Genetic mapping reveals a candidate gene *CmoFL1* controlling fruit length in pumpkin

Yimei Zhou[1†], Meng Zhao[1†], Qinghui Shen[1], Mengyi Zhang[1,2], Chenhao Wang[1], Yutong Zhang[1], Qinrong Yang[1], Yongming Bo[3], Zhongyuan Hu[1,2,4], Jinghua Yang[1,2,4], Mingfang Zhang[1,2,4] and Xiaolong Lyu[1*]

[1]Laboratory of Germplasm Innovation and Molecular Breeding, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, China, [2]Hainan Institute of Zhejiang University, Sanya, China, [3]Ningbo Weimeng Seed Company, Ningbo, China, [4]Key laboratory of Horticultural Plant growth, Development and Quality Improvement, Ministry of Agriculture, Hangzhou, China

Fruit length (FL) is an important economical trait that affects fruit yield and appearance. Pumpkin (*Cucurbita moschata* Duch) contains a wealth genetic variation in fruit length. However, the natural variation underlying differences in pumpkin fruit length remains unclear. In this study, we constructed a F_2 segregate population using KG1 producing long fruit and MBF producing short fruit as parents to identify the candidate gene for fruit length. By bulked segregant analysis (BSA-seq) and Kompetitive Allele-Specific PCR (KASP) approach of fine mapping, we obtained a 50.77 kb candidate region on chromosome 14 associated with the fruit length. Then, based on sequence variation, gene expression and promoter activity analyses, we identified a candidate gene (*CmoFL1*) encoding E3 ubiquitin ligase in this region may account for the variation of fruit length. One SNP variation in promoter of *CmoFL1* changed the GT1CONSENSUS, and DUAL-LUC assay revealed that this variation significantly affected the promoter activity of *CmoFL1*. RNA-seq analysis indicated that *CmoFL1* might associated with the cell division process and negatively regulate fruit length. Collectively, our work identifies an important allelic affecting fruit length, and provides a target gene manipulating fruit length in future pumpkin breeding.

Introduction

The fruit is an important organ in horticultural crops, which has a rich variety of shapes and sizes. Fruit length (FL) is a key indicator of fruit yield and appearance quality, directly affects fruit shape and size, and also exhibits a broad range of variation, thus emphasizing its importance in breeding program. Previous studies have showed that the fruit length and fruit shape were controlled by a single gene in some populations (Dou et al., 2018; Cheng et al., 2021; Ma et al., 2022), while extensive studies have confirmed that the fruit length and shape were commonly modulated by both quantitative trait loci (QTLs) and some environmental factors (Wang et al., 2020b, c). The genetic analysis of quantitative traits using mixed inheritance model with a major gene plus a polygene have elucidated the genetic rules regulating fruit length and fruit size in some crops (Ren et al., 2022; Yan et al., 2023). The genetic models that control fruit length were different in different crops.

Many important genes and quantitative trait loci (QTLs) regulating fruit length have been reported in different horticultural crops. In tomatoes, *OVATE* and *SUN* control fruit elongation. *OVATE* encodes a family of OVATE proteins (OFPs), which are negative growth regulators that control fruit length by blocking transcription, leading to the appearance of pear-shaped fruit in tomato (Liu et al., 2002; van der Knaap et al., 2014). *SUN* encodes a calmodulin-binding protein in the IQ-Domain (IQD) family that regulates plant growth and fruit shape by altering the way cells divide (Wu et al., 2011). In pepper, fruit length was significantly increased by down-regulation the expression of *CaOvate* through VIGS (Tsaballa et al., 2011). The gene *Smechr0301963*, which belongs to the *SUN* gene family, was predicted to be a key candidate gene for eggplant fruit length regulation (Wei et al., 2020). *CmCLV3*, which determines carpels number, is also thought to influence fruit shape in the andromonoecious genetic context in melon (Liu et al., 2020). In addition, overexpress the CRABS CLAW family gene *CmCRC* in melon can affect fruit elongation (Switzenberg et al., 2015). Depending on different genetic backgrounds, cucumber fruit length can be controlled by different single inherited sites, such as *CsSUN*, *CsTRM5*, *CsSF1*, *CsFUL1*, *CsSF2* genes (Pan et al., 2017; Wu et al., 2018; Xin et al., 2019; Zhao et al., 2019; Zhang et al., 2020). In watermelon, using materials with different genetic backgrounds, a major QTL controlling fruit shape was detected on chromosome 3, and the homologous gene *Cla011257* of *SUN* was found in this region (Sandlin et al., 2012; Dou et al., 2018; Legendre et al., 2020).

Pumpkin (*Cucurbita moschata* Duch) is one of the most important horticultural crops of cucurbits, which also serve as mainstays of diet to provide significant sources of nutrition. Pumpkin contains a wealth genetic variation in fruit shape and size, and fruit length exhibits great difference, varying from 5 cm to more than 100 cm (Oliveira et al., 2016; Paris, 2018; Ezin et al., 2022). With the development of pumpkin genomics (Montero-Pau et al., 2017; Sun et al., 2017), the mapping of fruit-related genes has been greatly promoted, and many QTLs that control fruit length and shape have been identified. In *Cucurbita pepo*, QTL mapping analysis was conducted on the size and shape of immature and mature fruits, and major QTLs (*IFSh_3, IFLe_3, IFWi_3, MFSh_3, MFLe_3*, and *MFWi_3*) were detected related to fruit shape (Esteras et al., 2012; Montero-Pau et al., 2017). Three QTLs with the total PVE of 20.25% were detected for FSI in pumpkin, and the *fsi2.1* and *fsi6.1* were overlapped with *fl2.1* and *fl6.1*, respectively, suggesting that fruit length plays a major role in determining fruit shape (Han et al., 2022).

In this study, to anchor the candidate gene of pumpkin fruit length, we developed F_2 population derived from the KG1 (extra-long pumpkin) and MBF (short-fruit pumpkin) hybrid. The fruit length of KG1 is about 107.13 cm, while that of MBF is about 12.25 cm. Through BSA-seq analysis and KASP approaches, we obtained a 50.77 kb candidate region on chromosome 14 associated with the fruit length. By sequence variation, gene expression and promoter activity analyses, we finally revealed a candidate gene *CmoFL1* controlling fruit length, which encodes E3 ubiquitin ligase. Our finding will help to clarify the genetic rules of fruit length regulation and provide valuable genetic resources for the production and breeding of new varieties of pumpkin with excellent fruit type.

Materials and methods

Plant material, statistical and genetic analyses

Two lab-preserved pumpkin accessions (KG1 and MBF) were selected and used to construct the fourth-generation population for genetic analysis of fruit length. KG1 produces a long fruit, while MBF produces a short fruit. F_1 plants were obtained by crossing KG1 (P_1) as female parent and MBF (P_2) as male parent in 2019, and the F_2 population was obtained by self-crossing F_1 plants. Seedlings of F_2 were raised in a greenhouse in Hangzhou City, China, in the spring of 2020. According to pumpkin cultivation technology, all plants adopt unified water fertilizer and pest management methods to ensure the same growth and development environment of individual plants. The pumpkin fruit length (cm) was measured with a tape and recorded on the day of flowering and at fruit maturity.

SPSS 23 was used to statistically analyze the phenotypic data of four generations of pumpkin populations, including the average, standard deviation, skewness and kurtosis. The fruit length was expressed as "mean ± standard deviation (mean ± SD)", and the significance was analyzed by ANOVA. P < 0.05 indicated that there was a significant difference.

Genetic analysis was performed using R software package SEA v2.0.1(https://cran.r-project.org/web/packages/SEA/index.html). Akaike's information criterion (AIC) value was calculated in each model to identify the existence of major genes affecting quantitative traits.

Cytological analysis

To compare the cytological characteristics of fruits from two parents, the ovaries of two parents at different stages before flowering and the fruits 35 days after pollination were picked respectively. Pulp with a size of about 1cm³ was cut from the part between pumpkin peel and heart cavity and fixed in FAA fixative solution. The fruit was embedded in paraffin (Surgipath), and then used to generate 8-μm–

thick sections in longitudinal directions using a model no. RM2235 microtome (Leica), before staining with fast green stain and imaging, the cell characteristics of ovary and fruit of the two parents at different stages were observed under microscope.

DNA extraction and library construction

The genomic DNA of the parents, and of the F_1 and F_2 populations, were extracted from fresh leaves of young seedlings using Cetyltrimethylammonium Bromide (CTAB) method (Porebski et al., 1997). The concentration and purity of the extracted DNA were determined by agarose gel electrophoresis and Nanodrop 2000 (Thermo Fisher Scientific, USA). 20 long-fruit F_2 individuals (FL > 45 cm) and 20 short-fruit F_2 individuals (FL < 20 cm) were mixed equally to construct L-pool and S-pool, respectively. Truseq Nano DNA HT Sample preparation Kit (Illumina) was used to generate sequencing libraries, and the Illumina HiSeq PE150 platform was used to sequence these libraries. The quality of the sequencing data was determined using FASTQC (Brown et al., 2017). The QC standard pipelines were employed as described in the literature (Lyu et al., 2022).

