Fine mapping and identification of the candidate gene BFS for fruit shape in wax gourd (*Benincasa hispida*)

Zhikui Cheng¹ · Zhengguo Liu¹ · Yuanchao Xu² · Lianlian Ma¹ · Jiyeing Chen¹ · Jiquan Gou¹ · Liwen Su¹ · Wenting Wu¹ · Yong Chen⁴ · Wenjin Yu¹ · Peng Wang¹,³

Received: 3 June 2021 / Accepted: 18 August 2021 / Published online: 4 September 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

**Key message** Non-synonymous mutations in the BFS gene, which encodes the IQD protein, are responsible for the shape of wax gourd fruits.

**Abstract** Fruit shape is an important agronomic trait in wax gourds. Therefore, in this study, we employed bulked segregant analysis (BSA) to identify a candidate gene for fruit shape in wax gourds within F₂ populations derived by crossing GX-71 (long cylindrical fruit, fruit shape index = 4.56) and MY-1 (round fruit, fruit shape index = 1.06) genotypes. According to BSA, the candidate gene is located in the 17.18 Mb region on chromosome 2. Meanwhile, competitive allele-specific PCR (KASP) markers were used to reduce it to a 19.6 Kb region. Only one gene was present within the corresponding region of the reference genome, namely Bch02G016830 (designated BFS). Subsequently, BFS was sequenced in six wax gourd varieties with different fruit shapes. Sequence analysis revealed two non-synonymous mutations in the round wax gourd and one non-synonymous mutation in the cylindrical wax gourd. Quantitative real-time PCR (qRT-PCR) analysis further showed that the expression of BFS in round fruits was significantly higher than in long cylindrical fruits at the ovary formation stage. Therefore, BFS is a candidate gene for determination wax gourd shape. The predicted protein encoded by the BFS gene belongs to the IQ67-domain protein family, which have the structural characteristics of scaffold proteins and coordinate Ca²⁺ CaM signaling from the membrane to the nucleus. Ultimately, two derived cleaved amplified polymorphic sequence (dCAPS) markers were developed to facilitate marker-assisted selection for wax gourds breeding.

Introduction

Wax gourds [*Benincasa hispida* (Thunb) Cogn. (2n = 2x = 24)] are an important crop species in the Cucurbitaceae family that represent a traditional vegetable variety used as a food source and for medicinal purposes; they are widely distributed in the tropical, subtropical, and temperate regions of Asia (Gu et al. 2013). The shape of wax gourds is a vital characteristic with a broad range of phenotypic variations. Consequently, fruit shape is an important trait in breeding programs.

Shape is one of the main indices used for quality evaluation and market classification of horticultural crops, including fruits and vegetables. The genes and quantitative trait loci (QTLs) that control fruit shape have been identified in many crops. For instance, in tomatoes, the regulation of multiple loci leads to numerous variations in fruit shape. In fact, *SUN, OVATE, FASCIIATED (FAS), and LOCULE NUMBER (LC)* have been identified as genes that affect tomato fruit size and shape (Liu et al. 2002; Cong et al. 2008; Xiao et al. 2008; Munos et al. 2011). The *FAS* and *LC* sites determine the number of ventricles and regulate fruit shape to a certain extent (Rodríguez et al. 2011); whereas the *SUN* gene encodes a member of the calmodulin-binding IQ67-domain (IQD) family proteins that induces fruit elongation. Moreover, *OVATE* gene encodes a negative growth regulator...
that controls fruit length by blocking transcription (van der Knaap et al. 2014). In addition, QTL $f_{s8.1}$, which causes tomatoes to take on lumpy and slightly elongated shapes, as well as $s_{ov1}$ and $s_{ov2}$, inhibitors of the $OVATE$ gene, can regulate fruit shape (Clevenger 2012; Rodriguez et al. 2013). Similarly, in cucumbers, numerous QTLs associated with fruit shape have been mapped (Zhu et al. 2016; Yang et al. 2018; Gao et al. 2020). For instance, Weng et al. (2015) used three QTL models to map nine fruit-size-related traits and identified three QTLs for fruit length and fruit diameter; four QTLs for ovary length, ovary diameter, and mature fruit length; and five QTLs for mature fruit diameter. Similarly, Pan et al. (2017) detected two QTLs, i.e., $f_{s1.2}$ and $f_{s2.1}$, using populations of WI7238 (long fruit) × WI7239 (round fruit). Further analysis showed that the homologous gene, Cs$SUN$, of the tomato fruit shape gene, $SUN$, was a candidate gene for $f_{s1.2}$. Meanwhile, the $SF1$ (short fruit 1) gene encodes a cucurbit-specific E3 ligase, and a mutation in this gene results in increased ubiquitination and degradation, leading to inhibition of cell division and short fruit formation (Xin et al. 2019). Moreover, a functional allele of CsFUL1 regulates fruit length by inhibiting CsSUP and auxin transport (Zhao et al. 2019). Indeed, QTL mapping analysis for fruit shape in watermelons showed two major, stable QTLs, located primarily on chromosomes 2 and 3 (Sandlin et al. 2012; Cheng et al. 2016; Liu et al. 2016). Additionally, Kim et al. (2015) identified a major QTL, $f_{si3.1}$, in a watermelon $F_2$ population. Subsequently, Dou et al. (2018) located $f_{si3.1}$ in the 26.817–26.863 Mb region on chromosome 3, determining that a 159-bp deletion in $Cl_{a011257}$ was responsible for fruit elongation; a total of 105 watermelon lines were tested, and all lines with long and thin fruits harbored this deletion. However, the $Cl_{a011257}$ allele reportedly also causes other variations in fruit shape (Legendre et al. 2020).

