



## OPEN ACCESS

## EDITED BY

Qiusheng Kong,  
Huazhong Agricultural  
University, China

## REVIEWED BY

Hao Li,  
Northwest A&F University, China  
Zhongyuan Hu,  
Zhejiang University, China  
Yongbing Zhang,  
Hami Melon Research Centre, China

## \*CORRESPONDENCE

Nan He  
henan@caas.cn  
Wenge Liu  
liuwenge@caas.cn

## SPECIALTY SECTION

This article was submitted to  
Plant Bioinformatics,  
a section of the journal  
Frontiers in Plant Science

RECEIVED 19 October 2022

ACCEPTED 17 November 2022

PUBLISHED 06 December 2022

## CITATION

Gong C, Lu X, Zhu H, Anees M, He N  
and Liu W (2022) Genome-wide  
association study provides genetic  
insights into natural variation in  
watermelon rind thickness and  
single fruit weight.  
*Front. Plant Sci.* 13:1074145.  
doi: 10.3389/fpls.2022.1074145

## COPYRIGHT

© 2022 Gong, Lu, Zhu, Anees, He and  
Liu. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s)  
are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Genome-wide association study provides genetic insights into natural variation in watermelon rind thickness and single fruit weight

Chengsheng Gong<sup>1,2</sup>, Xuqiang Lu<sup>1</sup>, Hongju Zhu<sup>1</sup>,  
Muhammad Anees<sup>1</sup>, Nan He<sup>1\*</sup> and Wenge Liu<sup>1\*</sup>

<sup>1</sup>Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou, China,

<sup>2</sup>Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Institute of Vegetable Crops, Jiangsu Academy of Agricultural Sciences, Nanjing, China

Rind thickness and fruit weight are agronomic traits closely related to quality and yield, which have attracted much attention from consumers and breeders. However, the genetic mechanism of these two traits is still not well understood in natural populations. In this study, rind thickness and single fruit weight in 151 watermelon accessions were determined in 2019 and 2020, and genome-wide association analysis was performed by integrating phenotypic and genotype data. Abundant phenotypic variation was found in the test population, and the watermelon with thinner rind thickness tended to have smaller fruit weights. Five significant SNPs were closely associated with rind thickness on chromosome 2 by Genome-wide association study (GWAS), i.e., 32344170, 32321308, 32304738, 32328501, and 32311192. And there were 21 genes were annotated in the candidate interval, most notably, *Cla97C02G044160* belonged to the MADS family, and part of the genes in this family played an important role in regulating organ size, further analysis of gene structure, gene expression level, and phylogenetic tree showed that *Cla97C02G044160* was a candidate gene for regulating target traits. In conclusion, our study provides molecular insights into the natural variation of watermelon rind thickness and single fruit weight, meanwhile, providing data support for molecular marker-assisted breeding.

## KEYWORDS

watermelon, GWAS, rind thickness, single fruit weight, transcriptome

## Introduction

As one of the five most popular consumed fresh fruits, watermelon [*Citrullus lanatus* (Thunb.) Matsum. & Nakai var. *lanatus*] is widely cultivated in the world and brought considerable economic benefits (Guo et al., 2013; Gong et al., 2021b). With the increasing market demand for high-quality watermelons, research on important phenotypic traits such as fruit shape, flesh color, fruit sugar content, and metabolic traits has been deepened (Ren et al., 2017; Dou et al., 2018; Zhang et al., 2020; Gong et al., 2021a). Rind thickness and single fruit weight are two important agronomic traits that have attracted much attention in the market. Specifically, small-fruit watermelon with thinner rind tends to have better flavor and is an important component of high-quality watermelon in the market, while large-fruit watermelon with thicker rind thickness contributes to the increase of watermelon yield and transportation.

The research on the genetic basis of important agronomic traits is helpful to accelerate the modern breeding process (Ren et al., 2014). Forward genetics carries out genetic basis research based on phenotypic variation, and the relevant research is often carried out by genetic map, Bulk Segregant Analysis (BSA), and Genome-wide association study (GWAS). Advances have been made in gene mapping of single fruit weight and rind thickness through forwarding genetics, and key QTL has been obtained in cucumber, melon, watermelon, and other Cucurbitaceae crops (Pan et al., 2020). For instance, Zhong et al. (2017) obtained QTL loci linked to 12 important agronomic traits such as rind thickness based on the construction of a high-density genetic map of pumpkin, providing theoretical support for mining candidate genes and molecular marker breeding. Research on QTL linked to fruit thickness in wax gourds has also made progress (Liu et al., 2018). In watermelon, (Wi et al. 2012) obtained QTLs linked to fruit size, rind thickness, and other traits through different genetic populations; Ren et al. (2014) realized the main QTL loci for 12 agronomic traits such as fruit length and sugar content by integrating four genetic maps, including QTLs related to rind thickness and fruit weight, i.e., *RTH2-1*, *RTH5*, *RTH6*, *FWT2-1*, *FWT2-2*, *FWT3*, *FWT5-1*, and *FWT5-2*. In addition, forward genetics combined with transcriptome data is helpful for further mining and identifying key candidate genes that regulate target traits. Compared with genetic map and BSA, GWAS could be used to analyze the association between polymorphic nucleic acids and traits in a constructed population, and quickly identify target loci and possible regulatory genes associated with traits (Hamblin et al., 2011; Yano et al., 2016).

At present, important progress has been made in the research on the physiological and genetic basis of plant fruit weight and rind thickness, and some key candidate genes have been identified. The size of normal organs is usually affected by the coordination of cell proliferation and cell expansion (Horiguchi et al., 2006). Genes can regulate the weight of fruit,

rind thickness, and the size of other traits by regulating the content of auxin, ethylene, cytokinin, and other plant hormones (Le et al., 2001; Wilmoth et al., 2005; Di Marzo et al., 2020). In addition, transcription factors (TFs) can also play an important role in regulating the size of plant tissues and organs, notably, the MADS family belongs to TFs and plays an important role in regulating the size of tissues and organs. For instance, in *Arabidopsis*, the TF of MADS gene *GORDITA* (*GOA*) has the function of regulating fruit growth, and mutation or reduced expression of this gene results in larger fruit (Prasad et al., 2010); *SEEDSTICK* regulating cytokinin levels and other genes to controls fruit weight of *Arabidopsis* (Di Marzo et al., 2020); In addition, rind thickness and fruit size have been extensively studied in tomato, and the genes such as *SIARF7* were considered to be key candidate genes related to regulating tomato rind thickness (De Jong et al., 2009). Furthermore, the construction of metabolic regulatory networks based on gene expression levels is helpful to elucidate the genetic basis of important agronomic traits (Umer et al., 2020). However, the current research on the key candidate genes regulating target traits is still far from enough.

