



# Genetic mapping reveals a candidate gene (*CIFS1*) for fruit shape in watermelon (*Citrullus lanatus* L.)

Junling Dou<sup>1,2</sup> · Shengjie Zhao<sup>1</sup> · Xuqiang Lu<sup>1</sup> · Nan He<sup>1</sup> · Lei Zhang<sup>2</sup> · Aslam Ali<sup>1</sup> · Hanhui Kuang<sup>2</sup> · Wenge Liu<sup>1</sup>

Received: 17 October 2017 / Accepted: 4 January 2018 / Published online: 23 January 2018  
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

## Abstract

**Key message** A 159 bp deletion in *CIFS1* gene encoding IQD protein is responsible for fruit shape in watermelon.

**Abstract** Watermelon [*Citrullus lanatus* (Thunb.) Matsum. & Nakai] is known for its rich diversity in fruit size and shape. Fruit shape has been one of the major objectives of watermelon breeding. However, the candidate genes and the underlying genetic mechanism for such an important trait in watermelon are unknown. In this study, we identified a locus on chromosome 3 of watermelon genome controlling fruit shape. Segregation analysis in F<sub>2</sub> and BC<sub>1</sub> populations derived from a cross between two inbred lines “Duan125” (elongate fruit) and “Zhengzhouzigua” (spherical fruit) suggests that fruit shape of watermelon is controlled by a single locus and elongate fruit (OO) is incompletely dominant to spherical fruit (oo) with the heterozygote (Oo) being oval fruit. GWAS profiles among 315 accessions identified a major locus designated on watermelon chromosome 3, which was confirmed by BSA-seq mapping in the F<sub>2</sub> population. The candidate gene was mapped to a region 46 kb on chromosome 3. There were only four genes present in the corresponding region in the reference genome. Four candidate genes were sequenced in this region, revealing that the CDS of *Cla011257* had a 159 bp deletion which resulted in the omission of 53 amino acids in elongate watermelon. An indel marker was derived from the 159 bp deletion to test the F<sub>2</sub> population and 105 watermelon accessions. The results showed that *Cla011257* cosegregated with watermelon fruit shape. In addition, the *Cla011257* expression was the highest at ovary formation stage. The predicted protein of the *Cla011257* gene fitted in IQD protein family which was reported to have association with cell arrays and Ca<sup>2+</sup>-CaM signaling modules. Clear understanding of the genes facilitating the fruit shape along with marker association selection will be an effective way to develop new cultivars.

## Introduction

Fruit shape is among one of vital characteristic traits for horticulture industry that exhibits a broad range of phenotypic variation, thus emphasizing its importance in breeding

Communicated by Michael J. Havey.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00122-018-3050-5>) contains supplementary material, which is available to authorized users.

✉ Wenge Liu  
lwghm@163.com  
Junling Dou  
junlingdou@163.com

<sup>1</sup> Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou 450009, China

<sup>2</sup> College of Horticulture and Forestry Sciences, Huazhong Agricultural University, Wuhan 430070, China

program. Many QTLs and genes related to fruit shape have been detected or cloned in different crops. In cucumber, various QTLs of fruit shape have been genetically mapped (Bo et al. 2015; Jiang et al. 2015; Weng et al. 2015). Using RIL populations, ten QTLs were identified for fruit length (FL), fruit diameter (FD), mature fruit length (MFL), and mature fruit diameter (MFD) (Miao et al. 2011). During the early cucumber development, seven kinesin genes (*CsKFI* to *CsKF7*) were proved to play the key role in exponential cell production and enlargement in fruit (Yang et al. 2013). The *sf1* gene was also proposed for controlling cucumber fruit length by regulating gibberellins and cytokinin biosynthesis, signal transduction, and auxin signaling (Wang et al. 2017). Another study identified 8 QTLs for mature and immature cucumber fruit length. A major-effect QTL fl3.2 was detected, which explained a maximum phenotypic variation of 38.87% (Wei et al. 2016). The *CsSUM*, the homologous gene of *SUN* in tomato, was the candidate gene for cucumber spherical fruit (Pan et al. 2017). In tomato, two categories