BSA-seq analysis

The *Cucurbita moschata* genome v1 (http://cucurbitgenomics.org/ftp/genome/Cucurbita_moschata/v1/) was used as a reference genome. The sequencing data were compared with the reference genome to analyze the sequencing depth and coverage. SAMtools software was used to convert BAM files. Single nucleotide polymorphism (SNP) variant calling was conducted using the GATK software with the Unified Genotype function and Variant Filtration (McKenna et al., 2010). The annotations of SNPs were performed using ANNOVAR software (Wang et al., 2010) based on the GFF3 files of the reference genome. The homozygous SNPs/InDels between two parents (KG1 and MBF) were extracted from the VCF files. In order to analyze the frequency differences of SNPs/InDels in offspring, the SNP indexes were calculated based on the read depth information for homozygous SNPs (Takagi et al., 2013) in the two offspring pools (L-pool and S-pool). Then, we calculated the ratio of different reads in the total number, which represented the SNP/InDel indexes of the base sites. We used sliding window methods to calculated the SNP/indel indexes of the whole genome and selected 1Mb as window, 1Kb as step size. The difference in the SNP/indel indexes between two pools was calculated as the Δ (All- index), 95% and 99% confidence levels were selected as the threshold line. Moreover, the G prime method was performed to calculate the G' value by using the QTLseqr package in R (Mansfeld and Grumet, 2018).

Fine mapping through KASP markers

To narrow down the genetic interval for fine mapping and to verify the accuracy of BSA-seq, we employed 18 pair of KASP primer combinations (FAM, HEX, Common) as markers based on SNPs/InDels generated from BSA-seq (Supplementary Table S1). The KASP markers were used for genotyping verification in both parents and part of F_2 population, and the primers with obvious genotyping effect were used for single plant polymorphism screening in all F_2 population. The KASP protocol was utilized as follows: Stage 1: preread stage, 30°C for 1 min; Stage 2: hold stage 94°C for 15 min; Stage 3: PCR stage (touchdown), 94°C for 20 s, 61°C for 1 min (decrease pf -0.6°C), recycling for 9 times (a total of 10 cycles), achieving a final annealing temperature of 55°C; Stage 4: PCR stage, 94°C for 20 s, 61°C for 1 min, recycling for 25 times; and Stage 5: postread stage, 30°C for 1 min. After the amplification, an ABI PRISM 7900HT (Applied Biosystems) was used to detect the fluorescence signal and validate the classification.

qRT-PCR analysis for candidate gene

The length of the ovary of parents from the formation of the ovary to the day of flowering were measured and recorded, and 5 biological replications for each variety were observed. Based on this, judge which period the ovaries of different lengths belong to. We took the ovaries of KG1 and MBF -12, -8, -4 and 0 days after anthesis (DAA), and stored them at - 80°C for further analyses.

The amplification primers were designed using Primer Premier 6.0 program based on relevant gene sequences from *Cucurbita Moschata* genome database (http://cucurbitgenomics.org/organism/9). The Easy RNA Kit (Easy-Do, Hangzhou, China) was used to extract RNA from the fruits of both parents before flowering according to the manufacturer's instructions. The cDNA was synthesized from 1 ng of the total RNA using ReverTra Ace qPCR RT Master Mix (Toyobo, Japan). We used the TOROGreen qPCR Master Mix (Toyobo, Japan) for qRT-PCR reactions by Thermo Fisher ABI Stepone Plus real-time fluorescence quantitative PCR. The relative expression levels of candidate genes were determined by using the $2^{-\Delta\Delta Ct}$ comparative method (Livak and Schmittgen, 2001), with gene *ACTIN* as an internal control. All analyses were conducted with three biological and technical replications. The primer sequences used for the above studies are listed in Supplementary Table S2.

Promoter activity analysis

Promoter activity of *CmoFL1* was analyzed using tobacco transient expression assay. The promoter sequence of *CmoFL1* from MBF and KG1 were cloned into upstream of the luciferase (LUC) coding sequence in the pGreen II 0800-LUC vector. The Renilla luciferase (REN) gene, driven by 35S promoter, was used as an internal control. By using Agrobacterium-mediated transformation, the plasmids were transformed into the leaves of tabacco (*N. benthamiana*). Grown two days in greenhouse, the relative luciferase activity LUC/REN was measured using the Dual-Luciferase Reporter Gene Assay Kit (Yeasen, Shanghai, China).

RNA-seq analysis

Total RNA was extracted from long-fruit KG1 and short-fruit MBF at -12 and 0 DAA. RNA sequencing of 1 μg RNA was performed using Illumina Truseq™ platform. Each stage had three biological replications. After filtering raw data and checking the sequencing error rate and G/C ratio, we obtained clean reads. The clean reads were mapped onto the Cucurbita moschata reference genome (http://cucurbitgenomics.org/ftp/genome/Cucurbita_moschata/v1/) using HISAT2 software. FPKM (fragments per kilobase of transcript per million mapped reads) values were calculated to estimate gene expression levels by StringTie software (Pertea et al., 2015). Transcripts with a fold change (FC) ≥ 1 and Padj ≤ 0.05 were identified as differentially expressed genes (DEGs) via DESeq2 software (Love et al., 2014).

Results

Phenotypic differences of fruit length in two parental lines

To investigate the inheritance pattern of the fruit length trait in pumpkin, we developed F_1 and F_2 populations by crossing KG1 (long fruit) and MBF (short fruit). The average FL of KG1 at maturity is 107.13 ± 14.33 cm, which is significantly longer than that of MBF, with an average FL of 12.25 ± 0.65 cm (Figure 1A; Supplementary Table S3). There was no significant difference in fruit diameter (FD) between the two parents. The average FD of KG1 was 12.98 ± 0.83 cm and that of MBF was 12.85 ± 0.70 cm and the fruit shape index (FSI) of KG1 and MBF were 8.32 ± 1.52 and 0.96 ± 0.07, respectively (Supplementary Table S3). We also investigate the fruit length between the two parents during ovary development, and found that the fruit length of KG1 increased rapidly before anthesis and was significantly longer than that of MBF after -12 DAA (Figures 1B, C).

Previous studies have showed that the organ shape and size were directly related to cell number and size (Pan et al., 2020). In order to investigate the difference of cell structure between two parents, we analyzed the paraffin sections from fruits of parents at different developmental stages (-12 DAA, -8 DAA, -4 DAA, 0 DAA, 35 DAA) (Figure 1D). Observation on the slices showed that there was no significant difference in cell size and morphology between MBF and KG1 during ovary development (Figure 1E). The cell number of MBF was consistently less than that of KG1 after -12 DAA, and the gap of cell number became larger with the development of ovary. The cell number of two parents showed a highly significant difference at -8, -4, 0 and 35 DAA (Figure 1F). These results revealed that the length difference between KG1 and MBF might mainly came from the number of cells.

Inheritance pattern of fruit length in pumpkin

Since there was a significant difference in fruit length between two parents before pollination, we measured the FL and FSI of F_2

individuals at both 0 DAA and 35 DAA (at maturity) (Supplementary Table S3). The correlation analysis of FL and FSI showed highly significant correlations between FL and FSI at 0 DAA and 35 DAA (Supplementary Table S4), indicating that before flowering, the shape of pumpkin ovary has been formed and remained stable during fruit growth. Furthermore, FL and FSI showed a wide range of variation in F_2 population. At 0 DAA and 35 DAA, FL varied in the range of 2.70–22.40 cm and 13.00–69.50 cm, respectively, while FSI varied in the range of 0.17–8.61 and 0.56–5.30, respectively (Supplementary Table S3). The distribution of FL and FSI were consistent with a positively skewed distribution (Figures 1G-J), which indicated that the pumpkin fruit length was a quantitative trait controlled by multiple genes. In addition, the average FL and FSI of F_1 fruits at maturity were 31.88 ± 4.01 cm and 1.96 ± 0.24 cm, respectively (Supplementary Table S3).

Moreover, we also found a significant difference of carpopodium length between two parents. KG1 has a carpopodium of about 85.43 cm, accounting for 79.74% of the total fruit length, while MBF has no carpopodium (Figure 2A). The carpopodium length of mature fruits were also measured in F_2 population, which was consistent with the variation of fruit length, and also displayed a skewed distribution in F_2 population (Figure 2B). To confirm the relevance between fruit length and carpopodium length, the correlation analysis was performed, which showed a significant positive correlation between them (Figure 2C), suggesting that the carpopodium length might determine the fruit length in pumpkin.

Genetic models for fruit length and carpopodium length were calculated based on the phenotypic data of P_1, P_2, F_1, F_2 populations. According to the lowest or relatively lower Akaike's information criterion (AIC) value and no significant (p < 0.05) parameter in the good-of-fit test (Supplementary Tables S5, S6), the optional genetic model for fruit length and carpopodium length was 1MG-A, indicating that a single major gene with an additive effect might be responsible for the pumpkin fruit length and carpopodium length.