In wax gourds, using a segregating population derived from crossing a landrace accession B214 and cultivated accession B227, two QTLs for fruit length ($f_{j4.1}$ and $f_{j10.1}$), and two QTLs for fruit diameter ($f_{jd3.1}$ and $f_{jd11.1}$) were mapped. The QTL $f_{jd3.1}$ located on chromosome 3 at 58.3 cm was a major QTL with an LOD value of 7.73 and explained 16.7% of the observed phenotypic variation (Liu et al. 2018). By resequencing 146 materials, Xie et al. (2019) constructed a genomic variation map of wax gourds that was subsequently combined with population genetics and linkage maps. Based on these data, they suggested that the homologous genes $Bhi10G00138$ and $Bhi10G00196$ are related to fruit size in other species.

The development of high-throughput sequencing technologies, as well as the availability of the wax gourd reference genome, genomic variation maps (Xie et al. 2019), bulked segregant analysis with whole genome resequencing (BSA-Sep) (Ramirez-Gonzalez et al. 2015; Zon et al. 2016), and next-generation sequencing (NGS) technologies have significantly accelerated the identification of candidate genes that control important agronomic traits. Meanwhile, combined NGS and BSA provides an effective means to accelerate the fine mapping and cloning of target genes (Tomita and Tanisaka 2019; Navarro-Escalante et al. 2020; Wang and Karamyshev 2020), as it has a low dependence on molecular markers, short genotyping time, and greatly shortens the time required for QTL mapping, thus improving the overall efficiency of research.

The genetic information controlling fruit shape in wax gourds remains unknown, and candidate genes controlling fruit shape have not been reported. Therefore, in the current study, we completed the high-quality genome map of wax gourd GX-19 (long cylindrical fruit) using the Illumina NovaSeq 6000 sequencing platform (San Diego, CA, USA). The genome size and contig N50 were 987 Mb and 18.87 Mb, respectively (unpublished data). This produced a high-quality genome map with gene predictions and functional annotations of the entire genome. Based on this, we sought to identify a candidate gene for fruit shape in wax gourds using BSA in $F_2$ populations derived from the crossing of GX-71 (long cylindrical fruit) and MY-1 (round fruit). To the best of our knowledge, this is the first report on candidate genes for fruit shape regulation in wax gourds, which will provide a theoretical basis for fruit shape breeding in this species.

**Materials and methods**

**Plant materials and phenotypic statistics**

Two inbred gourd lines (GX-71 and MY-1) were cultivated and used to obtain a fourth-generation population for genetic analysis of fruit shape. GX-71 fruit are long cylindrical, with green skin and flesh, while MY-1 are round, with white skin and flesh. $F_1$ plants were obtained by crossing GX-71 ($P_1$) and MY-1 ($P_2$), and the $F_2$ population was obtained by self-crossing $F_1$ plants. The parents and $F_1$ individuals were planted in the spring of 2020. To identify candidate genes for fruit shape, 276 and 6461 $F_2$ individuals were planted in the spring and autumn of 2020, respectively. The within-row spacing was 0.5 m, with a distance of 1.2 m between rows. To ensure full fruit development, only one well-developed fruit was retained for 8–16 nodes of the plant. The longitudinal and transverse diameters of the fruit were recorded 15 days post pollination (DPP). From ovary formation to fruit ripening, the longitudinal and transverse diameters of fruits of the two parents were measured. Three replicates were used for each measurement, and the longitudinal diameter to transverse diameter ratio was calculated to determine
the fruit shape index (FSI). All plant materials were grown in a field in Guangxi University (Guangxi, China).

**Cytological analysis**

To compare the cytological characteristics of fruits from the two parents, the mesocarp of the middle portion of the fruit at days 0 and 15 post pollination was cut into thin slices and immediately placed in formaldehyde alcohol acetic acid fixative solution (50% ethanol: 40% formaldehyde: glacial acetic acid = 16:1:1). The volume of the fixed solution was approximately four times that of the thin gourd slices. The bottle mouth was sealed with a sealing film and fixed at 25 °C for more than 48 h. The fixed pulp was dehydrated using an ethanol concentration gradient (70, 80, 90, 95, and 100%) and the thin slices were embedded in paraffin using xylene. The vertical and horizontal sections of these slices were sliced using a paraffin sectioning machine and pasted on a slide. The thin slices were stained using the safranin solid green double staining method and observed with a Zeiss microscope (Zeiss, Germany).

The cell sizes and cell numbers of parental lines were estimated using images obtained with the Image J software (https://imagej.nih.gov/ij/).

**DNA extraction**

The genomic DNA of the parent plants, and of the F1 and F2 populations, were extracted from leaf material using the cetyltrimethylammonium bromide (CTAB) method (Porebski et al. 1997). The concentration and purity of the extracted DNA were measured using a k5800 ultra-micro spectrophotometer (Kaiao, Beijing, China), and the DNA was evaluated by 1.2% agarose gel electrophoresis.

**Resequecing**

After the quality of DNA samples was confirmed, the DNA was randomly interrupted using ultrasonic fragmentation. The ends of DNA fragments were repaired, and an A added at the 3' end and sequencing connector added. The fragments were then purified and amplified by PCR to construct a sequencing library. After passing the quality inspection, the library was sequenced using the Illumina sequencing platform. The original image data files obtained through high-throughput sequencing were transformed into sequenced reads by base calling analysis. The sequenced reads contained low-quality reads with connectors. Raw reads were filtered to obtain clean reads to ensure the quality of subsequent information analyses. The main data filtering steps were as follows: (1) removed the adapter sequence; (2) if the proportion of N on a read was more than 10%, the paired reads were filtered out; and (3) low-quality reads (the number of bases with \( Q \leq 10 \) accounting for more than 50% of the whole read) were removed. The sequencing reads were compared with the reference genome using BWA software and re-located to the reference genome for subsequent mutation analysis (Li and Durbin 2009).