Research on the genetic basis of single fruit weight and rind thickness in watermelon can help meet the practical needs of production, progress has been made in the study of rind thickness and fruit weight, while the understanding of the variation and genetic basis of these two traits in the natural population is still insufficient. In the present work, variation analysis and correlation analysis were carried out after the single fruit weight and rind thickness were measured of 151 watermelon accessions in 2019 and 2020. Combined with high-quality SNP data, GWAS of the target traits was performed and significant SNPs were obtained. Further, combine bioinformatics analysis and transcriptome data mining to identify key candidate genes that regulate target traits. In conclusion, our study provides molecular insights into the natural variation of rind thickness and single fruit weight and contributes to marker-assisted breeding and new cultivars selection.

## Materials and methods

### Plant cultivation and management

151 watermelon accessions belong to cultivated cultivars (*C. lanatus cultivar* and *C. lanatus landrace*), with bright flesh color to facilitate the identification of rind thickness. These accessions were randomly selected from previous studies with resequencing data (Guo et al., 2019; Gong et al., 2022), and were collected in the National Mid-term Genebank for Watermelon and Melon, Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences. The plants were planted in Xinxiang (35.30°N, 113.88° E) in 2019 and Zhongmu (34.72°N, 113.98° E)

in 2020, respectively. The rind of 14 watermelon accessions with different thicknesses (grey background in Table S1) was sampled at the maturity period for qRT-PCR (Quantitative real-time polymerase chain reaction, qRT-PCR). In addition, two watermelon cultivars '97103' and '203Z' with different single fruit weights ('97103':3.2 kg, '203Z':2.0 kg) and rind thicknesses ('97103':1.1 cm, '203Z':0.5 cm) were mainly used for transcriptome determination and planted in Sanya, Hainan in the winter of 2021. And the plants were transplanted through seedling raising and according to the local climate conditions. Twin-vine pruning was used, and only one watermelon fruit per plant was retained. The main temperature and humidity in the greenhouse were as follows: during the daytime, the temperature was kept between 25°C to 35°C, and the humidity was 55%-70%, while at nighttime, the temperature was often above 15°C, and the humidity was maintained at 75% to 80%. Alternatively, irrigation, watering, and fertilization were managed according to the actual needs of watermelon growth and were managed in the same way. The cultivars were harvested according to their ripening characteristics, and then, the single fruit weight and rind thickness were measured.

## Determination of phenotypic data and sampling

For the 151 watermelon accessions, three mature watermelons with good growth were randomly selected from each cultivar for the determination of target traits, and the phenotype data after calculating the average value were used for further analysis. The electronic scale was used to measure the weight of the fruit, and after the watermelon fruit was cut lengthwise, the thickness of the rind was measured with a ruler. Watermelon accessions '97103' and '203Z' were used to perform transcriptome sequencing, and 14 watermelon accessions were used to qRT-PCR, the rind of watermelon without exocarp was sampled at mature period, and the watermelon was cut vertically into 1 cm thick slices along the largest diameter, and cuboid pericarp of about 1 cm in length and width was taken from the four vertical directions and mixed as a sample to be tested. Three watermelons of the same growth were individually sampled as triplicate biological replicates.

## GWAS and candidate gene identification

In the previous study, the linkage disequilibrium (LD) decay rates of the *C. lanatus* cultivar and *C. lanatus* landrace were relatively low (Guo et al., 2019), showing a high degree of domestication, and eliminating the influence of wild-type cultivars on genotype analysis due to differences in genetic background to a certain extent. The resequencing data of 126 watermelon accessions were obtained from (Guo et al. 2019)

(accession number WM in Table S1), and the other 25 were obtained from Gong et al. (2022) (accession number R in Table S1). The method of gene library construction and resequencing was consistent with a previous study (Guo et al., 2019). TruSeq Nano DNA High Throughput Library Prep Kit was used to construct about 400 bp sequences. Sequencing was performed on the Illumina HiSeq X or HiSeq 2000 platform at Berry Genomics (Beijing, China). And the watermelon genome 97103 v2 (<http://cucurbitgenomics.org/organism/21>) was used as the reference genome for the related analysis, and QQ plots were plotted to predict the reliability of the model. The Manhattan plot and QQ plot were drawn by R software. The SNP data of mutation detection were filtered according to the secondary allele frequency (MAF: 0.05) and locus integrity (INT: 0.8), and the high-quality SNP was obtained for further analysis. For GWAS, a Factored Spectrally Transformed linear mixed model (FaST-LMM) method was used for relevant analysis, and QQ plots were plotted to predict the reliability of the model, and EMMA eXpedited(EMMAX) and Linear Model are used to verify the reliability of the location results. The Manhattan plot and QQ plot were drawn by R software. And the modified Bonferroni correction was used to determine two significance thresholds, i.e.,  $-\log_{10}(P) 0.1/Ne$  ( $Ne =$  effective SNP number) and  $-\log_{10}(P) 0.01/Ne0.01/SNPs$  and  $0.1/SNPs$ , respectively. Criteria for identifying significant SNPs: when the SNP exceeded the top threshold in the Manhattan plot, it was considered to be highly significantly associated with the target trait and SNP. When the lower threshold line was exceeded, the correlation between the two was considered significant. When located below the threshold line, SNPs that are close to the threshold line and show a continuous distribution in a small interval were also considered may be associated with the target trait.

According to the analysis of LD decay in natural populations in previous studies (Guo et al., 2019), we took the interval of 100 kb upstream and downstream of the significant SNP locus as the candidate interval. Furthermore, combined gene annotation information, previous research reports, and transcriptome data to screen possible candidate genes. Gene sequence information was obtained from the online website of the cucurbitaceous crop genome (<http://www.cucurbitgenomics.org/>).

## Transcriptome sequencing and qRT-PCR

The rind samples of '97103' and '203Z' with three biological repeats were ground into powder in liquid nitrogen and then used for RNA extraction. RNA extraction was performed according to the instructions of the TRIzol Reagent (Life Technologies, California, USA). Qualified high-quality RNA was used for the next research, and NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, E7530) and NEBNext Multiplex Oligos for Illumina (NEB, E7500) was used for cDNA library construction. The cDNA library was sequenced at an Illumina HiSeq™ sequencing platform. The original sequencing data were filtered

to obtain clean data, which were compared with the watermelon genome 97103 v2, and further quality control was performed to obtain high-quality data. FPKM values (number of exon kilobases per million segments) were used to estimate gene expression levels (Florea et al., 2013). When the fold change of FPKM value between the two cultivars exceeded 1.8, the gene was defined as a differential gene. According to previous research reports, the genes related to seed expansion and hormones in watermelons were counted, and the heat map was drawn based on the expression level of Origin software (V 8.0.724).

Fourteen watermelon cultivars were set with three biological replicates. The powdered watermelon rind samples were extracted with the RNA isolation kit (Huayueyang Biotechnologies, China) for RNA extraction, and then synthesis of cDNA with NovoScript plus all-in-one 1st stand cDNA synthesis supermix (Novoprotein, China). Mix the solution to be tested using the SYBR Green real-time PCR mix, and an instrument LightCycler480 RT-PCR system (Roche, Swiss) was used for measurement. *CICAC* (Gene ID: *Clu97C09G174930*) was used as the reference gene. The original data obtained are calculated as the relative expression level by the  $2^{-\Delta\Delta CT}$  method (Livak and Schmittgen, 2001).