BSA-seq analysis anchors a candidate region on chromosome 14 for fruit length

To anchor the candidate QTL responsible for fruit length, we conducted BSA-seq based on genome resequencing data analysis. Among the 231 F_2 individuals, 20 long-fruit individuals (FL > 45 cm) and 20 short-fruit F_2 individuals (FL < 20 cm) were mixed equally to construct L-pool and S-pool, respectively. Then four pools (MBF, KG1, L-pool and S-pool) were sequenced on the Illumina HiSeq™PE150 platform. After filtering low-quality and short reads, a total of 3.53 Gb and 3.65 Gb clean reads was generated from KG1 and MBF, respectively, while 6.57 Gb and 6.80 Gb clean reads were generated from the L-pool and S-pool, respectively (Supplementary Table S7). The filtered reads of each sample were aligned to the reference genome of *Cucurbita moschata* genome v1 and the mapping rate was 97.54% ~ 98.31%, the average depth was 10.74× ~ 19.7× (Supplementary Table S7). In total, 1,898,813 SNPs
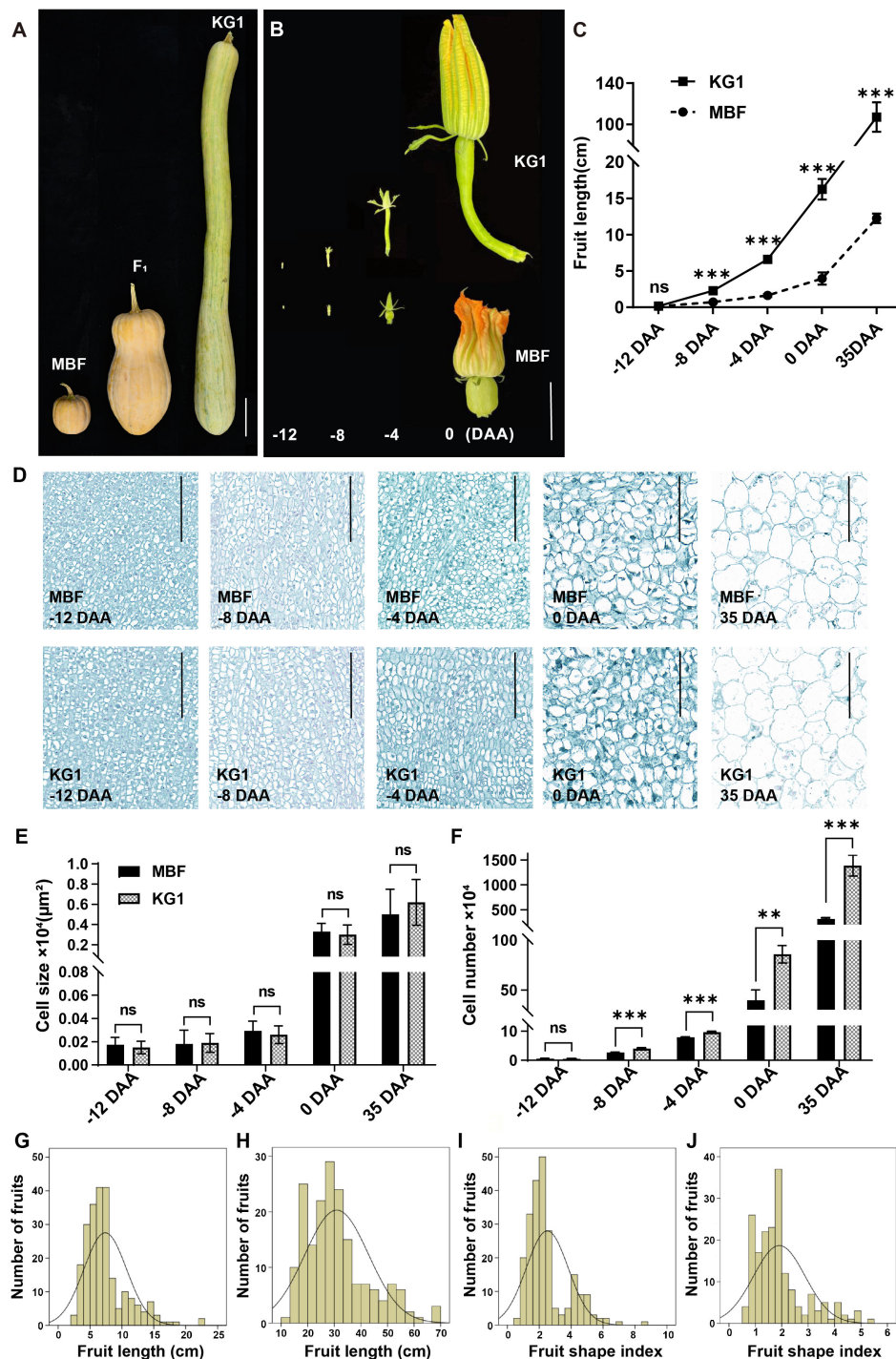
FIGURE 1

Phenotypic differences of fruit length in two parents. **(A)** The mature fruits of MBF, KG1 and F_1 generation. Scale bar, 10cm. **(B)** Dynamic changes of ovary morphology between KG1 and MBF from 12 days before anthesis to the day of anthesis. Scale bar, 5 cm. **(C)** The variations in fruit length of KG1 and MBF during ovary development. **(D)** Longitudinal section of fruit cells of MBF and KG1 at different developmental stages (-12 DAA, -8 DAA, -4 DAA, 0 DAA, 35 DAA). Scale bars, 100 µm. **(E)** Average cell size at different developmental stages of MBF and KG1. **(F)** Average number of cells at different developmental stages of MBF and KG1. **(G, H)** Frequency distribution of FL among F_2 individuals at 0 DAA **(G)** and 35 DAA **(H)**. **(F)**. **(I, J)** Frequency distribution of FSI among F_2 individuals at 0 DAA **(I)** and 35 DAA **(J)**. **(P<0.01), ***(P<0.001) indicates significant difference by Student's T-test, while ns shows non-significant different.

and 285,428 InDels were detected after removing SNPs and InDels with low coverage or discrepancies between the parental lines and the bulks. For identifying SNPs/InDels associated with the fruit length trait, the All index of the L-pool and S-pool were calculated and plotted to genome position (Figures 3A, B), and then Δ All index was calculated by subtracting the SNPs/indels index values of the two progenies pools (Figure 3C). Then, we identified the peak region above the threshold value as the candidate region for fruit
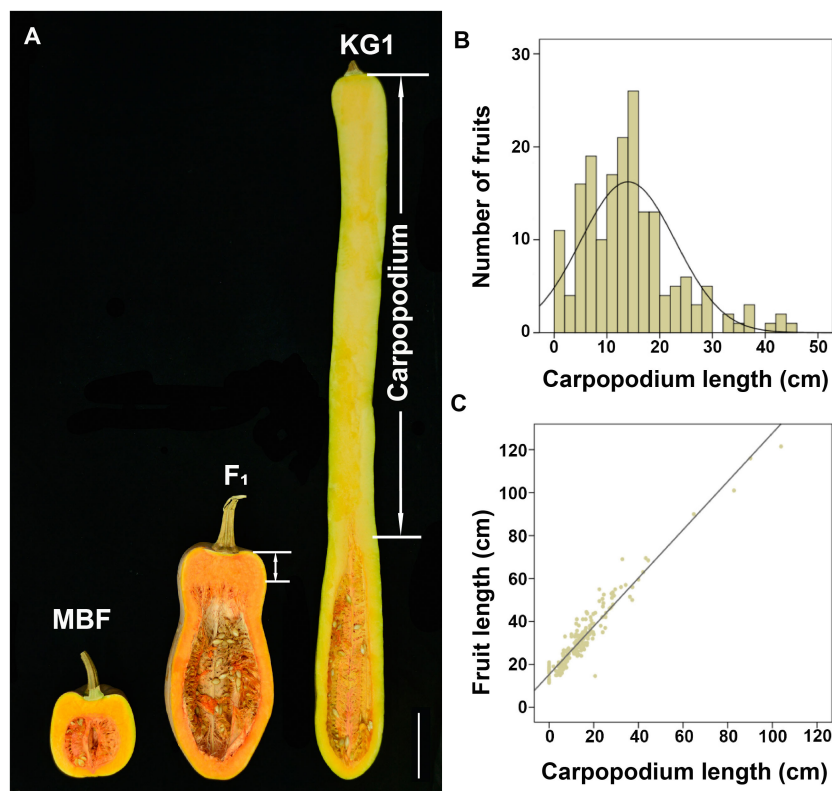
FIGURE 2

The correlation between fruit length and carpopodium length **(A)** Longitudinal section of mature fruit of MBF, KG1 and F$_1$ generation. Scale bar, 10cm. **(B)** Frequency distribution of mature fruit carpopodium length among F$_2$ individuals. **(C)** The correlation between fruit length and carpopodium length in F$_2$ population.

length trait. With the 95% significance level, we obtained 3 QTLs located on chromosome 4, 8 and 14 (Supplementary Table S8). Only one InDel was found in the candidate region of chromosome 4, and another was found in the candidate region of chromosome 8. Both InDels were located in intergenic regions. Thus, two QTLs on chromosomes 4 and 8 were excluded. We also analyzed the data with QTLseqr and calculated the G' value plotted to the genome position (Figure 3D). With the 95% significance level, a genome region (Cmo_Chr14: 159993–1750879) was obtained (Supplementary Table S9). Based on BSA results using different methods, the overlapping region (Cmo_Chr14: 718001–1711000) was referred to be the target region responsible for the length of pumpkin fruit.