**BSA-sep mapping approach**

To examine phenotype, 60 extreme plants (30 plants with long cylindrical fruits and 30 plants with round fruits) were selected from the 276 F2 plants. After single plant resequencing, two mixed pools, one long cylindrical pool and one round pool, were constructed. The two mixed pools and two parent pools were used for association analyses, and GX-19 was used as the reference genome.

SNP detection was performed using the GATK software kit (McKenna et al., 2010). Based on the results of the positioning of clean reads in the reference genome, GATK was used for local realization and other pretreatments to ensure the accuracy of SNP detection, as well as for SNP detection to determine the SNP site. Before association analyses were undertaken, SNPs were first filtered. The filtering criteria were as follows: SNPs with multiple genotypes were filtered; SNPs with read support < 4 were filtered; SNPs with consistent genotypes among pools and those with recessive pool genes not received from recessive parents were filtered. The Euclidean distance (ED) algorithm was then applied to analyze the association between SNPs with different genotypes in the two pools (Hill et al. 2013; Rym et al. 2013). SNP sites with different genotypes between the two pools were used to determine the depth of each base in each pool and to calculate the ED value of each site. To eliminate background noise, the original ED value was processed by power, and the second power of the original ED was taken as the correlation value. The DISTANCE method was used to fit the ED value.

**Fine mapping**

Based on BSA-Sep data, as well as the distribution and density of the physical SNP locations, KASP markers were developed at 1–3 Mb intervals in the candidate region. Preparation of the mixture for analysis and PCR amplification was performed according to the manufacturer’s instructions (LGC Genomics, Shanghai, China). The PCR reaction system occupied a volume of 3 μL, including 1.0 μL of DNA (8–15 ng μL\(^{-1}\)), 1.5 μL of 2× master mix and 0.5 μL of primer mix. PCR amplification was performed using landing PCR. The reaction conditions were as follows: heat treatment at 95 °C for 15 min; denaturation at 95 °C for 20 s, annealing and extension between 65 and 55 °C for 60 s, ten landing cycles (each cycle reduced by 1.0 °C); denaturation...
Cloning and sequencing analysis of candidate genes

Genes and whole coding sequences (CDSs) of candidate genes were cloned. The primers (Table S1) were designed based on the genome. Total RNA was isolated using a plant RNA purification kit (Tiangen, Beijing, China), according to the manufacturer’s instructions, and treated with RNase-free DNase solution to remove residual genomic DNA. The first strand of complementary DNA (cDNA) was synthesized using the enzyme reverse transcriptase, and primers for homologous cloning were designed using the CDS database of the research group. The 2 × 8 FastHiFi PCR Master Mix (Aidlab, Beijing, China) was used for PCR amplification. The PCR product was detected by 1.2% agarose gel electrophoresis, and the target strip was recovered and purified by gel cutting. A zero-background pTOPO-Blunt cloning kit (CV16) from Aidlab was then used to construct an expression vector, according to the manufacturer’s instructions: 1 µL of pTOPO-Blunt vector, 1 µL of 10 × Enhancer, 6 µL of sterile water, and 2 µL of PCR gel products were mixed, and ligated at 37 °C for 10 min. The vector was transformed into Trans5α chemically compatible cells according to the manufacturer’s instructions (TransGen Biotech, Beijing, China), and the correct PCR colony clones were selected for sequencing confirmation. All fragments were sequenced by Shanghai Shengong Biotechnology Co., Ltd. DNAMAN software was used for multiple sequence alignments.

Gene expression analysis

To analyze gene expression, RNA was extracted from the ovary and flesh at different developmental stages, as well as from peels, roots, stems, leaves, and female and male flowers at the flowering stage. cDNA was synthesized using reverse transcriptase RT Master Mix (RR036A) following the manufacturer’s instructions (TaKaRa, Beijing, China). The gene-specific primers (Table S1) of the candidate gene and reference gene, ACTIN (Bch10G018990), were designed using NCBI online Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). QuantStudio 6 Flex (Thermo Fisher, MA, USA) was used to evaluate the expression levels of target genes by qRT-PCR. A SYBR Green real-time PCR mixture was used for all reactions. A 20 µL RT–PCR reaction mixture containing 2 µL of cDNA, 1 µL of forward primers (10 µM), 1 µL of reverse primers (10 µM), 10 µL of 2 × SYBR Green real-time PCR mixture, and 6 µL of nuclease-free water was preheated at 95 °C for 30 s, followed by heating for 5 s at 95 °C and 34 s at 60 °C for 40 cycles. High-resolution melting was performed at 95 °C for 15 s, 60 °C for 1 min, and 95 °C for 15 s. A minimum of three replicates were tested for each sample. The original data from RT–PCR were obtained using QuantStudio 6 Flex software (Thermo Fisher, MA, USA), and the relative expression was determined by the 2^−ΔΔCT method with actin as the internal control.