## Functional analysis of candidate genes

The construction of systematic evolution process: Amino acid sequence of the target gene was obtained from the watermelon reference genome 97103 v2. NCBI BLASTp (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to obtain homologous genes of the target gene, database selection 'UniProtKB/Swiss-Prot (Swissprot)', and then download the full-length sequence of amino acids for further analysis. MEGE7 (7.0.26) was used for sequence alignment and phylogenetic tree construction, ClustalW was used for amino acid sequence alignment, and Construct/Test neighbor-joining Tree was used for phylogenetic tree construction. The conservative domain was predicted by NCBI Conserved Domains: [https://www.ncbi.nlm.nih.gov/Structure/cdd/docs/cdd\\_search.html](https://www.ncbi.nlm.nih.gov/Structure/cdd/docs/cdd_search.html); the transmembrane structure analysis was predicted through an online website: TMHMM: <http://www.cbs.dtu.dk/services/TMHMM/>.

## Results

### Natural variation and correlation analysis of rind thickness and single fruit weight

In 2019 and 2020, the coefficient of variation of rind thickness in 151 watermelon accessions was 26.17% and 27.20%, respectively, while the coefficient of variation of single fruit weight was 28.01% and 23.02%, which suggests that these cultivated cultivars had abundant variations in these two agronomic traits. In the process of agricultural production, 0.6cm

is an important reference to determine whether watermelon has a thicker rind, statistical analysis showed that there were 14 and 15 accessions with a rind thickness of less than 0.6 cm, 128 and 132 accessions between 0.6-1.6 cm, and 9 and 4 accessions of more than 1.6 cm in two years, respectively. For single fruit weight, there were 19 and 11 accessions with less than 2 kg, 111 and 122 accessions of 2-4 kg, meanwhile, 21 and 18 accessions with more than 4 kg, respectively (Table S1). At the same time, most of the kurtosis and skewness of rind thickness and single fruit weight was less than 1 in two years (Table 1) and combined with the statistical histogram (Figures 1A, B, Figures 1D, E), it was concluded that these two important agronomic traits were basically in line with the characteristics of normal distribution in the population. Correlation analysis showed that epidermal thickness had a significant positive relationship in both years (Figure 1C), as did single fruit weight (Figure 1F), implying that phenotypic differences may be regulated by both genes and the environment, and were the relatively stable quantitative traits.

To explore the relationship between rind thickness and single fruit weight, a correlation analysis between the two traits was performed. The results indicate that there was a positive correlation between rind thickness and single fruit weight, and the correlation values of the two traits were 0.51 and 0.47 in 2019 and 2020, respectively (Figure S1). Notably, when the rind thickness was less than 0.6, the average single fruit weight was 2.09 kg (SD=0.53) and 2.40 kg (SD=0.59), respectively; while when the rind thickness was greater than or equal to 0.6, the fruit weight was 3.16 kg (SD=0.82) and 3.29 kg (SD=0.70). This difference indicates that there was a positive correlation between rind thickness and single fruit weight, especially for the watermelon with thinner rind thickness, which tends to have smaller fruits.

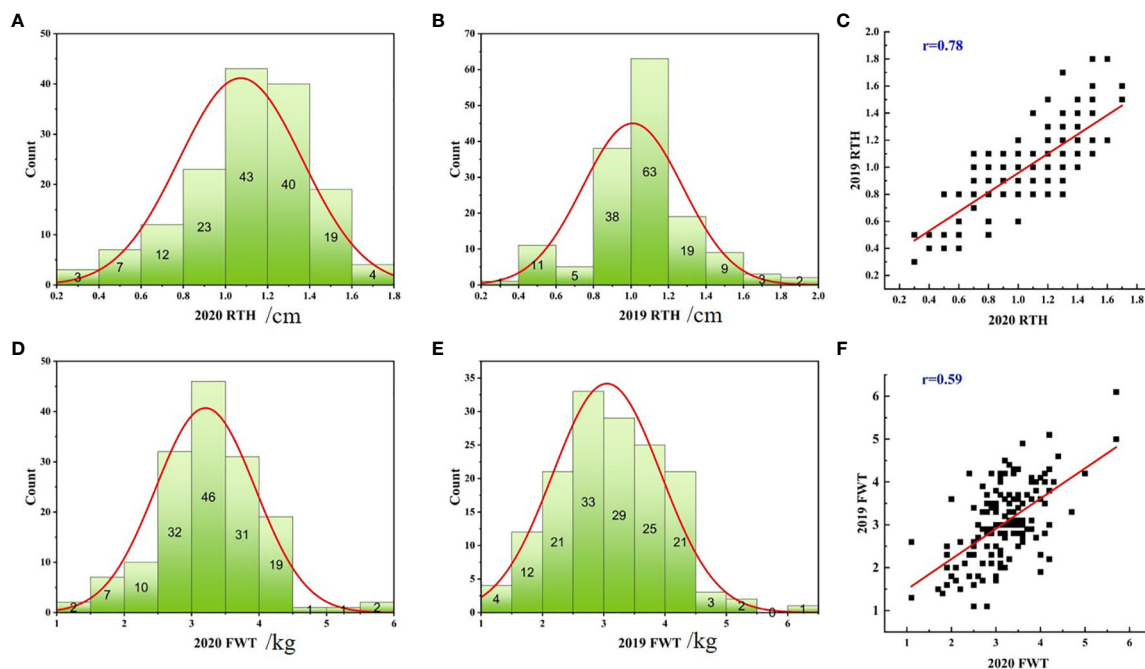
### GWAS and significant SNP acquisition of target traits

Significant SNPs were obtained by GWAS for rind thickness and single fruit weight in 2019 and 2020 (Figure 2), and the QQ plot indicates the reliability of the GWAS analytical model (Figure S2). The analysis by the FAST-LMM model of rind

