, A and B), are translocated through

stem flow (Figure 7C) to the fruit vascular bundle (Figure 7D), and *CIAGA2* regulated by CIN-FYC2 hydrolyzes Sta and Raf into Suc and Gal in the fruit vascular bundle, rather than the stem or carpopodium vascular bundle (Figure 7E). The sugar transporter *CIVST1* was recently demonstrated to unload Suc from the fruit vascular bundle into the intercellular space (Ren et al., 2020). Suc is hydrolyzed into hexose by invertase in the intercellular area or cell wall (Miron and Schaffer, 1991; Roitsch and Gonzalez, 2004). Next, *CISWEET3* transports hexose into FP cells from the intercellular space (Figure 7F). Meanwhile, the unknown plasma membrane-localized Suc transporters (shown in gray) in FP cells may uptake Suc from the intercellular space (Figure 7F). Alternatively, Suc can be synthesized from UDP-glucose (UDPG) and Fru-6-phosphate by sucrose-phosphate synthase (SPS) and sucrose-phosphate phosphatase or synthesized from Fru and UDPG via sucrose synthase (SuSy) in the cytosol. RNA-Seq data showed that hexokinase (Cl97C03G063270) in the sucrose-phosphate synthetic pathway and phosphoglycerate mutase (Cl97C08G153870) were downregulated in *claga2* mutants with decreased sugar content (Supplemental Data Set 5). The activity of SPS and SuSy in the sucrose synthesis direction was upregulated, while watermelon fruit was accumulating sugars from 26 DAP to 34 DAP, and it was suggested that both SPS and SuSy play roles in increasing Suc synthesis in watermelon fruit (Liu et al., 2013). Therefore, in addition to importing Suc by plasma membrane-localized Suc transporters, another pathway of Suc production could be through SPS or SuSy in the cytosol (Figure 7F). Next, *CITST2* responds to Suc and hexose storage in fruit vacuoles by importing sugars across the tonoplast (Ren et al., 2018). All four genes in sugar partitioning have undergone selection, leading to the orchestrated derivation of sweetness in modern watermelons from nonsweet ancestors. We have reported that the plant hormone abscisic acid and transcription factors are involved in regulating watermelon ripening and sugar accumulation (Wang et al., 2017; Ren et al., 2018; Zhou et al., 2020). However, how these factors, as well as other possible factors, such as microRNAs (miRNAs), long noncoding RNAs (lncRNAs), and DNA methylation, regulate these genes in sugar partitioning warrants further investigation (Lang et al., 2017; Ma et al., 2020). Overexpression of *CmTST2*, a melon ortholog of *CITST2*, could increase sugar accumulation in melon fruits (Cheng et al., 2018). Whether the orthologous genes of *CIAGA2*, *CIVST1*, *CISWEET3*, and *CITST2* play similar roles or underwent selection during evolution and domestication in other cucurbits warrants further research.

Evolution of oligosaccharide hydrolysis and sugar partitioning

The evolution of modern cultivated crops involved in effective sugar partitioning from source to human-edible tissues, such as fruit, seed, and root, has provided valuable knowledge for improving crop yield and quality in rice, maize, and tomato (Julius et al., 2017). Recently, it was reported