Fine–mapping of candidate gene associated with pumpkin fruit length

In order to confirm and narrow down the genetic interval obtained by BSA-seq analysis, KASP markers were used for fine mapping based on the initial candidate region. A total of 18 KASP markers were developed in the candidate region, and 259 F$_2$ individual plants were used for screen the recombinants and phenotype-genotype association analysis. We found that F$_2$_89 and F$_2$_21 were

genotyped identically by the CmoM17 and CmoM18 markers, but their fruit length differed. Similarly, F$_2$_89 and F$_2$_273 did not differ in their fruit length, but were genotyped differently using CmoM17 and CmoM18 markers. Based on the phenotype-genotype association analysis of recombinants, we finally narrowed down the target region to a 50.77 kb window between markers CmoM17 (791,681bp) - CmoM18 (842,449bp) through these three critical recombinants (Figures 4A-C). There were eight annotation genes in this region, *CmoCh14G001760* (Reverse transcriptase (RNA-dependent DNA polymerase), putative), *CmoCh14G001770* (Putative serine/threonine-protein kinase receptor), *CmoCh14G001780* (Nuclear transcription factor Y subunit B-8-like), *CmoCh14G001790* (E3 ubiquitin-protein ligase), *CmoCh14G001800* (DNA helicase), *CmoCh14G001810* (Pentatricopeptide repeat-containing protein, putative), *CmoCh14G001820* (DNA helicase), *CmoCh14G001830* (Small subunit rib) (Figure 4D; Supplementary Table S10). Sequence variation analysis was performed to identify the variations of these eight genes between two parents. There was no sequence variation in the coding sequences between two parents. We also analyzed the variants in the 2 kb upstream promoter region of these genes, and only *CmoCh14G001780* and *CmoCh14G001790* were found to have several variations in the promoter region co-segregating with the phenotype, indicating that these variations might regulate fruit length by affecting gene expression.
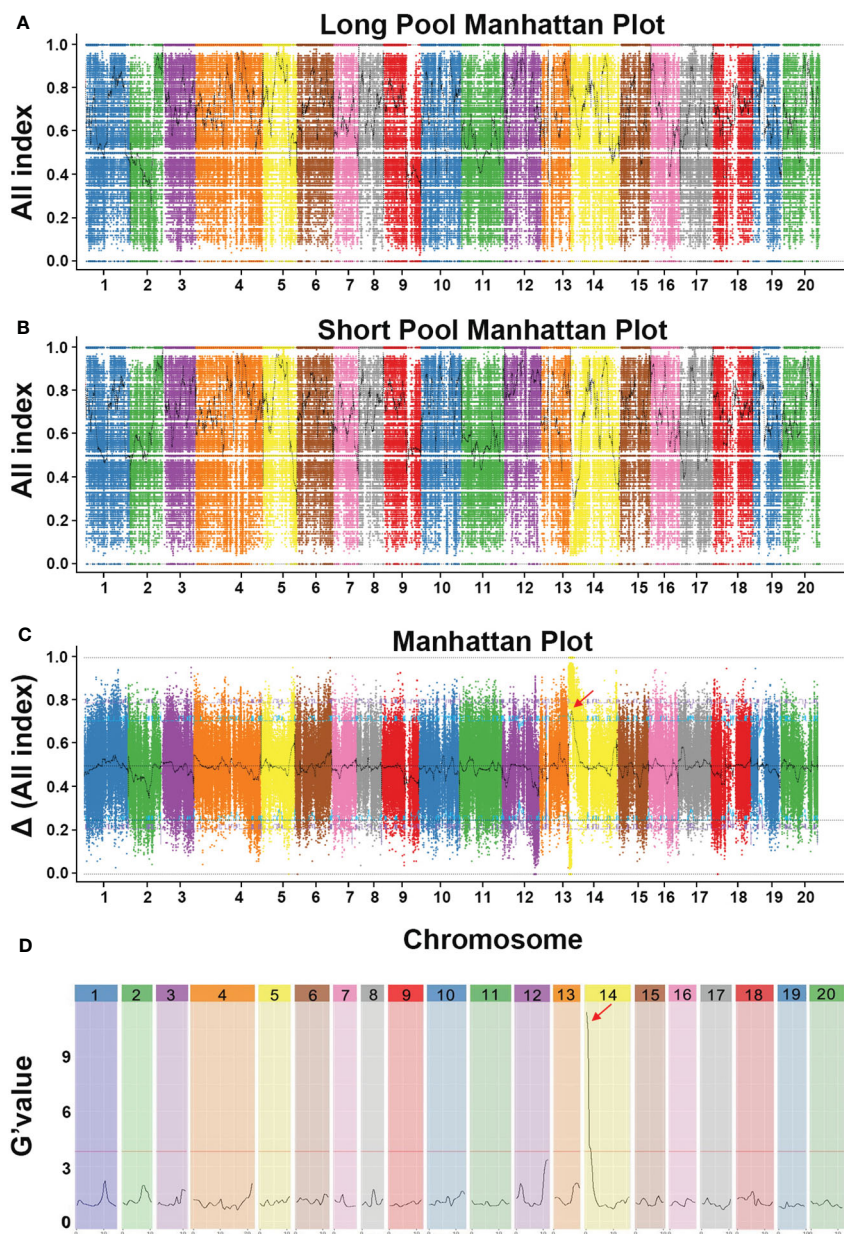
FIGURE 3
BSA-seq analysis results of the pumpkin fruit length. **(A)** Graph of the All-index of the long pool. **(B)** Graph of the All-index of the short pool.
(C) Graph of Δ All index values used for the association analysis. The x-axis indicates the 20 pumpkin chromosomes, and the y-axis indicates its
All-index value. Different chromosomes are distinguished by different colors. The black curve indicates the filter All index and Δ All index. The blue
line represents the 95% threshold line and the purple line represents the 99% threshold line. **(D)** Major quantitative trait loci for pumpkin fruit length
identified by QTLseqr.

Expression and promoter analysis of candidate gene

To further make sure which gene is responsible for the fruit
length, we analyzed the expression of *CmoCh14G001780* and
CmoCh14G001790. Since the shape of the pumpkin ovary has
been formed before flowering, and a significant difference in fruit
length was found between two parents at 0 DAA, we examined the
expression of these two candidate genes in the ovaries at four time
points (-12, -8, -4 and 0 DAA) before anthesis.

Expression analysis showed that the expression pattern of
CmoCh14G001780 in both parents showed a fluctuating trend that
did not correspond to the increase of fruit length (Figure 5A). While,
the expression pattern of *CmoCh14G001790* displayed an
approximately opposite trend with the development of ovary
(Figures 1C, 5B). The results indicated that in MBF the expression of
CmoCh14G001790 exhibited a gradual increase during the ovary
development, and it showed a gradual decrease in KG1 (Figure 5B).
Therefore, *CmoCh14G001790* was speculated as the most likely
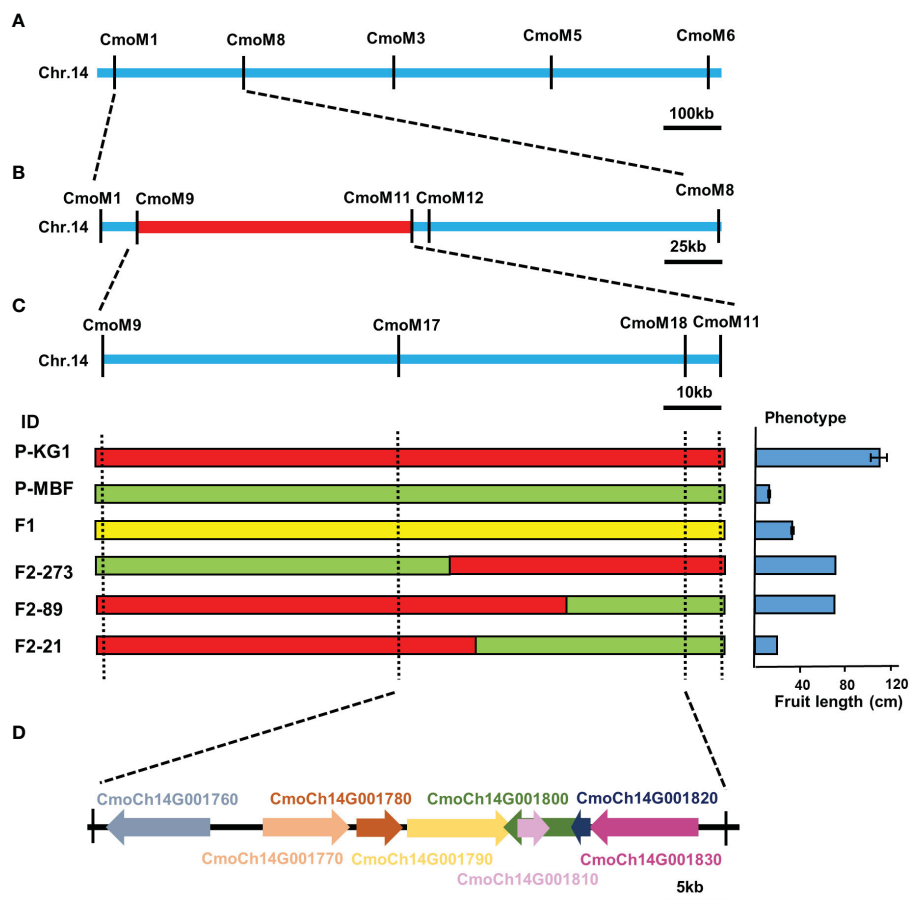candidate gene in regulating fruit length, which may exert a negative

FIGURE 4
Fine mapping of the candidate gene for pumpkin fruit length. **(A)** Primary mapping of target region using 259 F$_2$ individuals. The candidate gene was delimited to the region between markers CmoM1 and CmoM8. **(B)** Fine mapping of the candidate gene. The candidate gene was fine mapped in a 50.77kb region between markers CmoM17 and CmoM18. **(C)** Individuals with chromosome segment substitution on the target region, red indicates genotype of KG1, green indicates genotype of MBF, and yellow indicates heterozygous region. **(D)** In the mapping interval, there are eight genes, which may be candidate genes for target traits.

influence on this trait. We named this gene as *CmoFL1*(*fruit length 1*), which was predicted to encode an E3 ubiquitin-protein ligase.