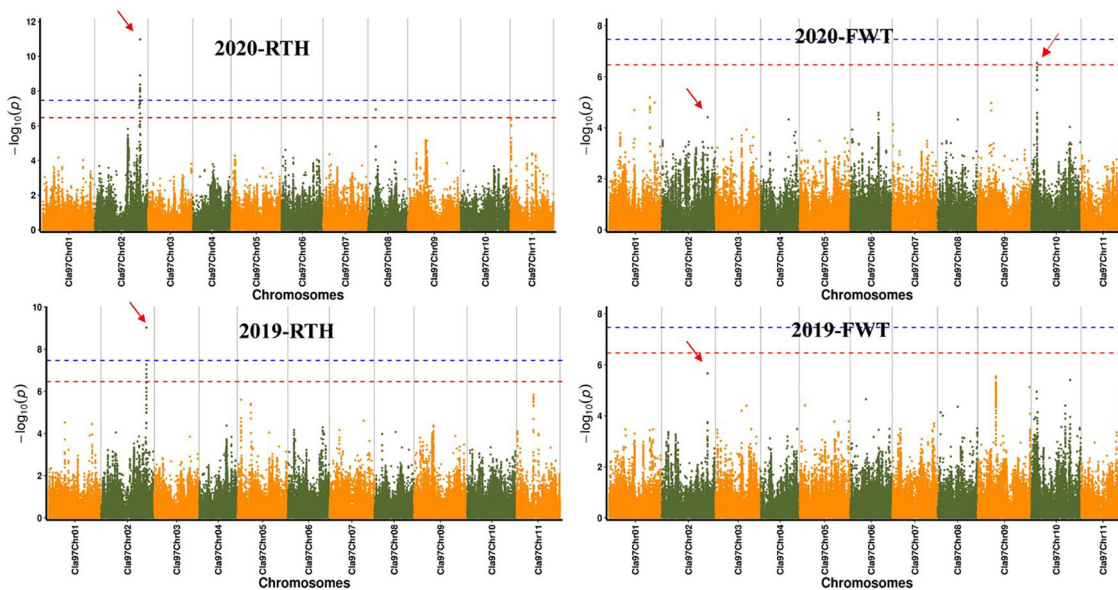
TABLE 1 Statistical analysis of rind thickness and single fruit weight of 151 watermelon accessions in 2019 and 2020.

Trait	Mean	SD	CV/%	Kurtosis	Skewness
2020 RTH	1.07	0.29	0.27	0.06	-0.54
2019 RTH	1.01	0.26	0.26	0.98	0.08
2020 FWT	3.21	0.74	0.23	1.15	0.13
2019 FWT	3.06	0.86	0.28	0.05	0.34





**FIGURE 1** Distribution histograms and correlation graphs of rind thickness and single fruit weight of 151 watermelon accessions in 2019 and 2020. Statistical histogram of the rind thickness in 2020 (A), the rind thickness in 2019 (B), the single fruit weight in 2020 (D) and the single fruit weight in 2019 (E), respectively. Correlation diagram of two-year data of rind thickness (C) and single fruit weight (F), respectively.



**FIGURE 2** Manhattan plot of watermelon rind thickness and single fruit weight by GWAS in 2019 and 2020. The abscissa represents the chromosome, the ordinate represents  $-\log_{10}(p)$ -value, and the red and blue line segments parallel to the coordinate axis represent the two threshold lines, SNPs over red represent highly significant positive associations with traits, SNPs over blue are significantly associated with traits, respectively.

thickness found that there were 5 and 16 SNPs that exceeded the threshold line in two years (Table S2), and five same significant SNPs (S2: 32344170, S2: 32321308, S2: 32304738, S2: 32328501, S2: 32311192) on chromosome 2 were located on chromosome 2 in 2019 and 2020, and these loci were within the range of 50 kb. Besides, S8:5576295 was a significant SNP of rind thickness on chromosome 8 in 2020, and S10:4123404 was a significant SNP of single fruit weight on chromosome 10. Interestingly, S2: 32344170, which was significantly associated with rind thickness, was also the most significant SNP on chromosome 2 for single fruit weight. It is worth noting that GWAS analysis by EMMAX and LM has relatively similar positioning results with FAST-LMM, which further explains the reliability of our positioning results (Figure S3).

## The acquisition of candidate interval and candidate gene

Notably, through the above analysis, we found that significant SNP loci were within a candidate interval of about 50 kb. Interestingly, S2: 32304738 and S2: 32311192 were located on the promoter and intron of *ClA97C02G044120*, respectively, and S2: 32321308 and S2: 32328501 were located on the intron and exon of *ClA97C02G044130*, respectively, and S2: 32344170 was the most significant SNP. To obtain allele information of significant SNP based on resequencing data, it is worth noting that when these alleles were in mutant alleles, they have thinner rind thickness and smaller fruit, and this difference was more obvious at S2: 32344170 (Figure 3). Moreover, it was found that when the rind thickness was less than 0.6 cm, all the alleles correspond to the mutant alleles (Table S1). These results further

illustrate the close correlation between significant SNP with target traits on chromosome 2, and S2: 32344170 could be used as potential molecular markers for marker-assisted breeding of target traits. Furthermore, the correlation analysis of regions and genes near the most significant locus S2: 32344170 was analyzed, and the 100 kb upstream and downstream intervals of this site were defined as candidate intervals for further analysis. LD block analysis found that SNP in the candidate interval had a relatively good correlation coefficient (Figure S4), indicating that there may be a linkage imbalance in the candidate interval. Interestingly, marker NW0249226 in the previous QTL *Qrth2-1* for rind thickness was only about 50 kb away from S2: 32344170.

A total of 21 genes were obtained in the candidate interval (Table 2). GO annotation analysis found that in the molecular\_function term, more genes were enriched in GTP binding, molecular\_function, protein binding, ATP binding, protein kinase activity, and transferase activity; in the cellular\_component term, an integral component of membrane, membrane, nucleus, and plasma membrane had more genes were annotated; in biological\_process term, more genes were enriched in protein phosphorylation, MAPK cascade, regulation of transcription, DNA-templated phosphorylation had annotated more genes. It is worth noting that protein phosphorylation, DNA-templated phosphorylation, etc. were often considered to be involved in auxin biosynthesis.

## Gene expression analysis of *ClA97C02G044160*

Combined with transcriptome data in '97103' and '203Z', the genes in the candidate interval were analyzed. It was found that

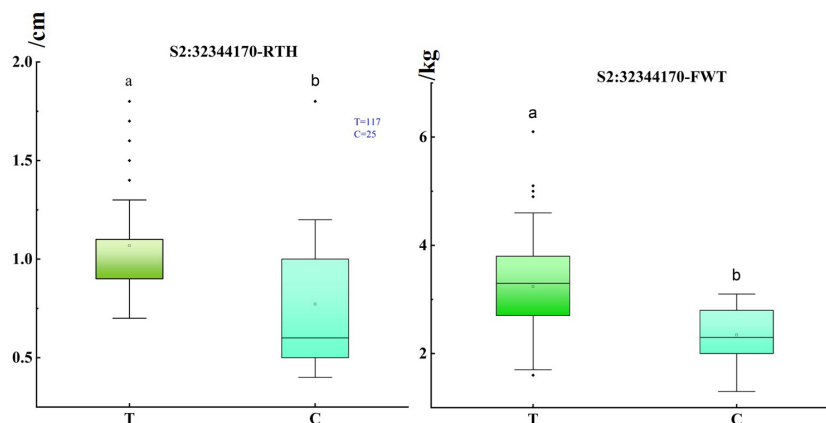


FIGURE 3

Box plot of single fruit weight and rind thickness under different bases of S2: 32344170. The abscissa represents the alleles with a mutation at the locus, the reference base near the origin, the other one was the key mutation locus, and the ordinate represents the rind thickness and fruit weight, respectively. Significance level: 0.05, a, and b represented significant differences in rind thickness under different alleles. The hollow origin in the boxplot represents the mean value.