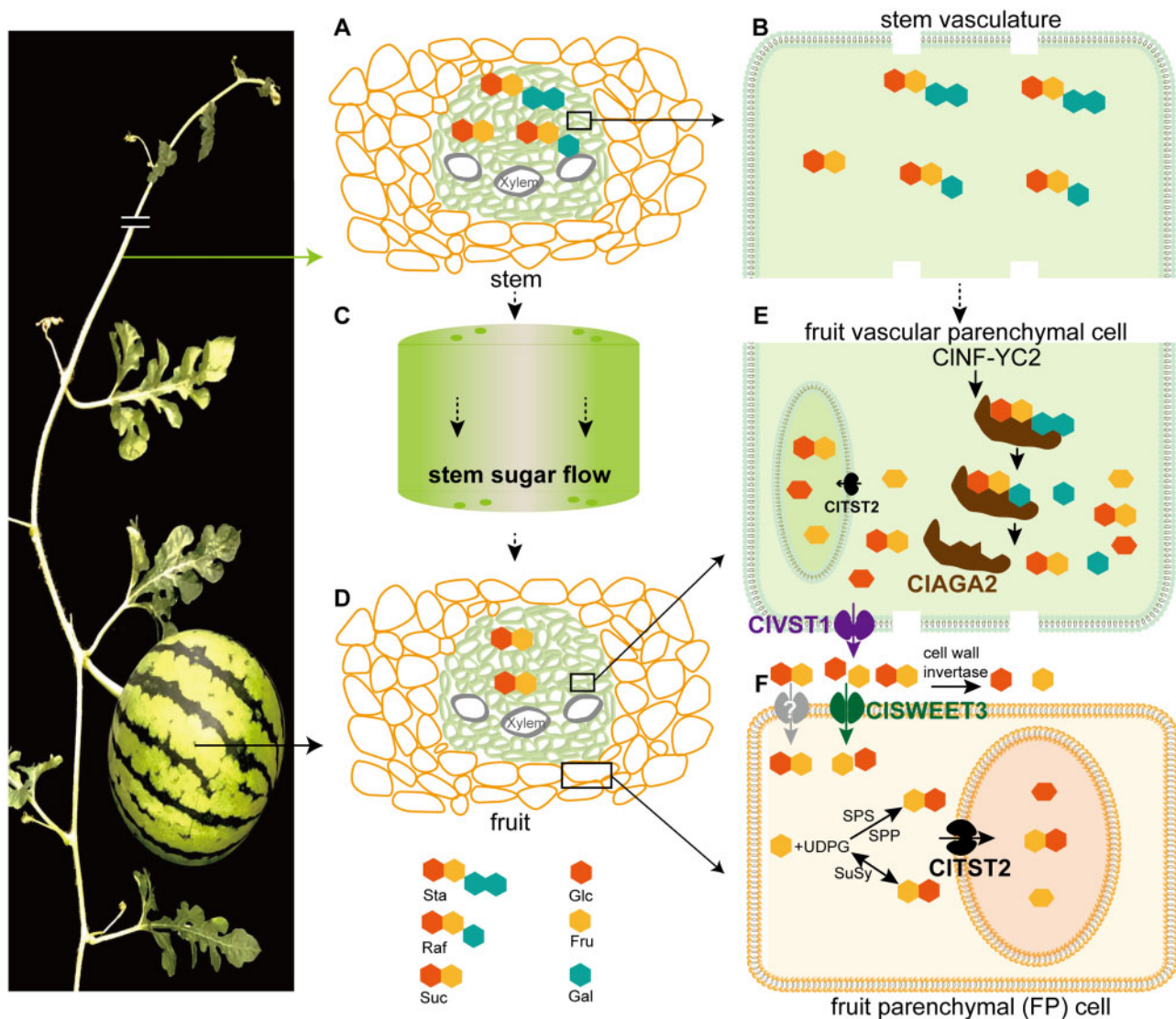


Figure 7 Proposed model for oligosaccharide hydrolysis and sugar transport in watermelon. A, Sta and Raf, along with Suc, are the main translocated sugars in the stem. B, The outstretched vascular parenchymal cell in (A). C, Sugar flow in stem. D, The fruit vascular bundle and FP cell. E, RFOs are hydrolyzed by CIAGA2 in fruit vascular parenchymal cells. CIVST1 unloads Suc from fruit vascular parenchymal cells into the intercellular space (Ren et al., 2020). F, CISWEET3 transports hexose to FP cells for storage from the intercellular space, and CITST2 accumulates sugars in vacuoles (Ren et al., 2018).

that simultaneous changes in sugar content, seed size, and oil content are driven by the domestication of SWEET family sugar transporters in soybean (Wang et al., 2020). Sugar content is an essential breeding trait for fruit crops, and it is important to characterize the evolution of oligosaccharide hydrolysis and sugar partitioning.

In our study, considerably more sugar and biomass accumulated in modern watermelon fruit than in its wild ancestor due to evolution. CC, the origin ancestor of sweet watermelon, yields small, white, and bitter fruits. Fruit size enlargement occurred in the closest ancestor CM and in the modern watermelon *C. lanatus* (Supplemental Figure 5). Hydrolysis and transport of sugars in sweet watermelons are clear gain-of-function processes during the speciation of CM from CC. Raf levels are considerably high in most of the

nonsweet wild watermelons, whereas in sweet watermelons, Raf is hydrolyzed to accumulate sugars, indicating that RFO hydrolysis and sugar accumulation are domesticated traits. Moreover, high levels of Raf in most nonsweet wild watermelons indicate that a low level of CIAGA2 mRNA could also provide necessary energy for fruit and seed development.

Nucleotide diversity (π), F_{ST} values and a cross-population composite likelihood ratio test (XP-CLR) showed that CIAGA2 was located in the selective sweep regions during the speciation process between the wild ancestor CC and the semi-wild CM (Guo et al., 2019). Semi-wild CM consists of both sweet and nonsweet watermelons and is believed to be the closest ancestor of cultivated watermelon (Guo et al., 2013). During the domestication process between the closest ancestor CM and cultivar CL group, CIVST1 was located in

the selective sweep regions (Ren et al., 2020). Furthermore, *CISWEET3* (Supplemental Figure 4) and *CITST2* were also selected during watermelon domestication from the closest ancestor CM to the cultivar CL group, while genes coding SPS, invertase, and sucrose synthase were not under selection (Guo et al., 2019). These results support that *CIAGA2*, *CIVST1*, *CISWEET3*, and *CITST2* were selected for during evolution and domestication, which attempted to increase the hydrolysis of RFOs, as well as subsequent sugar transport within the watermelon fruit. According to the selection and knockout phenotype, we proposed a model for the evolution of *CIAGA2* and domestication of *CIVST1*, *CISWEET3*, and *CITST2* in watermelon (Supplemental Figure 5). The nonsweet and small fruit-producing wild ancestor CC was initially selected at the *CIAGA2* locus during the process of obtaining moderate sweetness and larger fruit size in the semi-wild CM, while *CIVST1*, *CISWEET3*, and *CITST2* were selected during domestication from the semi-wild to cultivated *C. lanatus* to further increase fruit sugar accumulation and improve fruit quality (Supplemental Figure 5).

Owing to high divergence in the genome of wild watermelons and linkage drag, introgressing disease-resistant genes from wild ancestors into cultivars was impeded. Given that yield and sugar accumulation selection are critical traits for breeding selection, a molecular tool could be developed for breeding selection based on the polymorphic loci in *CIAGA2*, *CIVST1*, *CISWEET3*, and *CITST2*. Thus, our study not only helps to elucidate oligosaccharide hydrolysis and sugar repartitioning in other crops but also provides a potential molecular tool for fruit quality improvement and breeding.

Materials and methods

Plant materials and GWAS

The cultivated line 97103 (*C. lanatus*) with high sugar content and the nonsweet wild accession PI 296341 (*C. amarus*) growing in the field in Beijing under a temperature range of 20°C–25°C, were employed for reciprocal grafting. Briefly, young fruits of 97103 at 3 DAP were grafted to the lateral branches of PI 296341 by inclined cutting and vice versa (Figure 1C). Grafting sites were covered with plastic film to maintain moisture and avoid contamination for 5 days and subsequently fixed with a clamp for 15 days. In total, 40 ovaries were grafted with a survival ratio of ~50%. At 30 days after grafting, grafted 97103 and PI 296341 fruits were harvested for GC–MS-based metabolite analysis (Lisec et al., 2006). The GWAS population collected worldwide and consisting of 135 representative watermelon accessions was planted in the field in Beijing where the temperature ranged from 10°C to 35°C in the autumn season from July to October and three biological replicates sampled from the center fruit were performed for phenotyping in October. The Raf contents of fruits in the GWAS panel were investigated using ion chromatography (Thermo ICS-3000). Brix was measured using a pocket refractometer (model pal-1; ATAGO Co., Ltd., Tokyo, Japan) from the center juice of each watermelon fruit. The same center juice was used for