Results

Phenotypic characterization and inheritance of FSI

Considering that no significant increase was observed in the fruit shape index of GX-71 nor MY-1 at 15 DPP, FSI was measured during this stage (Fig. 1a). The final fruit shape indices of long cylindrical and spherical wax gourds were similar to their respective shape indices, recorded 4 days before flowering. Fruit shape can be predicted by ovary shape in wax gourds (Fig. 1b, c). Images of typical fruits for the two parental lines, as well as F1 and F2 populations, are shown in Fig. 1b and d. The mean FSI for the long cylindrical parent, GX-71, and the round parent, MY-1, were 4.56 and 1.06, respectively. The GX-71 × MY-1 F1 hybrid had an average FSI of 2.35 and appeared to be cylindrical. The fruits of the F2 population had three main shapes: long cylindrical, cylindrical, and round. We defined FSI for round, cylindrical, and long cylindrical as being < 1.3, 1.3–2.7, and > 2.7, respectively. In the spring of 2020, the F2 population was divided into 70 long cylindrical, 143 cylindrical, and 63 round spherical plants. These conformed to the Mendelian law of 1:2:1. In the autumn of 2020, 1469 F2 plants were divided into 373 long cylindrical, 733 cylindrical, and 363 spherical plants, which also conformed to the Mendelian law of 1:2:1. Collectively, these results indicate that fruit shape in this population is controlled by a single gene with incomplete dominance.

Analysis of paraffin-fixed longitudinal sections from parent fruits at different developmental stages revealed that cell size in GX-71 was significantly smaller than in MY-1, and the number of cells in GX-71 was significantly higher than that in MY-1 at the ovary development stage. There was no significant difference in cell size between GX-71 and MY-1 at day 15 post pollination (Fig. 1e–g).
After whole genome resequencing, 30.22 GB and 31.52 GB of clean data were obtained from GX-71 and MY-1, respectively, while an average of 5.49 GB was obtained for the F2 plants. The average Q30 and GC values were 92.70 and 35.57%, respectively. As compared to the reference genome, the mapping rate of the parents and F2 offspring ranged from 96.81% to 98.22%. The average genome coverage of the parents was > 27 ×, and the mapping rate was > 98% (at least 1 ×). The average genome coverage of F2 was 4.95 ×, and the mapping rate was > 95.21% (at least 1 ×; Table S2). Based on the mapping results of the clean reads in the reference genome, GATK was used to detect and filter SNPs, and 955,238 high-quality credible SNPs were obtained (Table S3). The ED algorithm was used to analyze the association of different SNPs, and the median + 3 SD of the fitting value of all sites was taken as the correlation threshold for analysis, and a value of 0.41 was obtained. According to the association threshold, a region on chromosome 2 was found to be associated with the fruit shape gene, which was located in the 50.83–68.01 Mb region, with a total length of 17.18 Mb (Fig. 2a).

To further narrow the candidate region, KASP primers were designed based on the single-nucleotide polymorphisms of the target parents (Table S4). 1469 F2
Fig. 2 Genetic mapping of the fruit shape gene in wax gourd. 

(a) Distribution of Euclidean distance (ED) association values on chromosomes (chr). The abscissa is the chromosome name, the colored point represents the ED value of each SNP locus, the black line represents the fitted ED value, and the red dotted line represents the significance association threshold. The higher the ED value, the better the association effect of this point.

(b) Genetic mapping of the fruit shape gene by KASP. Using 1469 F$_2$ individuals resulting from the GX-71 × MY-1 cross, the location of the fruit shape gene was narrowed down to a 1.98 Mb region between the markers fs57 and fs59 on chromosome 2.

(c) Genotyping of recombinant plants from an additional 4992 F$_2$ individuals resulting from the GX-71 × MY-1 cross. The FSI of the two parents and their F$_1$ offspring, and part of recombinant, are shown on the right. The location of the fruit shape gene was narrowed down to a 19.6 Kb region.
individuals were analyzed using 13 KASP markers. The candidate loci were located between the molecular markers, fs57 and fs59, with a physical distance of 1.98 Mb (Fig. 2b). To narrow down the potential mapping range of the candidate genes, the markers, fs56, fs57, fs59, and fs60, were used as flanking markers. After genotyping 4992 F₂ plants, 34 recombinant plants were identified; thus, new KASP markers were developed between fs57 and fs59 to genotype these plants. Through genotype–phenotype joint analysis, the fruit shape regulatory gene was found to be located between markers fs58.13 and fs58.15, with a physical distance of 19.6 Kb (Fig. 2c). Only one gene (Bch02G016830) was located within this interval, which we designated BFS (GenBank accession no. MZ032005.1). These results indicate that BFS may be a candidate gene for fruit shape determination in wax gourds.

Identification of candidate gene

To determine the BFS gene sequence, we cloned it along with CDSs from the two parental lines. Sequence alignment analysis showed that the full length of the BFS gene was 1607 bp, including 394 bp, 201 bp, and 671 bp exons, as well as 70 and 93 bp introns. Its sequence in GX-71 was the same as in the GX-19 reference genome; whereas its sequence in MY-1 harbored two non-synonymous mutations in the second and third exons. More specifically, a change from cytosine to guanine occurred resulting in the transformation of an alanine (A) to a glycine (G) at the 134th amino acid position (Fig. 3a, b).

To verify these differences, the BFS gene from four other cultivars, including one long cylindrical (HT-7), two cylindrical (GF-7, KX-8), and one round (YSB-1) fruit, were subjected to cloning and sequence analysis. Through sequence comparison the gene sequence of HT-7 was found to be the same as GX-71, while that of YSB-1 was the same as MY-1. Meanwhile, the gene sequence of GF-7 and KX-8 represented a new allele harboring a non-synonymous mutation in the third exon, i.e., a change from thymine to guanine, resulting in the transformation of a tyrosine (Y) to an aspartate (D) at the 238th amino acid position (Fig. 3a, b).

According to SMART (https://SMART.embl-heidelberg.de/) analysis, the protein structure included IQD domains i.e., two IQ motifs (115th–137th amino acids, 138th–159th amino acids) and DUF4005 (293rd–407th amino acids). A non-synonymous mutation resulted in amino acid substitution in the IQD domain (Fig. 3c).