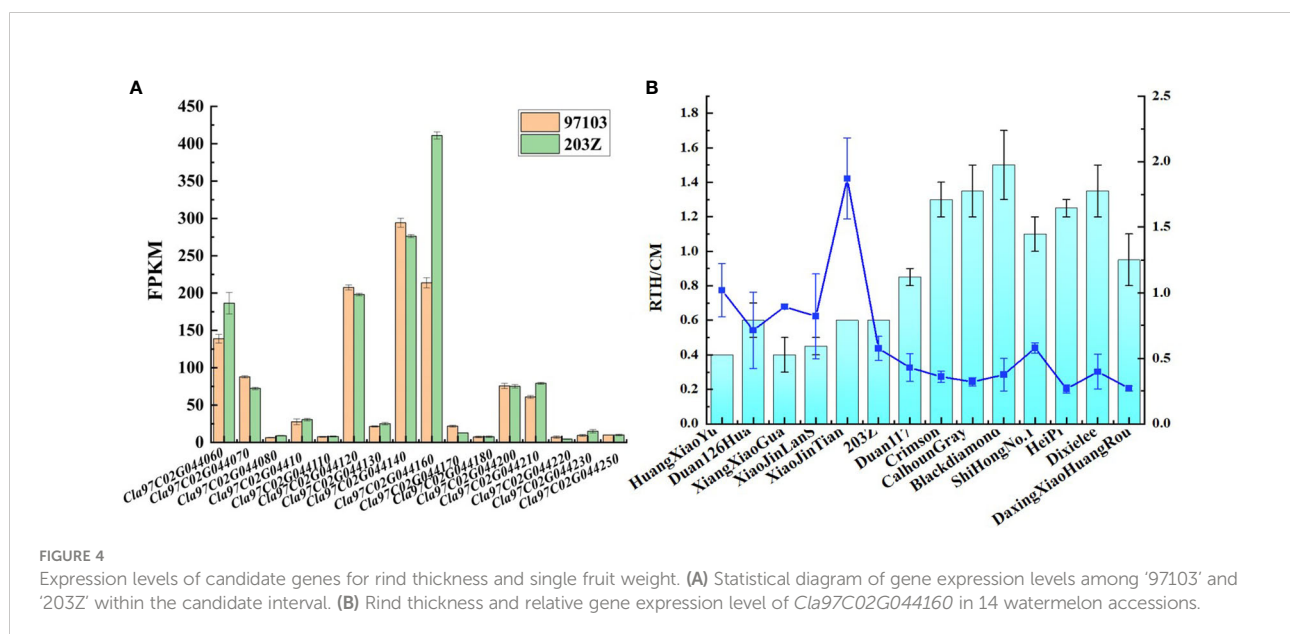
TABLE 2 Candidate gene information of rind thickness and single fruit weight on chromosome 2.

Gene	Location	Gene annotation
<i>Cla97C02G044060</i>	Cla97Chr02:32246890.32248741 (+)	transcription elongation factor B polypeptide 3-like
<i>Cla97C02G044070</i>	Cla97Chr02:32252262.32256301 (-)	Citrate synthase
<i>Cla97C02G044080</i>	Cla97Chr02:32264788.32269245 (-)	Ubiquinone biosynthesis O-methyltransferase, mitochondrial
<i>Cla97C02G044090</i>	Cla97Chr02:32278703.32280043 (-)	Leucine-rich repeat receptor-like protein kinase family protein
<i>Cla97C02G044100</i>	Cla97Chr02:32283295.32286529 (-)	Serine/threonine protein phosphatase 7 long form
<i>Cla97C02G044110</i>	Cla97Chr02:32290370.32300164 (-)	Serine/threonine-protein phosphatase
<i>Cla97C02G044120</i>	Cla97Chr02:32305922.32314246 (+)	elongation factor 2
<i>Cla97C02G044130</i>	Cla97Chr02:32320831.32328837 (+)	Protein kinase, putative
<i>Cla97C02G044140</i>	Cla97Chr02:32332095.32335291 (+)	elongation factor 2
<i>Cla97C02G044150</i>	Cla97Chr02:32337553.32338449 (-)	HVA22-like protein
<i>Cla97C02G044160</i>	Cla97Chr02:32349097.32356436 (+)	MADS box transcription factor AGAMOUS
<i>Cla97C02G044170</i>	Cla97Chr02:32360188.32365056 (-)	Hexosyltransferase
<i>Cla97C02G044180</i>	Cla97Chr02:32382369.32383409 (+)	Secretion-regulating guanine nucleotide exchange factor
<i>Cla97C02G044190</i>	Cla97Chr02:32384480.32385393 (-)	protodermal factor 1-like
<i>Cla97C02G044200</i>	Cla97Chr02:32387278.32388270 (-)	peroxisomal and mitochondrial division factor 2
<i>Cla97C02G044210</i>	Cla97Chr02:32390210.32394734 (-)	Mitogen-activated protein kinase
<i>Cla97C02G044220</i>	Cla97Chr02:32411247.32412302 (-)	glucan endo-1,3-beta-glucosidase 12-like
<i>Cla97C02G044230</i>	Cla97Chr02:32427610.32430022 (+)	WRKY transcription factor 100
<i>Cla97C02G044240</i>	Cla97Chr02:32433191.32433619 (-)	Unknown protein
<i>Cla97C02G044250</i>	Cla97Chr02:32434065.32437213 (-)	Receptor-like protein kinase
<i>Cla97C02G044260</i>	Cla97Chr02:32437502.32438686 (-)	Unknown protein

among the detected genes, only the gene expression level of *Cla97C02G044160* was more than 1.8 times between the two cultivars (Figure 4A). The analysis of this gene showed that *Cla97C02G044160* belongs to MADS family, and the typical conserved domains of MADS\_MEF2 and K-Box, and transmembrane domain analysis found that K-Box region has a transmembrane region (Figure S5). Further combined with

phylogenetic tree analysis, it was found that the gene had high homology with the previously reported genes *Osmads34*, *Osmads3*, and *Osmads13* regulating rice organ size (Figure S6) (Li et al., 2011; Zhang et al., 2016).

To fully understand the expression of candidate gene *Cla97C02G044160* in the natural population, RNA was extracted from 14 germplasm with different watermelon rind



thickness and fruit weight and further used for the qRT-PCR assay. The results showed that the gene expression level was significantly higher when the rind thickness was less than 0.6 cm than when the rind thickness was more than 0.6 cm. Specifically, The gene expression level of *Cl97C02G044160* in ‘HuangXiaoYu’, ‘Duan126Hua’, ‘XiangXiaoGua’, ‘XiaoJinLanS’, ‘XiaoJinTian’, ‘203Z’ was significantly higher than that of ‘Duan117’, ‘Crimson’, ‘CalhounGray’, ‘Blackdiamond’, ‘ShiHongNo.1’, ‘HeiPi’, ‘Dixielee’ and ‘DaxingXiaoHuangRou’ (Figure 4B). It should be noted that this trend of gene expression was significantly opposite to that of rind thickness, that is, the gene expression level of *Cl97C02G044160* in rind was higher in watermelon with thinner rind. These results suggest that *Cl97C02G044160* maybe was a key candidate gene for regulating watermelon rind thickness and single fruit weight.