Fru, Suc, and Glc content detection using HPLC (Shimadzu Co., Ltd., Japan). The released genome resequencing data of the watermelon accessions (Guo et al., 2019) were used for raffinose content GWAS. In total, 394,304 SNPs (Supplemental Data Sets 6 and 7) were obtained from 135 watermelon accessions by filtering with the following criteria: minor allele frequency (MAF) > 0.05, maximal heterozygous proportion = 0.05, maximal missing data proportion = 0.01. For population structure (Q) analysis, the SNP sites were further required to be four-fold degenerated sites with MAF > 0.05 and no pairwise $r^2 > 0.7$ in the 1-Mb surrounding region. The STRUCTURE software was run 20 times for k from 1 to 15 and found a best $k = 2$ with the CLUMPAK pipeline (Pritchard et al., 2000). Subsequently, the Q model was generated using the fastStructure program with $k = 2$ (Supplemental Data Set 8; Raj et al., 2014). The kinship matrix was generated using the Centered_IBS method in TASSEL 5 software (version 5.2.67; Supplemental Data Set 9). The phenotype data in Supplemental Data Set 2 were used for GWAS. GWAS was performed by the no compressed MLM with “P3D” variance component estimation using the Q + K model to estimate the association in TASSEL 5 software (version 5.2.67; Bradbury et al., 2007).

Vascular exudate sampling

In cucurbits, it is possible to obtain large amounts of pure vascular exudates through cuts of the petiole or stem (Tetyuk et al., 2013). To collect vascular exudates from the stem and carpopodium in watermelon, we followed the described method (Zhang et al., 2010; Zhang et al., 2012; Hu et al., 2016). In brief, immediately after detopping of stems and carpopodia when the temperature in the greenhouse in Beijing ranged from 25°C to 27°C from 9:00 to 10:00 in the morning under natural lighting, the earliest exudate was removed by clean paper to avoid contamination from either broken tissue debris or cells. The subsequent exudation is unlikely to come from the cut cells on the surface after the initial fluid is blotted away (Zhang et al., 2012). The phloem could be blocked, and the flow decreased substantially within a period of ~1–5 min after cutting (Zhang et al., 2010); therefore, only ~20 μ L of exudate within one minute was collected for each cut in both wild accession PI 296341 and cultivar 97103. In total, ~1 mL exudate from ~50 cuts per plant was collected. Average values of sugar content in exudate for three independent plants were used. The subsequent exudate was collected by pipette and immediately transferred to ice-cold microcentrifuge tubes and stored at –80°C until processing (Zhang et al., 2010). The sugar concentrations in stem and carpopodium vascular exudate were confirmed by using the 20 mM ethylenediamine tetraacetic acid (EDTA) solution collection method (King and Zeevaert, 1974; Tetyuk et al., 2013). The sugar content in vascular exudates was measured by ion chromatography (Thermo ICS-3000).

RNA extraction and RT-quantitative PCR

Total fruit RNA was extracted using the Quick RNA isolation kit (Huayueyang Biotechnologies Co., Ltd). LCM was performed to collect fruit phloem and surrounding FP cell tissues (as shown in [Figure 2D](#)) using a Leica LMD 6500. The Arcturus PicoPure RNA isolation Kit (Applied Biosystems) was used for total RNA extraction from the isolated fruit phloem and FP tissues. Potential DNA contamination was digested by DNase I (Qiagen). Five biological replicates sampled from fruits were used for each tissue. The TargetAmp™ 2-Round Aminoallyl-rRNA Amplification Kit 1.0 (Epicentre Biotechnologies) was used for pg level total cellular RNA amplification using T7-Oligo(dT) primer. cDNA was synthesized from total RNA using SuperScript III transcriptase (Invitrogen). SYBR Green I-based quantitative PCR was run on a Roche LightCycler 480. Five replicates were carried out using cDNA from five different fruits for each gene and the RT-qPCR experiment was performed twice. The watermelon *ACTIN* gene (ID: *Cl97C02G026960*) and *CYLS8* (ID: *Cl97C02G038590*) were used as the reference genes in RT-qPCR ([Kong et al., 2014](#)). The relative expression levels of *CIAGA2* (ID: *Cl97C04G070460*), *CISWEET3* (ID: *Cl97C01G000640*) and *CITST2* (ID: *Cl97C02G036390*) were normalized to those of the reference genes and averaged from the five fruit replicates. All gene sequences are available at the Cucurbit Genomics Database ([Zheng et al., 2019](#)).

DNA constructs

All genes were amplified with Phusion-HF DNA Polymerase (New England Biolabs, Ipswich, MA, USA). For these constructs, the *CIAGA2* and *CISWEET3* coding regions were amplified, and the stop codon was removed by PCR using gene-specific primers harboring the *attB1* and *attB2* sites ([Supplemental Data Set 10](#)) and then cloned via the BP enzyme reaction into pDONR221F1 (Invitrogen) followed by the LR reaction to transfer the gene sequence into the yeast expression vector pDRF-GFP or the plant p35S driving vector pX-EYFP_GW ([Chen et al., 2010b](#)). The *CISWEET3* gene was inserted between the *XbaI* and *XmaI* sites in the *Xenopus laevis* oocyte expression vector pGEMHE. The *CIAGA2* gene was fused in frame with a His tag in the *E. coli* expression vector pSKB3. gRNAs for the *CIAGA2* and *CISWEET3* genes were assembled into the CRISPR/Cas9 vector pBSE401 using the Golden Gate method ([Tian et al., 2017](#)).

Texas red dextran labeling

Xylem in the vascular system was labeled with Texas Red dextran solution (Invitrogen), as previously described ([Zhang et al., 2004](#)). We removed the cortical cell layer pedicel to avoid any damage to the vascular bundles. Afterward, the pedicel was encased by a cotton thread at one end, and the other end of the cotton thread was immersed in a tube with 200 mL (1 mg mL⁻¹) Texas Red dextran aqueous solution. Plants were allowed to translocate Texas Red dextran for 4 h. To distinguish the vascular system, the fruit tissues

were subsequently sectioned and examined using confocal laser scanning microscopy.

In situ hybridization

RNA in situ hybridization was performed using standard protocols ([Baker et al., 2016](#)) with slight modifications. Briefly, the specific coding sequence of *CIAGA2* was amplified from 97103 cDNA using Phusion Taq (Invitrogen), and the resulting products were ligated into the pSC-A-amp/kan vector using a StrataClone PCR cloning kit (Agilent Technologies). Specific probes were used for hybridization. Primers for specific probe synthesis are listed in [Supplemental Data Set 10](#). Sense and antisense probes were synthesized using SP6 and T7 RNA polymerases, respectively. The sense probe that cannot complement *CIAGA2* mRNA was used as the sense control. Fruits at 1 DAP were dissected by hand and fixed in 4% paraformaldehyde with 0.3% Triton X-100 for vacuum pumping for 20 min.