Gene expression analysis

BFS expression in the ovary and flesh of GX-71 and MY-1 was determined using qRT-PCR. During the ovary stage, BFS expression in MY-1 was higher than GX-71, especially in the early stage of ovarian formation. Although BFS expression in MY-1 and GX-71 flesh was also high at 0 and 3 DPP, BFS expression in MY-1 was higher than GX-71 at 6, 9 and 15 DPP (Fig. 4a). BFS expression was also evaluated in different tissues during the flowering stage and was found to be expressed in the roots, stem, young leaf, male and female flowers, flesh, and peel, with the highest expression observed in the flesh (Fig. 4b).

dCAPS marker development for fruit shape determination

Based on the two non-synonymous mutations observed in BFS, two dCAPS markers (Table S1) were developed, and primers were designed using dCAPS Finder 2.0 (http://helix.wustl.edu/dcaps/dcaps.html) and Primer Premier 5.0. Sixty homozygous wax gourd germplasms were screened to test for genotype and phenotype consistency. Of these 60 germplasms, 19, 25, and 16 were long cylindrical, cylindrical, and round, respectively. As predicted, the genotype was consistent with the phenotype (Fig. 5; Table S5).

Phylogenic analysis

To understand the relationship between the protein sequences of BFS and other homologous sequences, we analyzed the BFS protein sequence using NCBI BLAST (NCBI, Bethesda, MD, USA) and subsequently generated a phylogenetic tree with 1000 repeats using the neighbor-joining method. Results showed that BFS has a close phylogenetic relationship with plants of the Cucurbitaceae family, including Cucumis sativus, Cucumis melo, Cucumis maxima, and Cucurbita moschata, indicating that the gene is evolutionarily conserved in the Cucurbitaceae family (Fig. 6). Sequence alignment showed that BFS shared 81.32% sequence homology with AT3G16490, which is found in Arabidopsis thaliana; AT3G16490 encodes IQD26 proteins.

Discussion

Fruit shape gene and allelic variation in wax gourd

The wax gourd germplasm has three primary shapes: round, cylindrical, and long cylindrical, while the other fruiting types are natural or artificial hybrids of these three types. In wax gourds, based on the high-density genetic map of 140 F₂
individuals, Liu et al. (2018) detected nine QTLs responsible for determining fruit-related traits (fruit weight, fruit length, fruit diameter, and flesh thickness) on chromosomes 3, 4, 5, 6, 9, 10, and 11. In the present study, a candidate gene for wax gourd fruit shape was identified on chromosome 2, indicating that the gene for fruit shape is distinct from that for fruit weight, fruit length, fruit diameter, and flesh thickness. Based on the SNPs detected in these populations, KASP molecular markers were developed to expand the population and screen for recombinant plants. The fruit shape gene was determined to be limited to the 58.134–58.154-Mb region on chromosome 2 (Fig. 2). Annotation analyses confirmed the presence of a BFS gene in this region. BFS is a member of the SUN family; Cla011257, CmAUN-16, and FSl.2 are also families that control fruit shape in watermelons, melons, and cucumbers, respectively (Perpiñá et al. 2016; Pan et al. 2017; Dou et al. 2018). It is, therefore, suggested that BFS is a candidate gene for the determination of fruit shape in wax gourds. Sequence alignment between long cylindrical and round fruit plants revealed the presence of

![Mutation sites and amino acid changes in the BFS allele.](image)

**a** Comparison of BFS gene structure and base sequence in six cultivars. GX-71 and HT-7 are long cylindrical fruit, MY-1 and YSB-1 are round fruit, GF-7 and KX-8 are cylindrical fruit. White boxes, black boxes, and solid line represent 5′ and 3′ UTR, exons, and introns, respectively. **b** Comparison of amino acid sequence of BFS protein in six varieties; the red part indicates amino acid differences. **c** BFS: IQD domain, this shows that one of the two amino acid substitutions occurs in the domain.
two non-synonymous mutations, resulting in two amino acid changes in the protein sequence. Another new allele was also identified in the cylindrical wax gourd (Fig. 3). The germplasms were verified by dCAPS markers, and the genotype and phenotype were consistent (Fig. 5; Table S5). These fine mapping results have laid a solid foundation for revealing the genes that regulate fruit shape in wax gourds.

**Changes in cell size and number are responsible for changes in fruit shape**

Under the same environmental conditions, cell proliferation and expansion directly influence fruit size (Maeda et al. 2014). Cell proliferation is the basis for plant growth, development, and reproduction and generally occurs during the early stages of fruit growth and ultimately determines the number of cells in the fruit, thus affecting fruit size (Ho 1996; Cong et al. 2002). The duration of cell proliferation varies between plant taxa. For example, cell division in cucumbers and melons usually lasts only 5–8 days (Li et al. 2003; Fu et al. 2010; Colle et al. 2017), whereas in tomatoes, it lasts approximately 2 weeks (Ding et al. 2013). During the middle stage of fruit development, cell proliferation gradually stops, and cell expansion plays a leading role in fruit development. Cell expansion comprises cell elongation caused by changes in cell wall structure. It is generally believed that cell number and size both determine final fruit size.

In tomatoes, changes in cell number in different growth axes lead to a slender fruit shape (Clevenger et al. 2015; Wang et al. 2019). In cucumbers, inbred lines with longer fruits show higher cell numbers in the longitudinal direction than those with shorter fruits (Liu et al. 2020). In melons, cucumbers, and watermelons, the shape of mature fruit is related to the height of the ovary, and fruit shape can be determined before pollination (Wei et al. 2016; Pan et al. 2017; Dou et al. 2018). We also found that fruit shape can be predicted by ovary shape at different growth stages in wax gourds (Fig. 1a, c). In this study, the cell size and number of long cylindrical and round spherical wax gourds were compared using paraffin sections. Long cylindrical wax gourds had significantly higher cell numbers and lower cell areas.
than the round spherical wax gourds during the ovary formation stage (Fig. 1e-g). This difference in cell area decreased gradually after ripening, indicating that long cylindrical wax gourds grow faster than round spherical wax gourds at the ovary formation stage. Moreover, after entering the fruit growth stage, the cell area gradually increased, and the fruit continued to grow, indicating that the cell number at ovary stage and the cell area during the fruit stage regulates wax gourd fruit shape.