Through promoter sequence analysis, we found there was a SNP and an InDel in the promoter of *CmoFL1* co-segregating with the phenotype (Figure 6A). The SNP in the promoter of *CmoFL1* was located in a GT1CONSENSUS (GGAAAA), which disappeared when it was mutated from G to T (Figure 6B). GT1CONSENSUS were reported to involved in the regulation of transcription of many genes and can act as a transcription activator or repressor (Villain et al., 1996; Buchel et al., 1999; Zhou, 1999). To further study the SNP in the promoter of *CmoFL1* and its relevance to fruit length, a KASP marker based on this SNP was developed for genotyping, named CmoM19. Among F$_2$ individuals, we found that the fruit length in genotype GG (homozygous genotype with MBF) was significantly lower than TT (homozygous genotype with KG1), and the fruit length of heterozygous genotype G/T also showed significant different for that of GG (Figure 6C). To explore how the variations in the promoter region affect the expression of *CmoFL1*, we further analyzed the promoter activity of *CmoFL1* between two parents. We performed DUAL-LUC assay to explore whether the variation of GT1CONSENSUS affect the promoter activity. The results showed that the promoter activity of

CmoFL1 in the short fruit parent MBF was significantly higher than that in the long fruit parent KG1 (Figure 6D), indicating the absence of GT1CONSENSUS in KG1 could reduce the promoter activity and thus affect fruit length. Collectively, these results strongly suggested that *CmoFL1* was the candidate gene for the fruit length of pumpkin.

Transcriptomic profiles associated with cell division

To explore the regulatory pathway of *CmoFL1* gene regulating the fruit length in pumpkin, we performed a transcriptome analysis during fruit development. The samples of long-fruit KG1 and short-fruit MBF at -12 DAA and 0 DAA were collected for RNA-seq analysis. In total, we obtained 76.49 Gb clean data with Q20 ≥ 97.90%, Q30 ≥ 94.07%, and a 44.78%-45.34% G/C ratio (Supplementary Table S11). Then, we performed differential expression analysis between KG1 and MBF pumpkins at different developmental stages. The results showed that there were 1755 (505 up-regulated and 1250 down-regulated) and 6730 (3694 up-regulated and 3036 down-regulated) differentially expressed genes
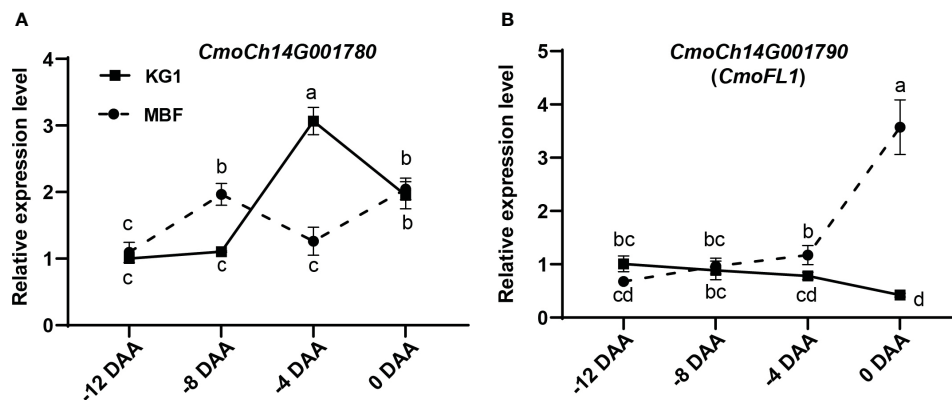
FIGURE 5
Expression analysis of candidate genes in different stages of fruit development. **(A, B)** Expression of candidate genes *CmoCh14G001780* **(A)** and *CmoCh14G001790* **(B)** during ovary development. Expression analysis suggested that *CmoCh14G001790* was the most likely candidate gene controlling pumpkin fruit length. Different letters indicated statistical significance (P < 0.05) as determined by one-way ANOVA.

(DEG) between two parents at -12 DAA and 0 DAA, respectively (Supplementary Table S12), indicating that the number of DEGs increased with the development of fruit length and was consistent with the phenotypic differences of fruit length development (Figures 1B, C). Then, Gene Ontology (GO) enrichment of the DEGs was performed between KG1 and MBF at 0 DAA. We found that the GO terms of the biological process categories "microtubule-based movement" and "microtubule-based process" and the molecular function categories "microtubule motor activity", "tubulin binding" and "microtubule binding" were most significantly enriched at 0 DAA (Figure 7A). Microtubules as an important cytoskeleton were reported participate in cell growth and cell division, and were regulated by microtubule associated proteins (Kaloriti et al., 2007; Sedbrook and Kaloriti, 2008; Zeng et al., 2009; Schaefer et al., 2017). These results suggested that the transcript levels of genes associated with microtubule were different between long-fruit KG1 and short-fruit MBF at 0 DAA, indicating that the formation of fruit length may be related to the cell division process in which microtubules were involved, as the cell number of KG1 was significantly higher than that of MBF at 0 DAA (Figure 1F).

Cell division was a highly coordinated process in which many plant cell cycle regulators have been reported to be involved (Vandepoele et al., 2002). To determine whether *CmoFL1* involved in regulating the cell division, we analyzed the orthologous genes in *Cucurbita moschata* that have been reported to be involved in cell cycle and cell division in *Arabidopsis* or maize (Supplementary Table S13), and determined the gene expression patterns during fruit development. Among the DEGs, CDKs and cyclins showed higher expression levels in both parents at -12 DAA, and significantly lower expression levels in short-fruit MBF at 0 DAA, which was opposite to the expression pattern of *CmoFL1* (Figure 7B). The CDKs family has been reported to drive cell through cell cycle phase, and it must bind to cyclins to be activated (Vandepoele et al., 2002). Four transcription factors (*E2F*, *OBP1*, *TCP15*, *ORC1*) involved in cell cycle and their downstream genes (*MCM3*, *CDC6*, *CDT1A*) (Berckmans and De Veylder, 2009) also showed significantly lower expression levels in short-fruit MBF at 0 DAA (Figure 7B). Cell cycle inhibitor Kip-related protein 4 (*KRP4*) and transcription factor

NTM1 were reported as the negative regulators of cell division (Kim et al., 2006; Vieira et al., 2012), and Cell Number Regulator1 (*ZmCNR01*) in maize was identified negatively regulated cell number and organ size (Guo et al., 2010). The cluster analysis showed that these negative regulators exhibited similar expression patterns to *CmoFL1* during fruit development. They all showed lower expression levels at -12 DAA, higher expression levels at 0 DAA, and significantly higher expression in short-fruit MBF than that of long-fruit KG1 at 0 DAA (Figure 7B). The results indicated that *CmoFL1* may negatively regulate cell division to influence the fruit length in pumpkin.

Discussion

Fruit length is an important crop commercial trait, which plays a vital role in fruit yield, external quality, consumer preference, packaging and transportation. The length of fruit directly affects the shape of fruit, which determines whether the fruit is oblate, round, oval, strip or other shapes, and also affects the weight of fruit, emphasizing the importance of fruit length trait in crop breeding. Previous studies have shown that in Cucurbitaceae crops, the shape of ovary at anthesis is highly correlated with the shape of mature fruit. The shape of fruit can be predicted before ovary pollination, which has been demonstrated in watermelon and melon (Dou et al., 2018; Ma et al., 2022), and a similar mechanism in pumpkin was found in this study. There was a significant correlation between FL and FSI in the parental line and F_2 population at 0 DAA and 35 DAA (Supplementary Table S4), and a significant difference in ovarian length between parental lines at 8 days before anthesis (Figures 1B, C).