CIAGA2 purification and alkaline *alpha*-galactosidase activity assay

Recombinant alkaline *alpha*-galactosidase with a His tag was purified from *E. coli* total protein using nickel resin affinity (Invitrogen Life Technology). The resin was incubated in wash buffer (300 mM NaCl, 50 mM NaH₂PO₄, and 10 mM imidazole) before the experiment. A total of 100 µL resin was incubated with 200 µL *E. coli* total protein at 4°C with shaking for 3 h. The resin affinity *CIAGA2* protein was used for the activity assay since purification with elution buffer can result in loss of function. The alkaline *alpha*-galactosidase activity assay was performed with eight *p*-nitrophenol synthetic substrates, Raf or Sta (Sigma Aldrich, USA), as described by Gao et al ([Gao and Schaffer, 1999](#)) at pH 7.5 in extracts from *E. coli* expressing *CIAGA2*. Total protein of watermelon fruit, roots, and leaves was assayed for alkaline *alpha*-galactosidase activity.

Generation of overexpression and mutant lines

The watermelon explant transformation, plant regeneration, and greenhouse care were performed as described ([Tian et al., 2017](#)). *Agrobacterium tumefaciens* strain EHA105, which harbored the DNA constructs, was used for overexpression and mutant transformation. The 35S-driven plant vector pX-EYFP_GW ([Chen et al., 2010b](#)) was used for overexpression. Briefly, surface-sterilized watermelon seeds were cocultivated in the dark for four days. The regenerated adventitious buds were excised and transferred onto selective elongation medium containing 0.1 mg L⁻¹ 6 BA, 0.01 mg L⁻¹ NAA, and 2 mg L⁻¹ Basta. Constructs were designed to produce defined deletions within each target gene-coding sequence using two sgRNAs alongside the Cas9 endonuclease gene. The CRISPR/Cas9-positive lines were further genotyped for indel mutations using a forward primer to the left of the sgRNA and a reverse primer to the right of the sgRNA ([Supplemental Data Set 10](#)). The Cas9-free T₂ population and homozygous mutants in *claga2* or *clsweet3* were harvested for phenotyping. Transgenic plants that did not show

any Basta resistance and no editing in the same segregating population were used as negative controls. The potential off-target sites in *claga2*, *clsweet3*, and *cltst2* mutants were predicted by CRISPR2 (<http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR>). Primers that would bind to the flanking sequence of these potential off-target sites were designed (Supplemental Data Set 11). Subsequently, the potential off-target loci were amplified and sequenced to determine whether off-targets occurred within these sites. Plants with no off-targets were found in these sites (Supplemental Data Set 11). Thermal asymmetric interlaced (TAIL) PCR was used to identify the insertion position of T-DNA insertion using long-specific (LP) primers from the overexpressing vector pX-EYFP_GW and short arbitrary degenerate primers (Supplemental Data Set 8; Liu et al., 1995). Overexpressors with T-DNA insertions in intergenic regions were employed for phenotype analysis (Supplemental Data Set 12). All lines were planted in the field for phenotyping between 10°C and 35°C in the autumn season of Beijing, China.

Determining transgene copy numbers

We performed DNA gel blotting to detect the copy numbers of T-DNA insertions and the stability of *CIAGA2*- and *CISWEET3*-overexpressing cells (Supplemental Figure 6). Total DNA was isolated from young leaves of T_1 - and T_2 -overexpressing plants that had Basta resistance, and 10 µg of total genomic DNA was digested with *Xba*I. The *CIAGA2*- or *CISWEET3*-overexpressing vector digested with *Eco*RV or *Bgl*II, respectively, was used as a positive control, and genomic DNA from WT watermelon was used as a negative control. The digested DNA was transferred to a Hybond-N+ positively charged nylon membrane (Amersham Pharmacia Biotech) after separation in a 0.8% agarose gel. A digoxigenin (DIG)-labeled probe corresponding to the bar gene was prepared by PCR, and hybridization was performed using a DIG-High Prime DNA Labeling and Detection Starter Kit II (Roche). Subsequently, DNA gel blot analysis proved that DNA containing a single-copy bar gene was used as standard DNA for SYBR Green I-based real-time PCR on a Roche LightCycler 480. DNA isolated from the remaining T_1 and T_2 overexpressors of *CIAGA2* OE-1, *CIAGA2* OE-2, *CISWEET3* OE-1, and *CISWEET3* OE-2, which showed Basta resistance, was utilized to detect and confirm the single-copy number by real-time PCR by comparing their threshold cycle values to that of equal amounts of DNA gel blot proving standard DNA containing a single-copy bar gene (Yuan et al., 2007) (Supplemental Data Set 13).

Y1H assay

Yeast cells were cotransformed with the *pAbAi* bait vector harboring the -1,512-bp promoter of *CIAGA2* and the pGADT7-Rec prey vector harboring the ORFs of the watermelon fruit cDNA library. As negative controls, we transformed yeast cells with empty pGADT7 and *pAbAi* vectors harboring the corresponding promoter. Transformed yeast cells were selected on SD-Leu-Ura medium plates

supplemented with 0.2 M AbA (Sigma Aldrich) in OE Biotech Co., Ltd. (Shanghai, China). The plates were incubated for 3–4 days at 30°C, and positive clones were sequenced. Identified candidates were confirmed by one-to-one Y1H with pGADT7-Rec prey vector harboring the -1,512-bp promoter of *CIAGA2* or motifs.

EMSA

The full-length *CINF*-YC2 ORF was fused in frame with His and expressed in *E. coli*. The *CINF*-YC2-His protein was purified by nickel resin affinity (Invitrogen Life Technology). EMSA was performed using a commercial kit (Thermo Fisher) according to the manufacturer's instructions. The 5'-end biotin-labeled motifs were used as probes, and the two-fold concentration sequences without biotin were used as competitors. Binding reactions were carried out using 10 and 20 ng of *CINF*-YC2-His protein and 10 ng of each biotin-labeled promoter motif at 25°C for 30 min. The DNA-protein complexes on 5% polyacrylamide gels were transferred to a nylon filter that was exposed to ultraviolet light to crosslink the samples. The membrane was then soaked in blocking buffer for 15 min and washed three times. We captured the pictures with an Azure Biosystems c300 imaging system.