## Transcriptome analysis of cultivars with different phenotypes

The genes whose expression was more than 1.8 times different between ‘97103’ and ‘203Z’ were counted. A total of 12 918 genes were screened for genes with an average content greater than 5. Furthermore, a total of 2117 genes were obtained by screening differentially expressed genes between the two cultivars. These 2117 genes were mainly divided into two categories, one has a higher gene expression in ‘97103’, there are 1241 genes, and the second group had higher gene expression in ‘203Z’, with 876 genes.

Among these genes, the statistics of genes related to organ expansion revealed that the expression levels of xyloglucan endotransglucosylase/hydrolase, cellulose synthase, pectinesterase, and pectinesterase inhibitor were higher in the watermelon ‘97103’ with thicker rind thickness and heavier fruit; Most of the key genes auxin efflux carrier component, auxin transporter-like protein, ABC transporter B family member related to auxin synthesis had relatively higher gene expression levels in ‘97103’; as well as ethylene synthesis-related genes S-adenosylmethionine, 1-aminocyclopropane-1-carboxylate synthase family protein, 1-aminocyclopropane-1-carboxylate oxidase 1, most of them have higher gene expression levels in ‘97103’; while the cytokinin-related genes cytokinin riboside 5’-monophosphate phosphoribohydrolase, adenylate isopentenyltransferase have higher content in ‘203Z’ (Figure 5). Combined with the GO annotation information in the candidate interval based on GWAS, it is speculated that the genes related to the regulation of auxin synthesis were likely to regulate target trait. Combined with the analysis of the expression level of candidate gene *Cl97C02G044160*, these genes with the same or opposite gene expression levels may have a potential regulatory relationship with

*Cl97C02G044160*, or may have synergistic effects to make watermelon show different single fruit weight or rind thickness.

## Discussion

Single fruit weight and rind thickness of watermelons are important agronomic traits closely related to domestication and yield. Bahari et al. (2012) screened watermelon cross combinations suitable for commercial cultivation by calculating the general combining ability (GCA) and special combining ability (SCA) of agronomic traits such as rind thickness and fruit weight, to screen out the hybrid combination more in line with the market demand. In the process of domestication, organ expansion is one of the important characteristics, Guo et al. (2019) found the average fruit weight of *C. Colocynthis*, *C. amarus*, *C. mucosospermus*, and *C. lanatus landraces* and cultivars was 0.2, 1.7, 3.7 and 3.4 kg, respectively. However, the artificial selection makes modern cultivated cultivars have a higher degree of domestication, correlation analysis of 151 modern cultivated watermelons showed that the watermelons with thinner rinds tended to have smaller fruits, thus establishing a link between phenotypes. In addition, the abundant variety of watermelon natural populations provides more potential parents for future breeding work.

For gene mapping of quantitative traits, the genetic map is one of the methods to obtain major QTLs for target traits. For instance, Ren et al. (2014) integrated four previously reported genetic maps into a new map of watermelon, and 58 QTLs for 12 traits were mapped into the new map. And interestingly, co-localization of rind thickness (*RTH2-1*) and fruit weight (*FWT2-2*) was found on chromosome 2 and chromosome 5. Moreover, *FWT2-2*, which controls the weight of watermelon fruit, was verified by selective clearance analysis of natural population accessions (Guo et al., 2019). GWAS can quickly obtain molecular markers and candidate genes that are significantly related to traits by detecting genetic variation polymorphisms of multiple individuals in the whole genome and analyzing their association with traits and made important progress in the study of important agronomic traits of watermelon (Dou et al., 2018; Ren et al., 2021). In the present study, 5 significant SNP loci on chromosome 2 were obtained in 2019 and 2020. Interestingly, the highest SNP locus S2: 32344170 was only about 50 kb away from QTL *FWT2-2* and *RTH2-1* (*RTH2-1* was included in QTL *FWT2-2*) (Ren et al., 2014), which regulated watermelon single fruit weight and rind thickness, indicating the reliability of the significant SNPs we obtained. The discovery of molecular markers can be applied to the molecular marker-assisted breeding of crops (Fang et al., 2021; Zhuo et al., 2021), in this study, we found that significant SNPs had significant differences in rind thickness under different bases. For example, under the