Duplication of the tomato SUN gene leads to an increase in its expression, promotes cell division, and leads to fruit elongation (Xiao et al., 2009; Wu et al., 2011). The Cla011257 gene, which is highly expressed during ovary formation and promotes ovary and fruit elongation, is missing 169 bp in watermelon (Dou et al. 2018). Deletion of the first exon of FS1.2 in cucumbers results in a reduction in its expression and the formation of round fruits (Pan et al. 2017). In this study, the expression pattern of BFS was found to be higher in round wax gourds at the ovary stage than in long cylindrical wax gourds (Fig. 4a), thus, slowing down cell division and promoting cell expansion in round wax gourds. Additionally, the number of cells in round gourds was lower than that in long cylindrical gourds, and the cell area was relatively larger in spherical wax gourds. Our results show that variations in BFS may slow cell division at the ovary formation stage. Collectively, these results provide a preliminary understanding of the molecular mechanisms underlying the regulatory effects of BFS on wax gourd fruit size; however, further analysis of cylindrical fruit cells and allele expression might shed more light on any additional effects, if any, of this mutation on fruit shape.

Association of the functional role of BFS with the IQD protein family

The amino acid series encoded by BFS contains an IQD domain (Fig. 3c). Phylogenetic analysis showed that BFS has high homology with AT3G16490 in the Arabidopsis genome, which belongs to the IQD protein family (Fig. 5). This protein family is defined by a central conserved IQ67 domain, which spans 67 amino acids and contains up to three IQ motifs that are necessary for binding to CaM Ca$^{2+}$ sensors (Bürstenbinder et al. 2013; Cai et al. 2016). Calcium (Ca$^{2+}$) is a second messenger in all eukaryotes and plays an important role in the complex process of precise cell growth and shape regulation (Cárdenas and Luis 2009; Steinhorst and Kudla 2013). Calmodulin/calmodulin-like (CaM/CML) protein polypeptides are the main Ca$^{2+}$ sensors that relay and
control the biochemical activities of different regulatory targets through complex Ca\(^{2+}\)-dependent and Ca\(^{2+}\)-independent interactions (Dodd 2010; Kudla et al. 2010). Ca\(^{2+}\) and CaM/CML affect cell division (Hepler et al. 2005). Meanwhile, the microtubule (MT) cytoskeleton forms a highly dynamic network and plays a central role in coordinating cell growth. In Arabidopsis species, IQD1 targets MT and interacts with the kinesin light chain associated protein-1 (KLCR1) and CaM/CML proteins, indicating that IQD1 participates in the regulation of Ca\(^{2+}\) CaM signaling and MT-related processes (Bürstenbinder et al. 2013, 2017). The IQD protein is a unique calmodulin protein that regulates the morphology and quantity of cells in Arabidopsis and rice plants (Duan et al. 2017; Sugiyama et al. 2017). SUN, found in tomatoes, is a member of the IQD family and a calmodulin-binding protein that controls the slendorness of the fruit shape by increasing the number of cells from the proximal to the distal end of the fruit and by decreasing the number of cells on the inner and outer sides of the fruit (Xiao et al. 2008; Wu et al. 2011). Thus, based on the results of this study, we hypothesize that the BFS gene may affect Ca\(^{2+}\) CaM signaling and modify cell division and the cytoskeleton, thus modifying fruit shape.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-021-03942-8.

Acknowledgements This study was supported by the National Nature Science Foundation of China (31960593). This study was also supported by Guangxi Major Program of Science and Technology (Guike AA17204026).

Author Contribution Statement Z.G.L., W.J.Y., and P.W conceived and designed the study. Z.K.C. and Y.C.X conducted all of the experiments. L.L.M., J.Q.G., J.Y.C., L.W.S., W.T.W., and Y.C participated in some experiments. Z.K.C. and Y.C.W wrote the manuscript. Z.G.L., W.J.Y., and P.W revised the manuscript. Z.K.C., Z.G.L., and Y.C.X contributed equally. All the authors reviewed the manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Bürstenbinder K, Savchenko T, Müller J, Adamson AW, Stamm G, Kwong R, Zipp BJ, Dinesh DC, Abel S (2013) Arabidopsis calmodulin-binding protein IQ67-domain 1 localizes to microtubules and interacts with kinesin light chain-related protein-1. J Biol Chem 288:1871–1882

Bürstenbinder K, Möller B, Plötner R, Stamm G, Hause G, Mitra D, Abel S (2017) The IQD Family of Calmodulin-Binding Proteins Links Calcium Signaling to Microtubules, Membrane Subdomains, and the Nucleus. Plant Physiol 173:1692–1708

Cai R, Zhang C, Yang Z, Zhu K, Wang Y, Jiang H, Yan X, Cheng B (2016) Genome-wide analysis of the IQD gene family in maize. Mol Genet Genomics 291:543–558

Cárdenas L (2009) New findings in the mechanisms regulating polar growth in root hair cells. Plant Signal Behav 4:4–8

Cheng Y, Luan F, Wang X, Gao P, Zhu Z, Liu S, Baloch AM, Zhang Y (2016) Construction of a genetic linkage map of watermelon (Citrus lanatus) using CAPS and SSR markers and QTL analysis for fruit quality traits. Sci Hortic 202:25–31