LUC protein expression assays in strawberry fruits and watermelon fruit protoplasts

Agrobacterium-mediated infiltration of strawberry fruits was carried out according to the protocol of Chen et al. (Chen et al., 2011). For LUC activity detection, *p35S::CINF*-YC2::*tNOS* was used as the effector, and 97103-*pCIAGA2::LUC* or PI 296341-*pCIAGA2::LUC* cell suspensions were mixed in equal volumes and coinfiltrated into white stage strawberry fruits using a needle syringe. Forty hours after infiltration, strawberry fruits were sprayed with luciferin solution and kept in the dark for 5 min to quench fluorescence. We captured images with a low-light cooled charged coupled device (CCD) imaging apparatus (NightSHADE LB985 with Indigo software). Five independent experiments were repeated two times. The double-LUC reporter assay vector contained a firefly LUC driven by 97103 or PI 296341 *CIAGA2* promoters or substituted promoters and an internal control renilla LUC (REN) driven by the 35S promoter, which was modified based on the pGreenII 0800-LUC reporter vector (Hellens et al., 2005). The constructed effector and reporter plasmids were cotransformed into watermelon fruit protoplasts. After 16 h, LUC and REN LUC activities were measured using a dual-LUC assay kit (Promega) on a Tecan M1000 reader. The results were calculated using the ratio of LUC to REN. At least five biological repeats were assayed for each combination.

Metabolome analysis

The freeze-dried watermelon fruits were crushed using a mixer mill (MM 400, Retsch) with a zirconia bead for 3 min at 30 Hz. One hundred milligrams of powder was weighed

and extracted overnight at 4°C with 1.0 mL of 70% aqueous methanol. Following centrifugation at 10,000 g for 10 min, the extracts were absorbed and filtered before LC-MS analysis. The samples were analyzed using an liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS/MS) system (HPLC, Shim-pack UFLC SHIMADZU CBM30A system, www.shimadzu.com.cn/; MS, Applied Biosystems 6500 Q TRAP, www.appliedbiosystems.com.cn/). Triple quadrupole (QQQ) scans were performed on a triple quadrupole-linear ion trap mass spectrometer (Q TRAP), API 6500 Q TRAP LC/MS/MS. The system was equipped with an ESI Turbo Ion-Spray interface and was operated in a positive ion model controlled by Analyst 1.6.3 software (AB Sciex). QQQ scans were performed as multiple reaction monitoring (MRM) experiments with collision gas (nitrogen) set to 5 psi at MetWare Biological Science and Technology Co., Ltd. (Wuhan, China). The declustering potential and collision energy for individual MRM transitions were further optimized. A specific set of MRM transitions was monitored for each period according to the metabolites eluted within this period.

Differential transcriptome expression analysis

A total amount of 2 µg total RNA per sample of WT and *claga2* mutants isolated from center of 34 DAP fruits with three biological replicates was used to construct RNA-Seq libraries using the NEB Next Ultra™ RNA Library Prep Kit for Illumina (NEB, USA) following the manufacturer's recommendations. The prepared libraries were sequenced on an Illumina HiSeq 4000 platform and paired-end reads 150 bp in length were generated. Raw RNA-Seq reads were first filtered with in-house Perl scripts to remove reads fitting any of the following rules: (1) N content was higher than 10% and (2) low-quality base content was larger than 50%. Next, Trimmomatic software (Bolger et al., 2014; v0.33) was further used to trim out adapter and low-quality bases. The resulting cleaned reads were aligned to the 97103 reference genome (Guo et al., 2019; v2) using Tophat2 (Trapnell et al., 2012; v2.1.0) with default parameters allowing up to one mismatch. Gene expression levels were normalized to fragments per kilobase of transcript per million mapped fragments. Differential expression analysis between WT and *claga2* mutants was performed using the DESeq R package (Anders and Huber, 2010; 1.10.1). We set the threshold for significantly differentially expressed genes as adjusted *P*-value < 0.05 and $|\log_2(\text{foldchange})| \geq 1$ (Storey and Tibshirani, 2003).

Sugar uptake analysis of transformed yeast and *Xenopus oocyte* cells

For the sugar uptake assay, at 3 days after injection of the RNA, oocytes were incubated with 10 mM sugar (0.1 µCi µL⁻¹ [¹⁴C] sugar) solution, as described by Chen et al. (2012), at different time points. EBY4000 yeast strain (Wieczorke et al., 1999) lacking 18 hexose transporters was transformed with the pDRF1 vector harboring the *CISWEET3* ORF. Yeast cells transformed with the pDRF1 empty vector

were used as negative controls. Transformed yeast cells were grown in SD-Ura selection medium supplemented with 50 mM Glc or Fru (Sigma Aldrich, USA). The plates were incubated for 3 days at 30°C, and positive clones were confirmed and sequenced.

Deciphering positive selection

We employed XP-CLR software with 30-kb nonoverlapping windows to test the cross-population composite likelihood ratio (XP-CLR) test (Chen et al., 2010a). The allele frequency differentiation test by XP-CLR between two populations was performed using two models: the Brownian motion to model neutralized genetic drift and the deterministic model to approximate the selective sweep effect on nearby SNPs (Chen et al., 2010a). We designated windows with the top 1% of the genome-wide empirical distribution as positive selection regions. Nucleotide diversity (π) was calculated using a 50-kb sliding window, and selective sweep was defined as the top 5% of the whole genome sequence diversity.

Accession numbers

Raw transcriptome sequencing reads have been deposited into the National Center for Biotechnology Information (NCBI) sequence read archive (SRA) under accession no. PRJNA563969. All gene sequences are available at the Cucurbit Genomics Database for *ACTIN* (ID: *Clag97C02G026960*), *CIYLS8* (ID: *Clag97C02G038590*) *CIAGA2* (ID: *Clag97C04G070460*), *CISWEET3* (ID: *Clag97C01G000640*), and *CITST2* (ID: *Clag97C02G036390*) under each gene ID.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure 1. Sequences of the *CIAGA2* promoter (−1,512) in cultivated sweet watermelons (*C. lanatus*) and ancestral species *C. colocynthis* and *C. amarus*.