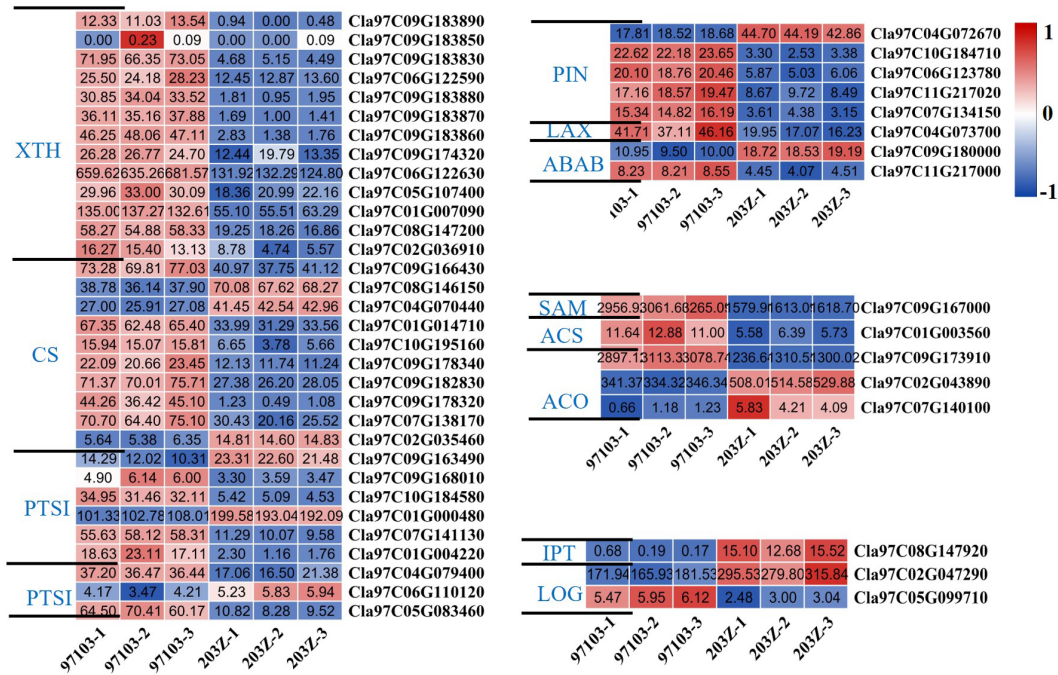


FIGURE 5

Heat map of differentially expressed genes related to seed expansion or hormone synthesis (auxin, ethylene, cytokinin). The darker the red color in the color scale, the higher the relative expression of the gene in the sample; the darker the blue color, the lower the expression of the gene in the sample. The value in the box represents the FPKM value in the sample. XTH, Xyloglucan endotransglucosylase/hydrolase; CS, Cellulose synthase; PTSI, Pectinesterase; PTSI, Pectinesterase inhibitor; PIN, Auxin efflux carrier component; LAX, Auxin transporter-like protein; ABAB, ABC transporter B family member; SAM, S-adenosylmethionine; ACS, 1-aminocyclopropane-1-carboxylate synthase family protein; ACO, 1-aminocyclopropane-1-carboxylate oxidase 1; LOG, Cytokinin riboside 5'-monophosphate phosphoribohydrolase; IPT, Adenylate isopentenyltransferase.

highest significant SNP, when the rind thickness of cultivated cultivars was less than 0.6 cm, the corresponding base was all C. In summary, these significant SNPs can be used as potential molecular markers in watermelon breeding.

The differences in phenotypic traits are often regulated by genes. Combined with transcriptome data, we obtained the possible candidate gene *Cla97C02G044160* (a MADS family gene) in the candidate interval. In previous reports, genes in the MADS family can negatively regulate the size of tissues and organs by regulating hormones, etc. For instance, *SEEDSTICK* regulates cytokinin levels and other genes to control the fruit size of *Arabidopsis* (Di Marzo et al., 2020); RNA interference of *SLMBP21* lead to longer sepals and improved fruit set in tomato, and maybe a significant correlation with auxin and ethylene (Li et al., 2017). In cucumber fruit, *CsFULIA* inhibits the expression of auxin transporters *PIN-FORMED1* (*PIN1*) and *PIN7* and then resulting in decreases in auxin (Zhao et al., 2019). Zheng et al. (2020) found that two MADS-box transcription factors, *VCM1* and *VCM2*, play an important role in tree thickening. The discovery that the MADS family gene *Cla97C02G044160* could negatively regulate the size of tissues and organs was also

consistent with this study, that is, the expression level of this gene in watermelon '203Z' with lighter fruit weight and the thinner rind was significantly higher than '97103'. In addition, combined with transcriptome data, it was further found that key genes related to organ expansion and auxin and ethylene had a higher accumulation amount in '97103', and our study further improved the transcriptional regulation basis of watermelon rind thickness and single fruit weight.

## Conclusions

In the current study, we identified natural variations and correlations between rind thickness and single fruit weight in 151 watermelon accessions, and the significant SNP loci obtained by GWAS may serve as potential molecular markers to participate in breeding. Further combined with transcriptome data and bioinformatics analysis, it was found that *Cla97C02G044160* may be involved in the phenotypic changes of the target trait. Our research provides rich data resources for the selection of new cultivars and molecular marker-assisted breeding.

## Data availability statement

The data presented in the study are deposited in the NCBI repository, accession numbers PRJNA641178 and PRJNA905076, all data is publicly available and can be used with the permission of the authors.

## Author contributions

WL, NH, and CG designed this experiment. WL, XL, NH, and HZ were mainly responsible for the collection of watermelon accessions and plants management. WL and CG were mainly responsible for the statistics of phenotypic data. CG was mainly responsible for the data collation and analysis, and completed the writing of the original manuscript. WL, NH, and MA were mainly responsible for the manuscript verification. All the authors agreed to the submission of the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was financially supported by the Agricultural Science and Technology Innovation Program (CAAS-ASTIP-2021-ZFRI), the China Agriculture Research System of MOF and MARA (CARS-25-03), National Key R&D Program of China (2018YFD0100704), the National Nature Science Foundation of China (31672178 and 31471893) and Natural Science Foundation of Henan Province (212300410312), and the Funding of Joint Research on Agricultural Variety Improvement of Henan Province (Grant No. 2022010503).

## References

- Bahari, M., Rafii, M. Y., Saleh, G. B., and Latif, M. A. (2012). Combining ability analysis in complete diallel cross of watermelon (*Citrullus lanatus* (Thunb.) matsum. Nakai) *ScientificWorldJournal* 2012, 543158–543158. doi: 10.1100/2012/543158
- De Jong, M., Wolters-arts, M., Feron, R., Martiani, C., and Vriezen, W. H. (2009). The solanum lycopersicum auxin response factor 7 (SlARF7) regulates auxin signaling during tomato fruit set and development. *Plant J.* 57 (1), 160–170. doi: 10.1111/j.1365-3113X.2008.03671.x
- Di Marzo, M., Herrera-Ubaldo, H., Caporali, E., Novák, O., Strnad, M., Balanzà, V., et al. (2020). *SEEDSTICK* controls arabidopsis fruit size by regulating cytokinin levels and FRUITFULL. *Cell Rep.* 30, 2846–2857.e2843. doi: 10.1016/j.celrep.2020.01.101
- Dou, J., Zhao, S., Lu, X., He, N., Zhang, L., Ali, A., et al. (2018). Genetic mapping reveals a candidate gene (*CIFS1*) for fruit shape in watermelon (*Citrullus lanatus* L.). *Theor. Appl. Genet.* 131, 947–958. doi: 10.1007/s00122-018-3050-5
- Fang, K., Xia, Z., Li, H., Jiang, X., Qin, D., Wang, Q., et al. (2021). Genome-wide association analysis identified molecular markers associated with important tea flavor-related metabolites. *Hortic. Res.* 8, 42. doi: 10.1038/s41438-021-00477-3
- Florea, L., Song, L., and Salzberg, S. L. (2013). Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues. *F1000Res* 2, 188–188. doi: 10.12688/f1000research
- Gong, C., Diao, W., Zhu, H., Umer, M. J., Zhao, S., He, N., et al. (2021a). Metabolome and transcriptome integration reveals insights into flavor formation of 'Crimson' watermelon flesh during fruit development. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.629361
- Gong, C., Zhao, S., Yang, D., Lu, X., Anees, M., He, N., et al. (2022). Genome-wide association analysis provides molecular insights into natural variation in watermelon seed size. *Hortic. Res.* 9, uhab074. doi: 10.1093/hr/uhab074
- Gong, C., Zhu, H., Lu, X., Yang, D., Zhao, S., Umer, M. J., et al. (2021b). An integrated transcriptome and metabolome approach reveals the accumulation of taste-related metabolites and gene regulatory networks during watermelon fruit development. *Planta* 254, 35. doi: 10.1007/s00425-021-03680-7
- Guo, S., Zhang, J., Sun, H., Salse, J., Lucas, W. J., Zhang, H., et al. (2013). The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat. Genet.* 45, 51–58. doi: 10.1038/ng.2470
- Guo, S., Zhao, S., Sun, H., Wang, X., Wu, S., Lin, T., et al. (2019). Resequencing of 414 cultivated and wild watermelon accessions identifies selection for fruit quality traits. *Nat. Genet.* 51, 1616–1623. doi: 10.1038/s41588-019-0518-4
- Hamblin, M. T., Buckler, E. S., and Jannink, J.-L. (2011). Population genetics of genomics-based crop improvement methods. *Trends Genet.* 27, 98–106. doi: 10.1016/j.tig.2010.12.003