of quantitative trait loci (QTLs) for fruit shape were suggested: One was locule number and flat shape controlled by *lc* (*locule number*) and *fas* (*fasciated*); the other was fruit shape controlled by *SUN*, *OVATE*, *LOCULE NUMBER*, and *FASCIATED* (Liu et al. 2002; Xiao et al. 2008; Munos et al. 2011; Rodriguez et al. 2011; van der Knaap et al. 2014). In addition, two suppressors of the ovate mutation (*Sov1* and *Sov2*) were proved to be regulator of fruit shape. The *self1* promotes fruit elongation by increasing cell layers in the ovary and the QTL *fs8.1* together with *SUN*, *OVATE* also controls tomato fruit shape (Paran and van der Knaap 2007; Rodriguez et al. 2013; Chusreeaom et al. 2014; Sun et al. 2015). In melon, the andromonoecious (*a*) was cloned and confirmed to control the fruit shape and carpel numbers (Boualem et al. 2008). Moreover, other QTLs associated with melon fruit shape (FS) were demonstrated by a number of QTL mapping (Paris et al. 2008; Fernandez-Silva et al. 2010; Tomason et al. 2013).

Watermelon [*Citrullus lanatus* (Thunb.) Matsum. & Nakai] belongs to Cucurbitaceae family and is one of the most important commercial crops worldwide having 11 chromosomes and a haploid genome of approximately 425 Mb ( $2n = 2x = 22$ ). The genome sequence and the genes annotation of watermelon were published in 2013 following cucumber and melon (Huang et al. 2009; Garcia-Mas et al. 2012; Guo et al. 2013), providing the opportunity to apply next-generation sequencing to fine map genes function. It was considered that watermelon fruit shape was controlled by an incompletely dominant gene, resulting in elongate (*OO*), oval (*Oo*), and spherical (*oo*) fruits (Weetman 1937; Guner and Wehner 2004). Tanaka et al. (1995) also determined that spherical (*O<sup>s</sup>*) and oval (*O<sup>+</sup>*) watermelon fruits are controlled by a single allele which shows incomplete dominance through the cross between a spherical fruit inbred line and an oval fruit inbred line. Furthermore, Poole and Grimball (1945) also found the same inheritance pattern in  $F_2$  populations of ‘Peerless’ × ‘Baby Delight’ and ‘Northern Sweet’ × ‘Dove’. Allele *Ob<sup>E</sup>* was proposed for elongate fruit, which was the dominant; allele *ob* for oblong fruit, which was the recessive, allele *Ob<sup>R</sup>* (was not the same as the *o* gene for round) for the round fruit (Lou and Wehner 2016).

Application of next-generation sequencing (NGS) and release of draft genome assemblies for two diploid watermelon inbred lines (97103 and Charleston Gray) (<http://cucurbitgenomics.org/>) greatly facilitated the construction of many genetic maps of watermelon recently (Sandlin et al. 2012; Zhang et al. 2012; Ren et al. 2014; Cheng et al. 2016). However, only a limited number of QTLs or genes have been reported for watermelon fruit shape. Using a consensus map of three mapping populations in watermelon, the major QTLs for fruit length (FL), fruit width (FWD), fruit shape index (FSI), and fruit weight (FWT) were found to be

stable across genetic backgrounds and environments. It was the first stable QTL for fruit size and shape in watermelon (Sandlin et al. 2012). Ren et al. (2014) constructed an integrated genetic map based on four mapping populations, and QTLs for fruit weight (FWT), fruit length (FL), fruit width (FWD), fruit shape index (FSI), and rind thickness (RTH) were detected and co-localized with QTL associated with Brix (BRX), suggesting the existence of pleiotropic effects on fruit maturity. The fruit shape index is the ratio of fruit length and width. There was positive correlation between fruit length, fruit width, and fruit shape index. Four QTLs associated with fruit shape index (FSI) were also detected through a genetic linkage map, and these QTLs explained the phenotypic variation of 31.9% (Sandlin et al. 2012; Cheng et al. 2016).