Supplemental Figure 2. Changes of metabolite and transcript levels in the starch and sucrose metabolism pathway in the *claga2* mutant.

Supplemental Figure 3. Expression profiles of sugar transporters in the watermelon genome during fruit development.

Supplemental Figure 4. XP-CLR and π values of *CISWEET3*.

Supplemental Figure 5. Model for the evolution and domestication of *CIAGA2*, *CIVST1*, *CISWEET3*, and *CITST2* in watermelon.

Supplemental Figure 6. DNA gel blot analysis for overexpressors of *CIAGA2* and *CISWEET3* indicated single copy T-DNA insertion in each overexpressor.

Supplemental Data Set 1. Metabolome and sugar content of fruit and mesocarp in cultivated watermelon 97103 and wild accession PI 296341.

Supplemental Data Set 2. List of 135 worldwide collected representative watermelon accessions used for GWAS for raffinose content.

Supplemental Data Set 3. Transcription factor CINF-YC2 (Cl97C09G176240) was identified by selecting watermelon fruit cDNA library using Y1H.

Supplemental Data Set 4. Relative content of significantly different carbohydrates, nucleotide, and derivatives between *claga2* mutant and WT fruit.

Supplemental Data Set 5. Up- and downregulated genes in the *claga2* mutants.

Supplemental Data Set 6. Genotype_for_Raf_GWAS_Chr1-5.

Supplemental Data Set 7. Genotype_for_Raf_GWAS_Chr6-11.

Supplemental Data Set 8. The population structure used for GWAS.

Supplemental Data Set 9. The kinship matrix used for GWAS.

Supplemental Data Set 10. Primers used in this study.

Supplemental Data Set 11. Sequencing to check the potential off-target loci in CRISPR/Cas9 mutants.

Supplemental Data Set 12. Sequencing PCR production of TAIL-PCR to identify the position of T-DNA insertion in overexpressors.

Supplemental Data Set 13. Estimating the copy numbers of bar gene in overexpressing plants using real-time PCR.

Acknowledgments

We thank Dr Zhangjun Fei for critical reading of this manuscript.

Funding

This research was supported by grants from the National Key R&D Program of China (2019YFD1000300), the Natural Science Foundation of Beijing Municipality (6182015), the National Natural Science Foundation of China (31772328, 31930096), the Beijing Scholar Program (BSP026), the Ministry of Agriculture and Rural Affairs of China (CARS-25), the Beijing Agriculture Innovation Consortium (BAIC10-2020), and the funding of the European Commission (SGA-CSA no. 664621 and no. 739582 under FPA no. 664620). Work at the Joint BioEnergy Institute was supported by the U.S. Department of Energy, Office of Science, Office of Biological, and Environmental Research through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

Conflict of interest statement. The authors declare that there is no conflict of interest.

References

- Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* **11**: R106
- Büchi R, Bachmann M, Keller F (1998) Carbohydrate metabolism in source leaves of sweet basil (*Ocimum basilicum* L.), a starch-storing and stachyose-translocating labiate. *J Plant Physiol* **153**: 308–315
- Baker RF, Leach KA, Boyer NR, Swyers MJ, Benitez-Alfonso Y, Skopelitis T, Luo A, Sylvester A, Jackson D, Braun DM (2016) Sucrose transporter *ZmSut1* expression and localization uncover new insights into sucrose phloem loading. *Plant Physiol* **172**: 1876–1898
- Bel AJEV (1993) Strategies of phloem loading. *Annu Rev Plant Physiol Plant Mol Biol* **44**: 253–281
- Blöchl A, Peterbauer T, Richter A (2007) Inhibition of raffinose oligosaccharide breakdown delays germination of pea seeds. *J Plant Physiol* **164**: 1093–1096
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**: 2633–2635
- Chen H, Patterson N, Reich D (2010a) Population differentiation as a test for selective sweeps. *Genome Res* **20**: 393–402
- Chen LQ, Hou BH, Lalonde S, Takanaga H, Hartung ML, Qu XQ, Guo WJ, Kim JG, Underwood W, Chaudhuri B, et al. (2010b) Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**: 527–532
- Chen LQ, Qu XQ, Hou BH, Sosso D, Osorio S, Fernie AR, Frommer WB (2012) Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science* **335**: 207–211
- Chen Q, Sun J, Zhai Q, Zhou W, Qi L, Xu L, Wang B, Chen R, Jiang H, Qi J, et al. (2011) The basic helix-loop-helix transcription factor MYC2 directly represses PLETHORA expression during jasmonate-mediated modulation of the root stem cell niche in *Arabidopsis*. *Plant Cell* **23**: 3335–3352
- Cheng J, Wen S, Xiao S, Lu B, Ma M, Bie Z (2018) Overexpression of the tonoplast sugar transporter *CmTST2* in melon fruit increases sugar accumulation. *J Exp Bot* **69**: 511–523
- Chomicki G, Renner SS (2015) Watermelon origin solved with molecular phylogenetics including Linnaean material: another example of museumics. *New Phytologist* **205**: 526–532
- Chrost B, Schmitz K (1997) Changes in soluble sugar and activity of α -galactosidases and acid invertase during muskmelon (*Cucumis melo* L.) fruit development. *J Plant Physiol* **151**: 41–50
- Dai N, Cohen S, Portnoy V, Tzuri G, Harel-Beja R, Pompan-Lotan M, Carmi N, Zhang G, Diber A, Pollock S, et al. (2011) Metabolism of soluble sugars in developing melon fruit: a global transcriptional view of the metabolic transition to sucrose accumulation. *Plant Mol Biol* **76**: 1–18
- Davidson A, Keller F, Turgeon R (2011) Phloem loading, plant growth form, and climate. *Protoplasma* **248**: 153–163
- Doudna JA, Charpentier E (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**: 1258096
- Gao Z, Schaffer AA (1999) A novel alkaline α -galactosidase from melon fruit with a substrate preference for raffinose. *Plant Physiol* **119**: 979–988
- Guo S, Sun H, Zhang H, Liu J, Ren Y, Gong G, Jiao C, Zheng Y, Yang W, Fei Z, et al. (2015) Comparative transcriptome analysis of cultivated and wild watermelon during fruit development. *PLoS One* **10**: e0130267
- Guo S, Zhao S, Sun H, Wang X, Wu S, Lin T, Ren Y, Gao L, Deng Y, Zhang J, et al. (2019) Resequencing of 414 cultivated and wild watermelon accessions identifies selection for fruit quality traits. *Nat Genet* **51**: 1616–1623
- Guo S, Zhang J, Sun H, Salse J, Lucas WJ, Zhang H, Zheng Y, Mao L, Ren Y, Wang Z, et al. (2013) The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat Genet* **45**: 51–58
- Hellens RP, Allan AC, Friel EN, Bolitho K, Grafton K, Templeton MD, Karunairetnam S, Gleave AP, Laing WA (2005) Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* **1**: 13
- Hepper FN (1990) Pharaoh's Flowers: The Botanical Treasures of Tutankhamun. HMSO, Chicago, IL, USA & London, UK
- Hu C, Ham BK, El-Shabrawi HM, Alexander D, Zhang D, Ryals J, Lucas WJ (2016) Proteomics and metabolomics analyses reveal the