Fruit carpopodium length is an important quality trait, which affects fruit length and commercial value, and longer carpopodium will increase fruit length. Some important genes related to carpopodium length have been detected in cucumbers successively. *CsFnl7.1* encoding a late embryogenesis abundant protein, was identified to control the fruit neck length in cucumber through bulked segregant analysis (Xu et al., 2020). *Cucumis sativus HECATE1* (*CsHEC1*) was found to be highly expressed in fruit neck and interacted with CsOVATE to
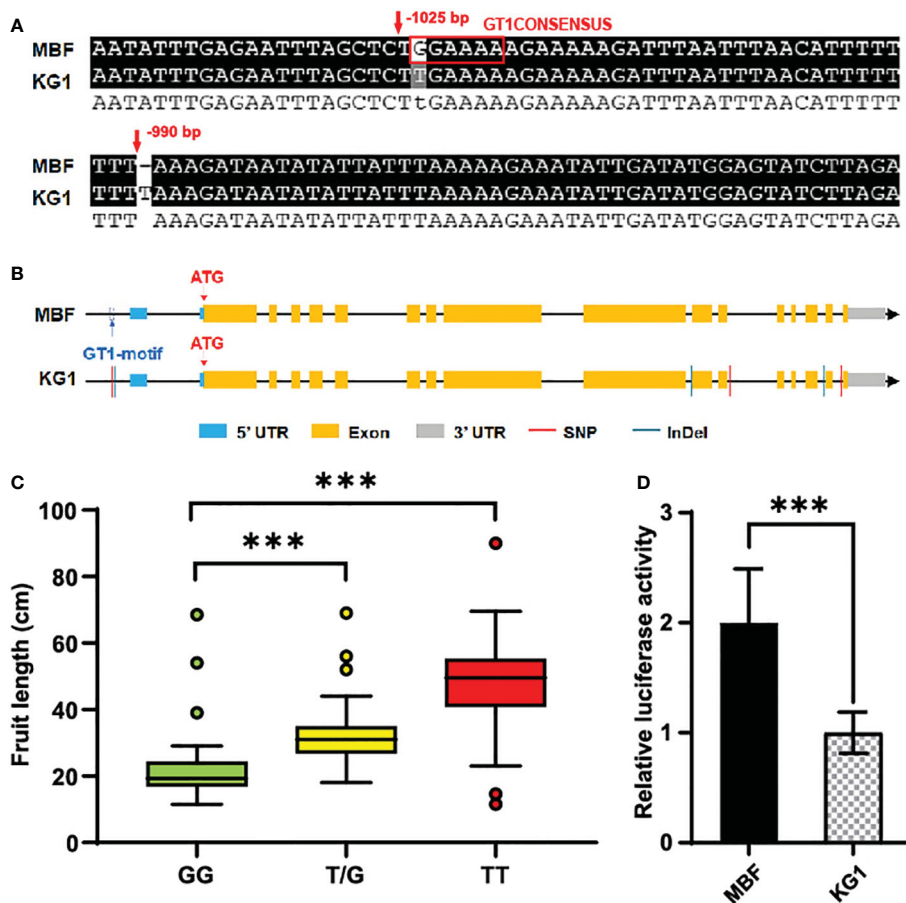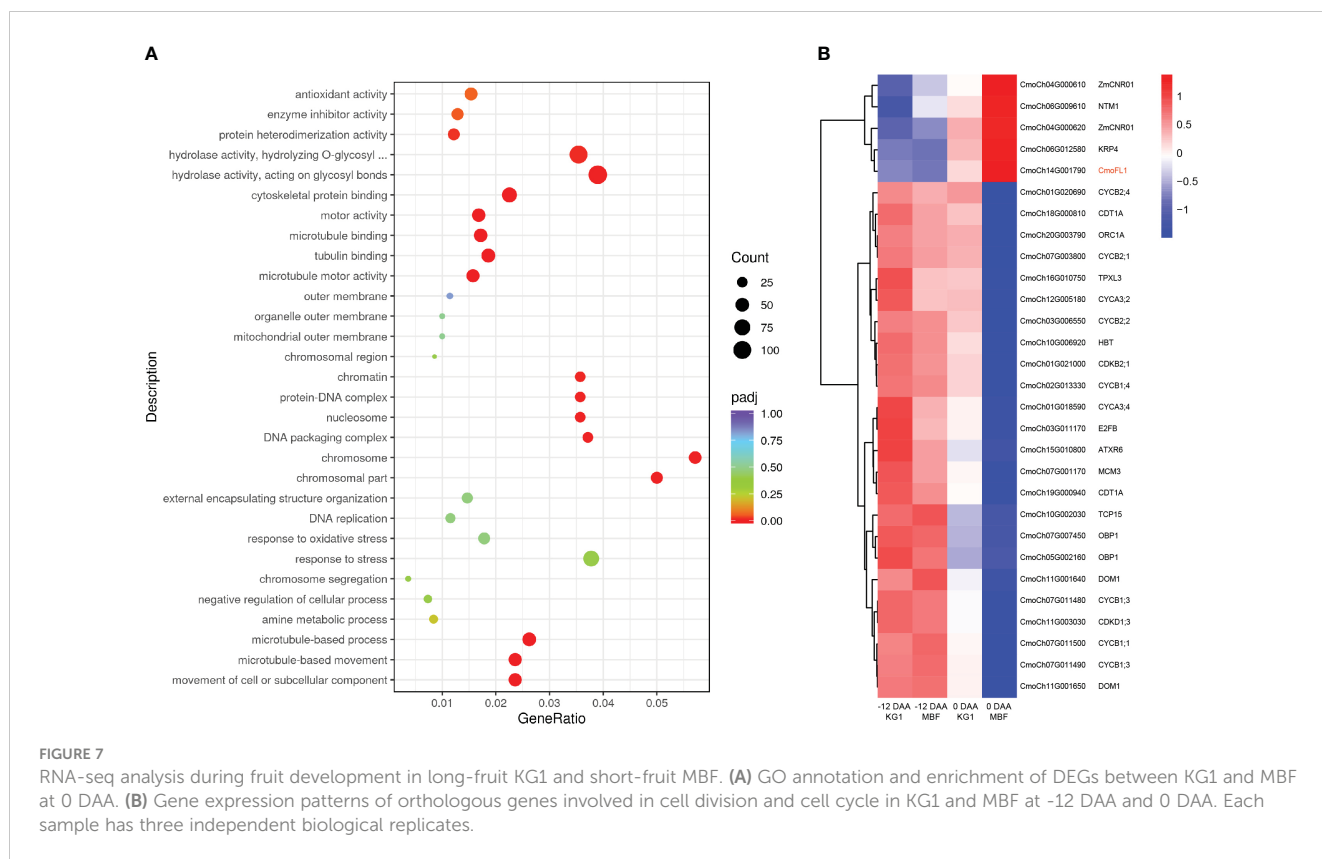
FIGURE 6
Promoter variations of candidate gene. **(A)** Partial results of sequence alignment of promoter region of *CmoFL1* between MBF and KG1. **(B)** Structure and sequence alignment of *CmoFL1* between MBF and KG1. **(C)** Association analysis between *CmoFL1* genotypes based on CmoM19 and fruit length in F₂ individuals. GG indicates homozygous type of MBF, TT indicates homozygous type of KG1 and G/T indicates heterozygous genotypes. **(D)** DUAL-LUC assay of promoter activity in MBF and KG1. The relative value of LUC/REN indicates promoter activity. ***(P < 0.001) indicates significant difference by Student's T-test.

regulate the fruit neck length variation via modulating auxin biosynthesis (Wang et al., 2022). However, the neck length had a limited effect on the fruit length in cucumber. In pumpkin, we found that the long-fruited pumpkin KG1 has the extra-long carpopodium, while short-fruited pumpkin MBF has no carpopodium (Figure 2A). The correlation analysis between fruit length and carpopodium length showed a significant positive correlation between parents and F₂ population (Figure 2C), suggesting that the genes associated with fruit length may also regulate the carpopodium length.

Protein ubiquitination plays key roles in multiple plant development stages and several abiotic stress responses. Importantly, diverse E3 ligases are involved in these regulatory pathways by mediating phytohormone and light signaling or other pathways (Shu and Yang, 2017). Extensive studies have shown that E3 ligases were involved in the morphological regulation of fruits and seeds. In rice, *GW2* encoding a previously unknown RING-type protein with E3 ubiquitin ligase activity, was found to play a negative role in regulating cell division in determining rice grain width and weight (Song et al., 2007). *Chang Li Geng 1* (CLG1), which encodes an E3 ligase, regulates grain size by targeting and degrading

the negative grain length regulator Gγ protein GS3 (Yang et al., 2021). In wheat, *TaGW2* is an orthologue of rice gene *OsGW2*, which is also negatively regulates grain width and weight (Hong et al., 2014; Simmonds et al., 2016; Liu et al., 2019). TaSDIR1–4A was confirmed as a negative regulator of grain size. Silencing of TaSDIR1–4A (a RING-type E3 ubiquitin ligase) led to an increase in wheat TGW (1000-grain weight), while overexpression of TaSDIR1–4A in Arabidopsis resulted in a decrease in seed size (Wang et al., 2020a). In *Brassica napus*, variation in the promoter region of the HECT E3 ligase gene *BnaUPL3.C03* made a major contribution to variation in seed weight per pod (Miller et al., 2019). In addition, E3 ligases were also found to play important roles in regulating fruit size. It is reported that the RING-type E3 ligase short fruit 1 (SF1) can ubiquitinate and degrade both itself and 1-aminocyclopropane-1-carboxylate synthase 2 (ACS2) to tightly control ethylene synthesis for dose-dependent cell division and fruit elongation in cucumber (Xin et al., 2019). In addition, UV-damaged DNA binding protein 1 (DDB1), the core component of Cullin4-RING E3 ubiquitin ligase complex (CRL4), has been shown to be profoundly involved in the cell division process (Iovine et al., 2011). In