Information regarding inheritance pattern of fruit shape is elusive and candidate genes controlling the fruit shape in watermelon have not been reported yet. In this study, we investigated the inheritance of watermelon fruit shape gene in  $F_2$  population of ‘Duan125’ (elongate fruit) × ‘Zhengzhouzigua’ (spherical fruit). We identified a candidate gene on chromosome 3 associated with watermelon fruit shape (*CIFSI*) through genome-wide association studies (GWAS) of 315 watermelon accessions and bulked-segregant analysis (BSA) by genotyping a pair of bulked DNA samples from two sets of individuals with opposite extreme phenotypes (elongate individuals and spherical individuals). This is the first report about the candidate gene in controlling watermelon fruit shape. This study will open up the ways to shorten the breeding period and lead to selection of desired fruit shape in breeding program. This work will assist in selection of plants with desired fruit shape at the early stages of seedling or growth that will ultimately hasten the breeding process. This new allele provides a valuable tool in understanding the inheritance mechanisms of *CIFSI* in watermelon.

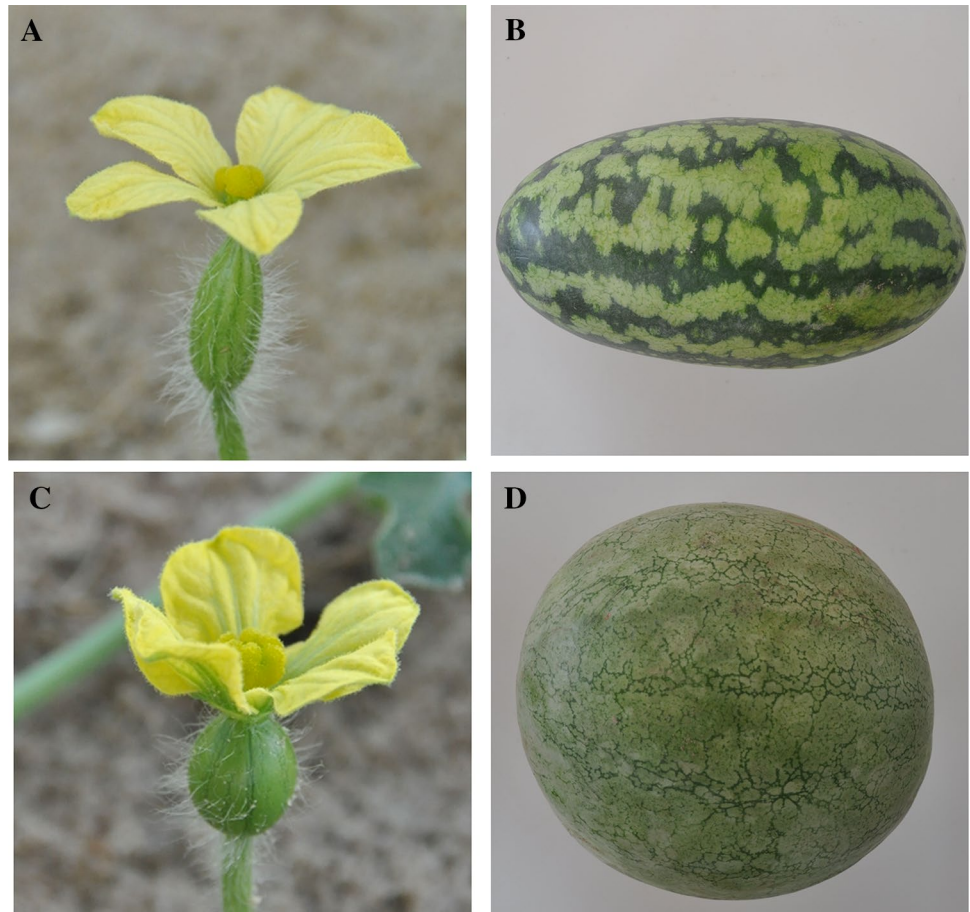
## Materials and methods

### Plant materials

Seeds of two inbred lines elongate watermelon ‘Duan125’ ( $P_1$ ) and spherical watermelon ‘Zhengzhouzigua’ ( $P_2$ ) were obtained from National Watermelon and Melon Germplasm Resource Library (Zhengzhou, China) to derive  $F_1$  population by crossing them. Later, selfing was performed in  $F_1$  to raise  $F_2$  (Fig. 1). Backcross population was obtained by hybridizing  $F_1$  with each parent to create  $BC_1P_1$  ( $F_1$  × Duan125) and  $BC_1P_2$  ( $F_1$  × Zhengzhouzigua).