Clevenger JP, Van Houten J, Blackwood M, Rodríguez GR, Jikumaru Y, Kamiya Y, Kusano M, Saito K, Vasa S, van der Knaap E (2015) Network analyses reveal shifts in transcript profiles and metabolites that accompany the expression of SUN and an elongated tomato fruit. Plant Physiol 168:1164–1178

Clevenger J (2012) Metabolic and genomic analysis of elongated fruit shape in tomato (Solanum lycopersicum). Dissertation, The Ohio State University

Colle M, Weng Y, Kang Y, Ophir R, Grumet R (2017) Variation in cucumber (Cucumis sativus L.) fruit size and shape results from multiple components acting pre-anthesis and post-pollination. Planta 246:641–658

Cong B, Liu J, Tanksley SD (2002) Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. Proc Natl Acad Sci USA 99:13606–13611

Cong B, Ba Rtero LS, Tanksley SD (2008) Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nat Genet 40:800–804

Ding J, Chen B, Xia X, Mao W, Shi K, Zhou Y, Yu J (2013) Cytokinin-induced parthenocarpic fruit development in tomato is partly dependent on enhanced gibberellin and auxin biosynthesis. PLoS One 8:e70080

Dodd AN, Kudla J, Sanders D (2010) The language of calcium signaling. Annu Rev Plant Biol 61:593–620

Dou J, Zhao S, Lu X, He N, Zhang L, Ali A, Kuang H, Liu W (2018) Genetic mapping reveals a candidate gene (CIFS1) for fruit shape in watermelon (Citrus lanatus L.). Theor Appl Genet 131:947–958

Duan P, Xu J, Zeng D, Zhang B, Geng M, Zhang G, Huang K, Huang L, Xu R, Ge S, Qian Q, Li Y (2017) Natural variation in the promoter of GSE5 contributes to grain size diversity in rice. Mol Plant 10:685–694

Fu FQ, Mao WH, Shi K, Zhou YH, Yu JQ (2010) Spatio-temporal changes in cell division, endoreduplication and expression of cell cycle-related genes in pollinated and plant growth substances-treated ovaries of cucumber. Plant Biol (Stuttgart) 12:98–107

Gao Z, Zhang H, Cao C, Han J, Li H, Ren Z (2020) QTL mapping for cucumber fruit size and shape with populations from long and round fruited inbred lines. Horticultural Plant Journal 6:132–144

Gu M, Fan S, Liu G, Guo L, Ding X, Xu Y, Zhang Y, Ji G, Huang C (2013) Extract of wax gourd peel prevents high-fat diet-induced hyperlipidemia in C57BL/6 mice via the inhibition of the PPARγ pathway. Evid Based Complement Alternat Med 2013:1–11

Hepler PK (2005) Calcium: a central regulator of plant growth and development. Plant Cell 17:2142–2155

Hill JT, Demarest BL, Bisgrove BW, Gorski B, Su Y, Yost HJ (2013) MMAPPR: mutation mapping analysis pipeline for pooled RNA-seq. Genome Res 23:687–697

Ho LC (1996) The mechanism of assimilate partitioning and carbohydrate compartmentation in fruit in relation to the quality and yield of tomato. J Exp Bot 47:1239–1243

Kim K, Hwang J, Han D, Park M, Kim S, Choi D, Kim Y, Lee GP, Kim S, Park Y (2015) Major quantitative trait loci and putative candidate genes for powdery mildew resistance and fruit-related traits revealed by an intraspecific genetic map for watermelon (Citrus lanatus var. lanatus). PLoS One 10:e145665

One 8:e70080

3993

Theoretical and Applied Genetics (2021) 134:3983–3995
Kudla J, Batistovic O, Hashimoto K (2010) Calcium signals: the lead currency of plant information processing. Plant Cell 22:541–563

Legendre R, Kuzy J, McGregor C (2020) Markers for selection of three alleles of CISU25-26-27a (Csa011257) associated with fruit shape in watermelon. Mol Breeding 40:19

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

Li Y, Yu J, Ye Q, Zhu Z, Guo Z (2003) Expression of CycD3 is transiently increased by pollination and N-(2-chloro-4-pyridyl)-N’-phenyleurea in ovaries of Lagenaria cucurbita. J Exp Bot 54:1245–1251

Liu J, Van Eek J, Cong B. Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. P Natl Acad Sci USA 99:13302–13306

Liu S, Gao P, Zhu Q, Luan F, Davis AR, Wang X (2016) Development of cleaved amplified polymorphic sequence markers and a CAPS-based genetic linkage map in watermelon (Citrullus lanatus [Thunb.] Matsum. et Nakai) constructed using whole-genome re-sequencing data. Breed 66:244–259

Liu W, Jiang B, Peng Q, He X, Lin Y, Wang M, Liang Z, Xie D, Hu K (2018) Genetic analysis and QTL mapping of fruit-related traits in wax gourd (Benincasa hispida). Euphytica 214:136

Liu X, Pan Y, Liu C, Ding Y, Meng H (2020) Cucumber fruit size and shape variations explored from the aspects of morphology, histology, and endogenous hormones. Plants 9:772

Maeda S, Gunji S, Hanai K, Hirano T, Kizawa Y, Ohbayashi I, Abe T, Sawa S, Tsukaya H, Ferjani A (2014) The conflict between cell proliferation and expansion primarily affects stem organogenesis in Arabidopsis. Plant Cell Physiol 55:1994–2007

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

Munos S, Ranc N, Botton E, Berard A, Rolland S, Duffe P, Carretero Y, Le Paslier MC, Delande C, Bouzayen M (2011) Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiol 156:2244–2254