FIGURE 7
RNA-seq analysis during fruit development in long-fruit KG1 and short-fruit MBF. **(A)** GO annotation and enrichment of DEGs between KG1 and MBF at 0 DAA. **(B)** Gene expression patterns of orthologous genes involved in cell division and cell cycle in KG1 and MBF at -12 DAA and 0 DAA. Each sample has three independent biological replicates.

the present study, we identified that *CmoFL1* encoding E3 ubiquitin ligase was a candidate gene for the fruit length of pumpkin. The expression analysis revealed that the expression pattern of *CmoFL1* in the MBF exhibited a gradual increase during ovary development, while it showed a gradual decrease in KG1 (Figure 5B). In addition, the result of paraffin section indicated that the cell number of KG1 fruit was significantly higher than that of MBF fruit (Figure 1F). RNA-seq analysis suggested that *CmoFL1* showed consistent expression patterns with cell division negative regulators (Figure 7B). These results suggested that high expression level of *CmoFL1* may inhibit the increase of cell number in MBF. Therefore, we speculated that the regulation pattern of *CmoFL1* in pumpkin fruit length may be similar to other E3 ligases, which negatively regulate cell division and thus affect the variation of fruit length.

In conclusion, we reported the potential candidate gene *CmoFL1* encoding E3 ubiquitin ligase was responsible for the fruit length of pumpkin. The results of this study can provide a reference for understanding the genetic mechanism of pumpkin fruit morphogenesis and will also enable marker-assisted breeding of pumpkin fruit length improvement.

Data availability statement

The dataset presented in the study can be found in online repository (http://www.ncbi.nlm.nih.gov), accession number PRJNA1099576.

Author contributions

YMZ: Data curation, Formal analysis, Writing – original draft. MZ: Methodology, Data curation, Writing – original draft. QS: Data curation, Formal analysis, Writing – original draft. MYZ: Data curation, Formal analysis, Writing – original draft. CW: Data curation, Formal analysis, Writing – original draft. YTZ: Data curation, Investigation, Writing – original draft. QY: Data curation, Investigation, Writing – original draft. YB: Data curation, Investigation, Writing – original draft. ZH: Data curation, Supervision, Writing – review & editing. JY: Data curation, Supervision, Writing – review & editing. MFZ: Data curation, Supervision, Writing – review & editing. XL: Data curation, Visualization, Writing – review & editing.

Funding

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.2024.1408602/full#supplementary-material

References

Berckmans, B., and De Veylder, L. (2009). Transcriptional control of the cell cycle. *Curr. Opin. Plant Biol.* 12, 599–605. doi: 10.1016/j.pbi.2009.07.005

Brown, J., Pirrung, M., and McCue, L. A. (2017). FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics* 33, 3137–3139. doi: 10.1093/bioinformatics/btx373

Buchel, A. S., Brederode, F. T., Bol, J. F., and Linthorst, H. J. M. (1999). Mutation of GT-1 binding sites in the *Pr-1A* promoter influences the level of inducible gene expression in *vivo*. *Plant Mol. Biol.* 40, 387–396. doi: 10.1023/A:1006144505121

Cheng, Z., Liu, Z., Xu, Y., Ma, L., Chen, J., Gou, J., et al. (2021). Fine mapping and identification of the candidate gene *BFS* for fruit shape in wax gourd (*Benincasa hispida*). *Theor. Appl. Genet.* 134, 3983–3995. doi: 10.1007/s00122-021-03942-8

Dou, J. L., Zhao, S. J., Lu, X. Q., He, N., Zhang, L., Ali, A., et al. (2018). Genetic mapping reveals a candidate gene (*ClFS1*) for fruit shape in watermelon (*Citrullus lanatus* L.). *Theor. Appl. Genet.* 131, 947–958. doi: 10.1007/s00122-018-3050-5

Esteras, C., Gómez, P., Monforte, A. J., Blanca, J., Vicente-Dólera, N., Roig, C., et al. (2012). High-throughput SNP genotyping in *Cucurbita pepo* for map construction and quantitative trait loci mapping. *BMC Genomics* 13, 80. doi: 10.1186/1471-2164-13-80

Ezin, V., Gbemenou, U. H., and Ahanchede, A. (2022). Characterization of cultivated pumpkin (*Cucurbita moschata* Duchesne) landraces for genotypic variance, heritability and agro-morphological traits. *Saudi J. Biol. Sci.* 29, 3661–3674. doi: 10.1016/j.sjbs.2022.02.057

Guo, M., Rupe, M. A., Dieter, J. A., Zou, J. J., Spielbauer, D., Duncan, K. E., et al. (2010). *Cell number regulator1* affects plant and organ size in maize: implications for crop yield enhancement and heterosis. *Plant Cell* 22, 1057–1073. doi: 10.1105/tpc.109.073676

Han, X. X., Min, Z. Y., Wei, M. H., Li, Y. Q., Wang, D. H., Zhang, Z. Q., et al. (2022). QTL mapping for pumpkin fruit traits using a GBS-based high-density genetic map. *Euphytica* 218, 106. doi: 10.1007/s10681-022-03057-1

Hong, Y., Chen, L., Du, L., Su, Z., Wang, J., Ye, X., et al. (2014). Transcript suppression of *TaGW2* increased grain width and weight in bread wheat. *Funct. Integr. Genomics* 14, 341–349. doi: 10.1007/s10142-014-0380-5

Iovine, B., Iannella, M. L., and Bevilacqua, M. A. (2011). Damage-specific DNA binding protein 1 (DDB1): a protein with a wide range of functions. *Int. J. Biochem. Cell Biol.* 43, 1664–1667. doi: 10.1016/j.biocel.2011.09.001

Kaloriti, D., Galva, C., Parupalli, C., Khalifa, N., Galvin, M., and Sedbrook, J. C. (2007). Microtubule associated proteins in plants and the processes they manage. *J. Integr. Plant Biol.* 49, 1164–1173. doi: 10.1111/j.1672-9072.2007.00541.x

Kim, Y. S., Kim, S. G., Park, J. E., Park, H. Y., Lim, M. H., Chua, N. H., et al. (2006). A membrane-bound NAC transcription factor regulates cell division in *Arabidopsis*. *Plant Cell* 18, 3132–3144. doi: 10.1105/tpc.106.043018

Legendre, R., Kuzy, J., and McGregor, C. (2020). Markers for selection of three alleles of *ClSUN25–26-27a* (*Cla011257*) associated with fruit shape in watermelon. *Mol. Breed.* 40, 19. doi: 10.1007/s11032-020-1104-2

Liu, S., Gao, P., Zhu, Q. L., Zhu, Z. C., Liu, H. Y., Wang, X. Z., et al. (2020). Resequencing of 297 melon accessions reveals the genomic history of improvement and loci related to fruit traits in melon. *Plant Biotechnol. J.* 18, 2545–2558. doi: 10.1111/pbi.13434

Liu, H., Li, H., Hao, C., Wang, K., Wang, Y., Qin, L., et al. (2019). *TaDA1*, a conserved negative regulator of kernel size, has an additive effect with *TaGW2* in common wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* 18, 1330–1342. doi: 10.1111/pbi.13298

Liu, J. P., Van Eck, J., Cong, B., and Tanksley, S. D. (2002). A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc. Natl. Acad. Sci. United States America* 99, 13302–13306. doi: 10.1073/pnas.162485999

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Lyu, X., Shi, L., Zhao, M., Li, Z., Liao, N., Meng, Y., et al. (2022). A natural mutation of the *NST1* gene arrests secondary cell wall biosynthesis in the seed coat of a hull-less pumpkin accession. *Horticulture Res.* 9, uhac136. doi: 10.1093/hr/uhac136

Ma, J., Li, C. C., Zong, M., Qiu, Y. H., Liu, Y. M., Huang, Y. T., et al. (2022). *CmFSI8/CmOFP13* encoding an OVATE family protein controls fruit shape in melon. *J. Exp. Bot.* 73, 1370–1384. doi: 10.1093/jxb/erab510

Mansfeld, B. N., and Grumet, R. (2018). QTLseqr: an R package for bulk segregant analysis with next-generation sequencing. *Plant Genome* 11, 180006. doi: 10.3835/plantgenome2018.01.0006

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Miller, C., Wells, R., McKenzie, N., Trick, M., Ball, J., Fatihi, A., et al. (2019). Variation in expression of the HECT E3 ligase *UPL3* modulates LEC2 levels, seed size, and crop yields in brassica napus. *Plant Cell* 31, 2370–2385. doi: 10.1105/tpc.18.00577

Montero-Pau, J., Blanca, J., Esteras, C., Martínez-Pérez, E. M., Gómez, P., Monforte, A. J., et al. (2017). An SNP-based saturated genetic map and QTL analysis of fruit-related traits in Zucchini using Genotyping-by-sequencing. *BMC Genomics* 18, 94. doi: 10.1186/s12864-016-3439-y