For segregation analysis,  $F_2$  population was grown in two experiments conducted in 2016 winter, 2017 spring with 348 and 420  $F_2$  individuals, respectively. The  $BC_1P_1$  population

**Fig. 1** Materials of two different fruit shapes in watermelon. **a** Ovary of elongate watermelon ‘Duan125’. **b** Matured fruit of elongate watermelon ‘Duan125’. **c** Ovary of spherical watermelon ‘Zhengzhouzigua’. **d** Matured fruit of spherical watermelon ‘Zhengzhouzigua’



(120 plants) and the  $BC_1P_2$  (238 plants) population were investigated only in spring 2017 (Table S1). For each plant, the fruit length and width were recorded at mature stage. The fruit length and width were measured from ovary formation to fruit maturation to calculate shape index in two parents. At least ten fruits per genotype were evaluated. All the materials were grown in the experimental area of ‘Xinxiang’, Zhengzhou fruit research institute (Zhengzhou, China).

### Bulked-segregant analysis of DNA

DNAs were isolated using the CTAB method (Porebski et al. 1997) from fresh leaves of  $F_2$  plants in the 2016 winter for BSA-seq and CAPS marker analysis. Two DNA pools, elongate watermelon fruit pool (E-pool) and spherical watermelon fruit pool (S-pool), were constructed by mixing an equal amount of DNAs from 30 elongate watermelon plants and 30 spherical watermelon plants. Pair-end sequencing libraries (read length 100 bp) with insert size of around 500 bp were prepared for sequencing with an Illumina Genome Analyzer IIx machine. The short reads from E-pool and S-pool were aligned to the ‘97103’ reference genome (Guo et al. 2013) with the BWA software (Li and Durbin 2009). Alignment files were converted to SAM/BAM

files using SAM tools (Li et al. 2009), and then applied to the SNP-calling filter ‘Coval’ previously developed (Abe et al. 2012) to increase SNP-calling accuracy. SNP-index was calculated for all the SNP positions. We excluded SNP positions with SNP-index of  $< 0.6$  and read depth  $< 6$  from the two sequences, as these may represent spurious SNPs called due to genomic repeat sequence, sequencing, or alignment errors.

Two parameters, SNP-index and  $\Delta$  (SNP-index) (Abe et al. 2012; Takagi et al. 2013) were calculated to identify candidate regions for watermelon fruit shape QTL. An SNP-index is the proportion of reads harboring the SNP that are different from the reference sequence.  $\Delta$  (SNP-index) was obtained by subtraction of SNP-index of E-pool from that of S-pool. Thus, SNP-index = 0 if the entire short reads contain genomic fragments from mutation line; SNP-index = 1 if all the short reads are from reference genome ‘97103’. An average of SNP-index of SNPs located in a given genomic interval was calculated using a sliding window analysis with 1 Mb window size and 10 kb increment. The SNP-index graphs for E-pool and S-pool, as well as corresponding  $\Delta$  (SNP-index) graph were plotted.

To generate confidence intervals of the SNP-index value under the null hypothesis of no QTL, we carried out

computer simulation. We first made two pools of progeny with a given number of individuals by random sampling. From each pool, a given number of alleles corresponding to the read depth were sampled. We calculated SNP-index for each pool and derived the  $\Delta$  (SNP-index). This process was repeated 10,000 times for each read depth and confidence intervals were generated. These intervals were plotted for all the genomic regions that have variable read depths.

## GWAS

A total of 315 watermelon accessions which contained different fruit shape were re-sequenced and 2.3T data were obtained which had 85.42% average genome coverage and  $9.24 \times$  average sequence depth. The watermelon accessions were genotyped using 4,661,625 evenly spaced SNPs. The association between fruit shape and each SNP was tested using a unified mixed model (Yu et al. 2006; Zhang et al. 2010). This mixed linear model included principal components (Price et al. 2006) as a fixed effect to account for population structure, kinship matrix (Loiselle et al. 1995), as well as to explain the familial relatedness. Utilizing the Bayesian information criterion, a backward elimination procedure was implemented to determine the optimal number of principal components to include in the mixed model (Schwarz 1978). The false discovery rate was controlled at 5% using the Benjamini and Hochberg (1995) procedure. A likelihood ratio-based  $r^2$  statistic was used to assess the goodness of fit of each SNP (Sun et al. 2010). All analyses were conducted using the Genome Association and Prediction Integrated Tool software package (Lipka et al. 2012).