## Acknowledgments

We sincerely thank the editors and reviewers for their contributions to the manuscript review. We also thank Henan Key Laboratory of Fruit and Cucurbit Biology for providing transcriptome detection and analysis in our experiment.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.1074145/full#supplementary-material>

- Horiguchi, G., Ferjani, A., Fujikura, U., and Tsukaya, H. (2006). Coordination of cell proliferation and cell expansion in the control of leaf size in *Arabidopsis thaliana*. *J. Plant Res.* 119, 37–42. doi: 10.1007/s10265-005-0232-4
- Le, J., Vandenbussche, F., van der Straeten, D., and Verbelen, J. P. (2001). In the early response of *Arabidopsis* roots to ethylene, cell elongation is up- and down-regulated and uncoupled from differentiation. *Plant Physiol.* 125, 519–522. doi: 10.1104/pp.125.2.519
- Li, N., Huang, B., Tang, N., Jian, W., Zou, J., Chen, J., et al. (2017). The MADS-box gene *SLMBP21* regulates sepal size mediated by ethylene and auxin in tomato. *Plant Cell Physiol.* 58, 2241–2256. doi: 10.1093/pcp/pcx158
- Li, H., Liang, W., Yin, C., Zhu, L., and Zhang, D. (2011). Genetic interaction of *OsMADS3*, *DROOPING LEAF*, and *OsMADS13* in specifying rice floral organ identities and meristem determinacy. *Plant Physiol.* 56 (1), 263–274. doi: 10.1104/pp.111.172080
- Liu, W., Jiang, B., Peng, Q., He, X., Lin, Y., Wang, M., et al. (2018). Genetic analysis and QTL mapping of fruit-related traits in wax gourd (*Benincasa hispida*). *Euphytica* 214 (8), 1–8. doi: 10.1007/s10681-018-2166-7
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-delta delta C(T)</sup> method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Pan, Y., Wang, Y., McGregor, C., Liu, S., Luan, F., Gao, M., et al. (2020). Genetic architecture of fruit size and shape variation in cucurbits: a comparative perspective. *Theor. Appl. Genet.* 133, 1–21. doi: 10.1007/s00122-019-03481-3
- Prasad, K., Zhang, X., Tobón, E., and Ambrose, B. A. (2010). The *Arabidopsis* b-sister MADS-box protein, *GORDITA*, represses fruit growth and contributes to integument development. *Plant J.* 62, 203–214. doi: 10.1111/j.1365-313X.2010.04139.x
- Ren, Y., Guo, S., Zhang, J., He, H., Sun, H., Tian, S., et al. (2017). A tonoplast sugar transporter underlies a sugar accumulation QTL in watermelon. *Plant Physiol.* 176, 836–850. doi: 10.1104/pp.17.01290
- Ren, Y., Li, M., Guo, S., Sun, H., Zhao, J., Zhang, J., et al. (2021). Evolutionary gain of oligosaccharide hydrolysis and sugar transport enhanced carbohydrate partitioning in sweet watermelon fruits. *Plant Cell.* 33, 1554–1573. doi: 10.1093/plcell/koab055
- Ren, Y., McGregor, C., Zhang, Y., Gong, G., Zhang, H., Guo, S., et al. (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. doi: 10.1186/1471-2229-14-33
- Umer, M. J., Bin Safdar, L., Gebremeskel, H., Zhao, S., Yuan, P., Zhu, H., et al. (2020). Identification of key gene networks controlling organic acid and sugar metabolism during watermelon fruit development by integrating metabolic phenotypes and gene expression profiles. *Hortic. Res.* 7, 193–193. doi: 10.1038/s41438-020-00416-8
- Wilmoth, J. C., Wang, S., Tiwari, S. B., Joshi, A. D., Hagen, G., Guilfoyle, T. J., et al. (2005). *NPH4/ARF7* and *ARF19* promote leaf expansion and auxin-induced lateral root formation. *Plant J.* 43, 118–130. doi: 10.1111/j.1365-313X.2005.02432.x
- Wi Sandlin, K., Prothro, J., Heesacker, A., Khalilian, N., Okashah, R., Xiang, W., et al. (2012). Comparative mapping in watermelon [*Citrullus lanatus* (Thunb.) matsum. *Nakai*] *Theor. Appl. Genet.* 125, 1603–1618. doi: 10.1007/s00122-012-1938-z
- Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P.-C., Hu, L., et al. (2016). Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat. Genet.* 48, 927–934. doi: 10.1038/ng.3596
- Zhang, J., Sun, H., Guo, S., Ren, Y., and Xu, Y. (2020). Decreased protein abundance of lycopene  $\beta$ -cyclase contributes to red flesh in domesticated watermelon. *Plant Physiol.* 183, 1171–1183. doi: 10.1104/pp.19.01409
- Zhang, Y., Yu, H., Liu, J., Wang, W., Sun, J., Gao, Q., et al. (2016). Loss of function of *OsMADS34* leads to large sterile lemma and low grain yield in rice (*Oryza sativa* L.). *Mol. Breed.* 36 (11), 147. doi: 10.1007/s11032-016-0578-4
- Zhao, J., Jiang, L., Che, G., Pan, Y., Li, Y., Hou, Y., et al. (2019). A functional allele of *CsFUL1* regulates fruit length through repressing *CsSUP* and inhibiting auxin transport in cucumber. *Plant Cell.* 31, 1289–1307. doi: 10.1105/tpc.18.00905
- Zheng, S., He, J., Lin, Z., Zhu, Y., Sun, J., and Li, L. (2020). Two MADS-box genes regulate vascular cambium activity and secondary growth by modulating auxin homeostasis in populus. *Plant Commun.* 23, 100134. doi: 10.1016/j.xplc.2020.100134
- Zhong, Y. J., Zhou, Y. Y., Li, J. X., Yu, T., Wu, T. Q., Luo, J. N., et al. (2017). A high-density linkage map and QTL mapping of fruit-related traits in pumpkin (*Cucurbita moschata* Duch.). *Sci. Rep.* 7 (1), 12785. doi: 10.1038/s41598-017-13216-3
- Zhuo, X., Zheng, T., Li, S., Zhang, Z., Zhang, M., Zhang, Y., et al. (2021). Identification of the *PmWEEP* locus controlling weeping traits in prunus mume through an integrated genome-wide association study and quantitative trait locus mapping. *Hortic. Res.* 8, 131–131. doi: 10.1038/s41438-021-00573-4