- cucurbit sieve tube system as a complex metabolic space. *Plant J* **87**: 442–454
- Hu L, Sun H, Li R, Zhang L, Wang S, Sui X, Zhang Z** (2011) Phloem unloading follows an extensive apoplasmic pathway in cucumber (*Cucumis sativus* L.) fruit from anthesis to marketable maturing stage. *Plant Cell Environ* **34**: 1835–1848
- Julius BT, Leach KA, Tran TM, Mertz RA, Braun DM** (2017) Sugar transporters in plants: new insights and discoveries. *Plant Cell Physiol* **58**: 1442–1460
- Jung B, Ludewig F, Schulz A, Meißner G, Wöstefeld N, Flügge U-I, Pommerrenig B, Wirsching P, Sauer N, Koch W, et al.** (2015) Identification of the transporter responsible for sucrose accumulation in sugar beet taproots. *Nat Plants* **1**: 14001
- King RW, Zeevaert JA** (1974) Enhancement of Phloem exudation from cut petioles by chelating agents. *Plant Physiol* **53**: 96–103
- Kong Q, Yuan J, Gao L, Zhao S, Jiang W, Huang Y, Bie Z** (2014) Identification of suitable reference genes for gene expression normalization in qRT-PCR analysis in watermelon. *PLoS One* **9**: e90612
- Lang Z, Wang Y, Tang K, Tang D, Datsenka T, Cheng J, Zhang Y, Handa AK, Zhu J-K** (2017) Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. *Proc Natl Acad Sci U S A* **114**: E4511–E4519
- Lemoine R, La Camera S, Atanassova R, Dedaldechamp F, Allario T, Pourtau N, Bonnemain JL, Laloi M, Coutos-Thevenot P, Maurousset L, et al.** (2013) Source-to-sink transport of sugar and regulation by environmental factors. *Front Plant Sci* **4**: 272
- Lin IW, Sosso D, Chen LQ, Gase K, Kim SG, Kessler D, Klinkenberg PM, Gorder MK, Hou BH, Qu XQ, et al.** (2014) Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature* **508**: 546–549
- Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR** (2006) Gas chromatography mass spectrometry–based metabolite profiling in plants. *Nat Protocols* **1**: 387
- Liu J, Guo S, He H, Zhang H, Gong G, Ren Y, Xu Y** (2013) Dynamic characteristics of sugar accumulation and related enzyme activities in sweet and non-sweet watermelon fruits. *Acta Physiol Plant* **35**: 3213–3222
- Liu YG, Mitsukawa N, Oosumi T, Whittier RF** (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J* **8**: 457–463
- Ma L, Shi Y-N, Grierson D, Chen K-S** (2020) Research advance in regulation of fruit quality characteristics by microRNAs. *Food Qual Saf* **4**: 1–8
- Ma S, Li Y, Li X, Sui X, Zhang Z** (2018) Phloem unloading strategies and mechanisms in crop fruits. *J Plant Growth Regul* **38**: 494–500
- McCaskill A, Turgeon R** (2007) Phloem loading in *Verbascum phoeniceum* L. depends on the synthesis of raffinose-family oligosaccharides. *Proc Natl Acad Sci U S A* **104**: 19619–19624
- Miron D, Schaffer AA** (1991) Sucrose phosphate synthase, sucrose synthase, and invertase activities in developing fruit of *Lycopersicon esculentum* Mill. and the sucrose accumulating *lycopersicon hirsutum* Humb. and Bonpl. *Plant Physiol* **95**: 623–627
- Oliva R, Ji C, Atienza-Grande G, Huguet-Tapia JC, Perez-Quintero A, Li T, Eom J-S, Li C, Nguyen H, Liu B, et al.** (2019) Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat Biotechnol* **37**: 1344–1350
- Platt FM, Dwek RA, Butters TD** (2005) Imino sugar inhibitors for treating the lysosomal glycosphingolipidoses. *Glycobiology* **15**: 43R–52R.
- Pritchard JK, Stephens M, Donnelly P** (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959
- Raj A, Stephens M, Pritchard JK** (2014) fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **197**: 573–589
- Ren Y, Guo S, Zhang J, He H, Sun H, Tian S, Gong G, Zhang H, Levi A, Tadmor Y, et al.** (2018) A tonoplast sugar transporter underlies a sugar accumulation QTL in watermelon. *Plant Physiol* **176**: 836–850
- Ren Y, McGregor C, Zhang Y, Gong G, Zhang H, Guo S, Sun H, Cai W, Zhang J, Xu Y** (2014) An integrated genetic map based on four mapping populations and quantitative trait loci associated with economically important traits in watermelon (*Citrullus lanatus*). *BMC Plant Biol* **14**: 33
- Ren Y, Sun H, Zong M, Guo S, Ren Z, Zhao J, Li M, Zhang J, Tian S, Wang J, et al.** (2020) Localization shift of a sugar transporter contributes to phloem unloading in sweet watermelons. *New Phytol* **227**: 1858–1871
- Renner SS, Sousa A, Chomicki G** (2017) Chromosome numbers, Sudanese wild forms, and classification of the watermelon genus *Citrullus*, with 50 names allocated to seven biological species. *TAXON* **66**: 1393–1405
- Rennie EA, Turgeon R** (2009) A comprehensive picture of phloem loading strategies. *Proc Natl Acad Sci U S A* **106**: 14162–14167
- Roitsch T, Gonzalez MC** (2004) Function and regulation of plant invertases: sweet sensations. *Trends Plant Sci* **9**: 606–613.
- Saminathan T, García M, Ghimire B, Lopez C, Bodunrin A, Nimmakayala P, Abhuri VL, Levi A, Balagurusamy N, Reddy UK** (2018) Metagenomic and metatranscriptomic analyses of diverse watermelon cultivars reveal the role of fruit associated microbiome in carbohydrate metabolism and ripening of mature fruits. *Front Plant Sci* **9**: 4
- Schmitz K, Cuyper B, Moll M** (1987) Pathway of assimilate transfer between mesophyll cells and minor veins in leaves of *Cucumis melo* L. *Planta* **171**: 19–29
- Schulz E** (1991) Holocene environments in the central Sahara. *Hydrobiologia* **214**: 359–365
- Shi H, Ye T, Zhong B, Liu X, Jin R, Chan Z** (2014) AtHAP5A modulates freezing stress resistance in *Arabidopsis* through binding to CCAAT motif of AtXTH21. *New Phytol* **203**: 554–567
- Sonnewald U, Fernie AR** (2018) Next-generation strategies for understanding and influencing source–sink relations in crop plants. *Curr Opin Plant Biol* **43**: 63–70
- Sosso D, Luo D, Li QB, Sasse J, Yang J, Gendrot G, Suzuki M, Koch KE, McCarty DR, Chourey PS, et al.** (2015) Seed filling in domesticated and rice depends on SWEET-mediated hexose transport. *Nat Genet* **47**: 1489–1493
- Storey JD, Tibshirani R** (2003) Statistical significance for genome-wide studies. *Proc Natl Acad Sci U S A* **100**: 9440–9445
- Tetyuk O, Benning UF, Hoffmann-Benning S** (2013) Collection and analysis of *Arabidopsis* phloem exudates using the EDTA-facilitated Method. *J Vis Exp* **80**: e51111–e51111
- Tian S, Jiang L, Gao Q, Zhang J, Zong M, Zhang H, Ren Y, Guo S, Gong G, Liu F, Xu Y** (2017) Efficient CRISPR/Cas9-based gene knockout in watermelon. *Plant Cell Rep* **36**: 399–406
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L** (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protocols* **7**: 562
- Wang H, Yan S, Xin H, Huang W, Zhang H, Teng S, Yu YC, Fernie AR, Lu X, Li P, et al.** (2019) A subsidiary cell-localized glucose transporter promotes stomatal conductance and photosynthesis. *Plant Cell* **31**: 1328–1343
- Wang S, Liu S, Wang J, Yokosho K, Zhou B, Yu Y-C, Liu Z, Frommer WB, Ma JF, Chen L-Q, et al.** (2020) Simultaneous changes in seed size, oil content, and protein content driven by selection of SWEET homologues during soybean domestication. *Natl Sci Rev* **7**: 1776–1786
- Wang Y, Guo S, Tian S, Zhang J, Ren Y, Sun H, Gong G, Zhang H, Xu Y** (2017) Abscisic acid pathway involved in the regulation of watermelon fruit ripening and quality trait evolution. *PLOS One* **12**: e0179944
- Webb JA, Gorham PR** (1964) Translocation of photosynthetically assimilated C¹⁴ in straight-necked squash. *Plant Physiol* **39**: 663–672