## Fine mapping through CAPS markers

Watermelon genome sequence was obtained from the watermelon database (<http://www.icugi.org>); the sequence was compared with the re-sequenced data to identify SNPs via a filter pipeline (Takagi et al. 2013). To minimize the genetic interval for fine mapping and to verify the accuracy of BSA-Seq, The 298 cleaved amplified polymorphic sequences (CAPS) markers were developed based on SNPs generated from BSA-seq (Table S2).

The PCR was carried out in a total volume of 10  $\mu$ L containing 5  $\mu$ L  $2 \times$  Power Taq PCR Master Mix (BioTeke, China), 0.5  $\mu$ L of 10  $\mu$ M per primer, 1  $\mu$ L of 200 ng of genomic DNA, and 3  $\mu$ L RNase-free water. All amplifications were performed on a EasyCycler (Analytik Jena, Germany) under the following conditions: 5 min at 95  $^{\circ}$ C; 28 cycles of 30 s at 94  $^{\circ}$ C, 30 s at 56  $^{\circ}$ C, 50 s at 72  $^{\circ}$ C, and a final extension step at 72  $^{\circ}$ C for 10 min. Amplified PCR products were digested using suitable restriction endonucleases according to the manufacturer's instructions at 37 or 65  $^{\circ}$ C for 4–10 h. The digests were resolved by

electrophoresis in 1.0% agarose gel or PAGE gel and visualized on a Versa Doc (Bio-Rad) after staining with ethidium bromide (EB).

## Candidate gene prediction and testing marker for linkage to fruit shape of watermelon

The predicted genes in the mapping interval were downloaded from the watermelon database (<http://www.icugi.org/>). All predicted genes were sequenced and aligned to elongate and spherical watermelon to check the bases' difference and to obtain the candidate gene. The candidate gene function was retrieved through NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). An indel marker (Table S3) was developed according to the candidate gene sequence and tested on 100  $F_2$  individuals in 2017 spring. Marker was also tested to confirm the linkage on 105 watermelon progenies which contain 95 spherical watermelons and 10 elongate watermelons.

## Real-time PCR analysis for candidate gene expression

The ovaries and fruits from different developmental stages and others tissues from flowering stage were collected for RNA extraction to analyse the gene expression. Total RNA was extracted using the plant total RNA purification kit (GeneMark, China) following the manufacturer's instructions. The cDNA was synthesized with reverse transcriptase M-MLV (RNase H-) following the manufacturer's instructions (Takara, Japan).

The primers of candidate gene and reference gene *Actin* (Kong et al. 2015) used in quantitative reverse transcription polymerase chain reaction (qRT-PCR) were designed based on Cucurbit Genomics Database (<http://www.icugi.org>) (Table S3). Expression levels of the target gene were evaluated by qRT-PCR using a LightCycler480 RT-PCR system (Roche, Swiss). All reactions were performed using the SYBR Green real-time PCR mix according to the manufacturer's instructions. Each 20  $\mu$ L RT-PCR reaction mixture containing 1  $\mu$ L cDNA, 1  $\mu$ L forward primer (10  $\mu$ M), 1  $\mu$ L reverse primer (10  $\mu$ M), 10  $\mu$ L  $2 \times$  SYBR Green real-time PCR mix, and nuclease-free water to final volume of 20  $\mu$ L was preheated at 95  $^{\circ}$ C for 5 min, followed by 45 cycles of 95, 60, and 72  $^{\circ}$ C for 30 s. High-resolution melting was performed under the following conditions: 1 min at 95  $^{\circ}$ C, 1 min at 40  $^{\circ}$ C, 1 s at 65  $^{\circ}$ C, and continuous at 95  $^{\circ}$ C. All experiments were performed in triplicate. The raw data of qRT-PCR were analyzed using LCS480 software 1.5.0.39 (Roche, Swiss) and the relative expression was determined using the  $2^{-\Delta\Delta CT}$  method (Livak and Schmittgen 2001).