Navarro-Escalante L, Zhao C, Shukle R, Stuart J (2020) BSA-Seq discovery and functional analysis of candidate hessian fly (Mayetiola destructor) avirulence genes. Front Plant Sci 11:956

Pan Y, Liang X, Gao M, Liu H, Meng H, Weng Y, Cheng Z (2017) Round fruit shape in W17239 cucumber is controlled by two interacting quantitative trait loci with one putatively encoding a tomato SUN homolog. Theor Appl Genet 130:573–586

Perpiñá G, Esteras C, Gibon Y, Monforte AJ, Picó B (2016) A new genomic library of melon introgression lines in a cantaloupe background for dissecting desirable agronomical traits. BMC Plant Biol 16:154

Porebski S, Bailey LG, Baum BR (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Rep 15:8–15

Ramírez-Gonzalez RH, Segovia V, Bird N, Fenhwick P, Holdgate S, Berry S, Jack P, Caccamo M, Uauy C (2015) RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. Plant Biotechnol J 13:613–624

Rodríguez GR, Munño S, Anderson C, Sim S, Michel A, Causse M, Gardener BMM, Francis D, van der Knaap E (2011) Distribution of SUN, OVA TE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. Plant Physiol 156:275–285

Rodríguez G, Kim HJ, Van D (2013) Mapping of two suppressors of OVA TE (sov) loci in tomato. Heredity 111:256–264

Rym F, Hiroki T, Muluneh T, Akira S, Satoshi N, Hiroki Y, Shailendra S, Shiveta S, Hiroyuki K, Hideo M (2013) MutMap+: Genetic mapping and mutant identification without crossing in rice. PLoS One 8:e68529

Sandlin K, Prothro J, Heesacker A, Khalilian N, Okashah R, Xiang W, Bachlava E, Caldwell DG, Taylor CA, Seymour DK, White V, Chan E, Tolla G, White C, Safran D, Graham E, Knapp S, McGregor C (2012) Comparative mapping in watermelon (Citrullus lanatus [Thunb.] Matsum. et Nakai). Theor Appl Genet 125:1603–1618

Steinhorst L, Kudla J (2013) Calcium – a central regulator of pollen germination and tube growth. Biochim Biophys Acta 1833:1573–1581

Sugiyama Y, Wakuazaki M, Toyooka K, Fukuda H, Oda Y (2017) A novel plasma membrane-anchored protein regulates xylem cell-wall deposition through microtubule-dependent lateral inhibition of Rho GTase domains. Curr Biol 27:2522–2528

Tomita M, Tanisaka T (2019) The gametic non-lethal gene Gal on chromosome 5 is dispensable for the transmission of the co-induced semidwarfing gene d60 in rice. Biology 8:94

van der Knaap E, Chakrabarti M, Chu YH, Clevenger JP, Ilia-Berenguer E, Huang ZJ, Keyhaninejad N, Mu Q, Sun L, Wang YP, Wu S (2014) What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. Front Plant Sci 5:227

Wang D, Karamyshiev AL (2020) Next generation sequencing (NGS) application in multigene parameter gene expression analysis. Methods Molecular Biol 2102:17–34

Wang Y, Clevenger JP, Ilia-Berenguer E, Meulia T, van der Knaap E, Sun L (2019) A comparison of sun, ovate, fas1 and auxin application on tomato fruit shape and gene expression. Plant Cell Physiol 60:1067–1081

Wei Q, Fu W, Wang Y, Qin X, Wang J, Li J, Lou Q, Chen J (2016) Rapid identification of fruit length loci in cucumber (Cucumis sativus L.) using next-generation sequencing (NGS)-based QTL analysis. Sci Rep 6:27496

Weng Y, Colle M, Wang Y, Yang L, Grumet R (2015) QTL mapping in multiple populations and development stages reveals dynamic quantitative trait loci for fruit size in cucumbers of different market classes. Theor Appl Genet 128:1747–1763

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

Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science 319:1527–1530

Xiao H, Radovich C, Welty N, Hou J, Li D, Meulia T, van der Knaap E (2009) Integration of tomato reproductive developmental landmarks and expression profiles, and the effect of SUN on fruit shape. BMC Plant Biol 9:49

Xie D, Xu Y, Wang J, Liu W, Zhang Z (2019) The wax gourd genomes offer insights into the genetic diversity and ancestral cucurbit karyotype. Nat Commun 10:5158

Xin T, Zhang Z, Li S, Zhang S, Li Q, Zhang Z, Huang S, Yang X (2019) Genetic regulation of ethylene dosage for cucumber fruit elongation. Plant Cell 31:1063–1076

Yang L, Liu H, Zhao J, Pan Y, Cheng S, Lietzow CD, Wen C, Zhang X, Weng Y (2018) LITTLELEAF (LL) encodes a WD40 repeat domain-containing protein associated with organ size variation in cucumber. Plant J 95:834–847

Zhao J, Jiang L, Che G, Pan Y, Li Y, Hou Y, Zhao W, Zhong Y, Ding L, Yan S, Sun C, Liu R, Yan L, Wu T, Li X, Weng Y, Zhang X (2019) A functional allele of CsFULL1 regulates fruit length through repressing CsSUP and inhibiting auxin transport in cucumber. Plant Cell 31:1289–1307

Zhu WY, Huang L, Chen L, Yang JT, Wu JN, Qu ML, Yao DQ, Guo CL, Lian HL, He HL (2016) A high-density genetic linkage map for cucumber (Cucumis sativus L.): based on specific length
amplified fragment (SLAF) sequencing and QTL analysis of fruit traits in cucumber. Front Plant Sci 7:437
Zou C, Wang P, Xu Y (2016) Bulked sample analysis in genetics, genomics and crop improvement. Plant Biotechnol J 14:1941–1955

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.