- Wieczorke R, Krampe S, Weierstall T, Freidel K, Hollenberg CP, Boles E** (1999) Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. *FEBS Lett* **464**: 123–128
- Wingenter K, Schulz A, Wormit A, Wic S, Trentmann O, Hoermiller II, Heyer AG, Marten I, Hedrich R, Neuhaus HE** (2010) Increased activity of the vacuolar monosaccharide transporter TMT1 alters cellular sugar partitioning, sugar signaling, and seed yield in *Arabidopsis*. *Plant Physiol* **154**: 665–677
- Yamauchi Y, Ejiri Y, Tanaka K** (2002) Glycation by ascorbic acid causes loss of activity of ribulose-1,5-bisphosphate carboxylase/oxygenase and its increased susceptibility to proteases. *Plant Cell Physiol* **43**: 1334–1341
- Yuan JS, Burris J, Stewart NR, Mentewab A, Stewart CN** (2007) Statistical tools for transgene copy number estimation based on real-time PCR. *BMC Bioinformatics* **8**: S6
- Zhang B, Tolstikov V, Turnbull C, Hicks LM, Fiehn O** (2010) Divergent metabolome and proteome suggest functional independence of dual phloem transport systems in cucurbits. *Proc Natl Acad Sci U S A* **107**: 13532–13537
- Zhang C, Yu X, Ayre BG, Turgeon R** (2012) The origin and composition of cucurbit "phloem" exudate. *Plant Physiol* **158**: 1873–1882
- Zhang H, Fna J, Guo S, Ren Y, Gong G, Zhang J, Weng Y, Davis A, Xu Y** (2016) Genetic diversity, population structure, and formation of a core collection of 1197 *Citrullus* accessions. *Hortscience* **51**: 23–29
- Zhang LY, Peng YB, Pelleschi-Travier S, Fan Y, Lu YF, Lu YM, Gao XP, Shen YY, Delrot S, Zhang DP** (2004) Evidence for apoplasmic phloem unloading in developing apple fruit. *Plant Physiol* **135**: 574–586
- Zheng Y, Wu S, Bai Y, Sun H, Jiao C, Guo S, Zhao K, Blanca J, Zhang Z, Huang S, et al.** (2019) Cucurbit Genomics Database (CuGenDB): a central portal for comparative and functional genomics of cucurbit crops. *Nucleic Acids Res* **47**: D1128–D1136
- Zhou M, Guo S, Tian S, Zhang J, Ren Y, Gong G, Li C, Zhang H, Xu Y** (2020) Overexpression of the watermelon ethylene response factor ClERF069 in transgenic tomato resulted in delayed fruit ripening. *Horticult Plant J* **6**: 247–256
- 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