Oliveira, R. L., Goncalves, L. S. A., Rodrigues, R., Baba, V. Y., Sudre, C. P., d., M. H., et al. (2016). Genetic divergence among pumpkin landraces. *Semina: Cienc. Agrar.* 37, 547–556. doi: 10.5433/1679-0359.2016v37n2p547

Pan, Y. P., Liang, X. J., Gao, M. L., Liu, H. Q., Meng, H. W., Weng, Y. Q., et al. (2017). Round fruit shape in WI7239 cucumber is controlled by two interacting quantitative trait loci with one putatively encoding a tomato SUN homolog. *Theor. Appl. Genet.* 130, 573–586. doi: 10.1007/s00122-016-2836-6

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

Paris, H. S. (2018). Consumer-oriented exploitation and conservation of genetic resources of pumpkins and squash, Cucurbita. *Israel J. Plant Sci.* 65, 202–221. doi: 10.1163/22238980-00001036

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Porebski, S., Bailey, L. G., and Baum, B. R. (1997). Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Rep.* 15, 8–15. doi: 10.1007/BF02772108

Ren, X., Li, H., Yin, Y., Duan, L., Wang, Y., Liang, X., et al. (2022). Genetic analysis of fruit traits in wolfberry (*Lycium* L.) by the major gene plus polygene model. *Agronomy* 12, 1403. doi: 10.3390/agronomy12061403

Sandlin, K., Prothro, J., Heesacker, A., Khalilian, N., Okashah, R., Xiang, W., et al. (2012). Comparative mapping in watermelon [*Citrullus lanatus* (Thunb.) Matsum. et Nakai]. *Theor. Appl. Genet.* 125, 1603–1618. doi: 10.1007/s00122-012-1938-z

Schaefer, E., Belcram, K., Uyttewaal, M., Duroc, Y., Goussot, M., Legland, D., et al. (2017). The preprophase band of microtubules controls the robustness of division orientation in plants. *Science* 356, 186–189. doi: 10.1126/science.aal3016

Sedbrook, J. C., and Kaloriti, D. (2008). Microtubules, MAPs and plant directional cell expansion. *Trends Plant Sci.* 13, 303–310. doi: 10.1016/j.tplants.2008.04.002

Shu, K., and Yang, W. Y. (2017). E3 ubiquitin ligases: ubiquitous actors in plant development and abiotic stress responses. *Plant Cell Physiol.* 58, 1461–1476. doi: 10.1093/pcp/pcx071

Simmonds, J., Scott, P., Brinton, J., Mestre, T. C., Bush, M., del Blanco, A., et al. (2016). A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theor. Appl. Genet.* 129, 1099–1112. doi: 10.1007/s00122-016-2686-2

Song, X., Huang, W., Shi, M., Zhu, M., and Lin, H. (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39, 623–630. doi: 10.1038/ng2014

Sun, H., Wu, S., Zhang, G., Jiao, C., Guo, S., Ren, Y., et al. (2017). Karyotype stability and unbiased fractionation in the paleo-allotetraploid cucurbita genomes. *Mol. Plant* 10, 1293–1306. doi: 10.1016/j.molp.2017.09.003

Switzenberg, J. A., Beaudry, R. M., and Grumet, R. (2015). Effect of CRC::*etr1–1* transgene expression on ethylene production, sex expression, fruit set and fruit ripening in transgenic melon (*Cucumis melo* L.). *Transgenic Res.* 24, 497–507. doi: 10.1007/s11248-014-9853-5

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., et al. (2013). QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* 74, 174–183. doi: 10.1111/tpj.12105

Tsaballa, A., Pasentsis, K., Darzentas, N., and Tsaftaris, A. S. (2011). Multiple evidence for the role of an *Ovate*-like gene in determining fruit shape in pepper. *BMC Plant Biol.* 11, 46. doi: 10.1186/1471-2229-11-46

Vandepoele, K., Raes, J., De Veylder, L., Rouzé, P., Rombauts, S., and Inzé, D. (2002). Genome-wide analysis of core cell cycle genes in Arabidopsis. *Plant Cell* 14, 903–916. doi: 10.1105/tpc.010445

van der Knaap, E., Chakrabarti, M., Chu, Y. H., Clevenger, J. P., Illa-Berenguer, E., Huang, Z. J., et al. (2014). What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00227

Vieira, P., Engler, G., and Engler, J. D. (2012). Whole-mount confocal imaging of nuclei in giant feeding cells induced by root-knot nematodes in Arabidopsis. *New Phytol.* 195, 488–496. doi: 10.1111/j.1469-8137.2012.04175.x

Villain, P., Mache, R., and Zhou, D. X. (1996). The mechanism of GT element-mediated cell type-specific transcriptional control. *J. Biol. Chem.* 271, 32593–32598. doi: 10.1074/jbc.271.51.32593

Wang, Y., Bo, K., Gu, X., Pan, J., Li, Y., Chen, J., et al. (2020c). Molecularly tagged genes and quantitative trait loci in cucumber with recommendations for QTL nomenclature. *Horticulture Res.* 7, 3. doi: 10.1038/s41438-019-0226-3

Wang, X. F., Li, H., Gao, Z. H., Wang, L. N., and Ren, Z. H. (2020b). Localization of quantitative trait loci for cucumber fruit shape by a population of chromosome segment substitution lines. *Sci. Rep.* 10, 11030. doi: 10.1038/s41598-020-68312-8

Wang, K., Li, M. Y., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38, e164. doi: 10.1093/nar/gkq603

Wang, J. Y., Wang, R. T., Mao, X. G., Zhang, J. L., Liu, Y. N., Xie, Q., et al. (2020a). RING finger ubiquitin E3 ligase gene *TaSDIR1–4A* contributes to determination of grain size in common wheat. *J. Exp. Bot.* 71, 5377–5388. doi: 10.1093/jxb/eraa271

Wang, Z. Y., Zhou, Z. Y., Wang, L. M., Yan, S. S., Cheng, Z. H., Liu, X. F., et al. (2022). The CsHEC1-CsOVATE module contributes to fruit neck length variation via modulating auxin biosynthesis in cucumber. *Proc. Natl. Acad. Sci. United States America* 119, e2209717119. doi: 10.1073/pnas.2209717119

Wei, Q., Wang, J., Wang, W., Hu, T., Hu, H., and Bao, C. (2020). A high-quality chromosome-level genome assembly reveals genetics for important traits in eggplant. *Horticulture Res.* 7, 153. doi: 10.1038/s41438-020-00391-0

Wu, S., Xiao, H., Cabrera, A., Meulia, T., and van der Knaap, E. (2011). SUN regulates vegetative and reproductive organ shape by changing cell division patterns. *Plant Physiol.* 157, 1175–1186. doi: 10.1104/pp.111.181065

Wu, S., Zhang, B., Keyhaninejad, N., Rodríguez, G. R., Kim, H. J., Chakrabarti, M., et al. (2018). A common genetic mechanism underlies morphological diversity in fruits and other plant organs. *Nat. Commun.* 9, 4734. doi: 10.1038/s41467-018-07216-8

Xin, T., Zhang, Z., Li, S., Zhang, S., Li, Q., Zhang, Z., et al. (2019). Genetic regulation of ethylene dosage for cucumber fruit elongation. *Plant Cell* 31, 1063–1076. doi: 10.1105/tpc.18.00957

Xu, X., Wei, C., Liu, Q., Qu, W., Qi, X., Xu, Q., et al. (2020). The major-effect quantitative trait locus Fnl7.1 encodes a late embryogenesis abundant protein associated with fruit neck length in cucumber. *Plant Biotechnol. J.* 18, 1598–1609. doi: 10.1111/pbi.13326

Yan, W., Zhen-Shan, L., Xiao-Qin, Y., Zhi-Yi, W., Lan, M., Hong-Xia, T., et al. (2023). Inheritance analysis of fruit-related traits in Chinese cherry [*Cerasus pseudocerasus* (Lindl.) G.Don] breeding progenies. *Scientia Hortic.* 307, 111519. doi: 10.1016/j.scienta.2022.111519

Yang, W. S., Wu, K., Wang, B., Liu, H. H., Guo, S. Y., Guo, X. Y., et al. (2021). The RING E3 ligase CLG1 targets GS3 for degradation via the endosome pathway to determine grain size in rice. *Mol. Plant* 14, 1699–1713. doi: 10.1016/j.molp.2021.06.027

Zeng, C. J. T., Lee, Y. R. J., and Liu, B. (2009). The WD40 Repeat Protein NEDD1 Functions in Microtubule Organization during Cell Division in *Arabidopsis thaliana*. *Plant Cell* 21, 1129–1140. doi: 10.1105/tpc.109.065953

Zhang, Z., Wang, B. W., Wang, S. H., Lin, T., Yang, L., Zhao, Z. L., et al. (2020). Genome-wide target mapping shows histone deacetylase complex1 regulates cell proliferation in cucumber fruit. *Plant Physiol.* 182, 167–184. doi: 10.1104/pp.19.00532

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

Zhou, D. X. (1999). Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends Plant Sci.* 4, 210–214. doi: 10.1016/S1360-1385(99)01418-1