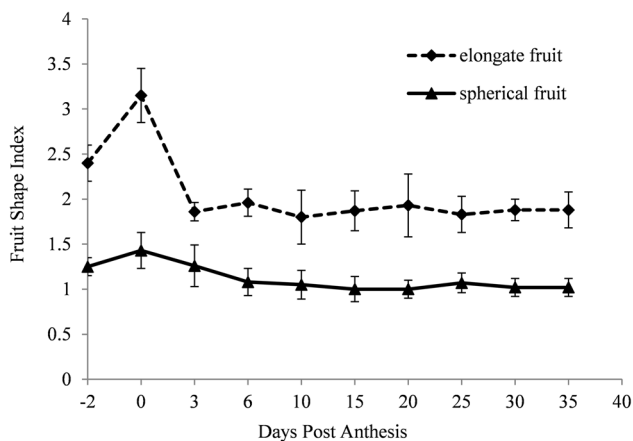


## Results

### Inheritance of fruit shape and changes of fruit shape index in watermelon

The watermelon fruit has three major kinds of shapes: elongate, oval, and spherical. We defined the fruit shape index of  $> 1.8$ ,  $1.4$ – $1.6$ , and  $1.0$ – $1.1$ , for elongate, oval, and spherical fruit, respectively. Segregation ratio of watermelon fruit shape (*CIFS1*) between two  $F_2$  populations in 2 years is presented in Table S1. All fruits length and width were measured to calculate the fruit shape index at fruit ripening stage. The *CIFS1* in the four segregating populations ( $F_2$ -2016 Winter,  $F_2$ -2017 Spring,  $BC_1P_1$ -2017 Spring, and  $BC_1P_2$ -2017 Spring) showed that the elongate fruit allele was incompletely dominant over the spherical fruit. The  $\chi^2$  test showed the segregation ratio of two  $F_2$  populations was in accordance with 1:2:1 (Table S1), suggesting that *CIFS1* was controlled by a single gene with incomplete dominance in this population.

To evaluate the effect of *CIFS1* on development, fruit shape index was monitored during fruit development in two parents. The dynamics of the fruit shape index in elongate and spherical watermelon followed the similar pattern from 2 days before anthesis (DBA) in ovaries to fruit ripening with the highest value occurring at anthesis stage which was 3.15 and 1.43, respectively (Fig. 2). The fruit shape index of elongate and spherical watermelon was stable from 25 day post-anthesis (DPA) to fruit ripening, which was 1.83 and 1.02, respectively (Fig. 2). However, the final fruit shape index of both elongate and spherical watermelon was similar to the respective shape index recorded 2 days before anthesis.

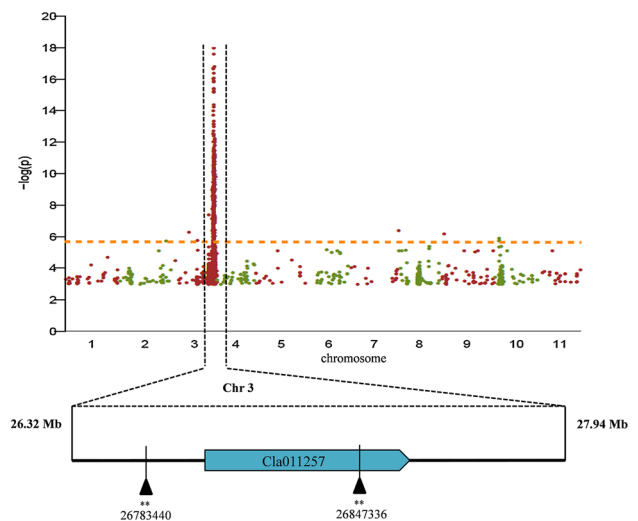


**Fig. 2** Dynamic changes of fruit shape index between elongate and spherical watermelon during the ovary and fruit development. At least three fruits per stage were evaluated

### GWAS identifies *CIFS1* gene located on chromosome 3

GWAS utilized a total of 315 watermelon accessions including elongate, oval, and spherical fruit shape to capture maximum genetic diversity, and was genotyped with a high-density, genome-wide coverage 4,661,625 and evenly spaced SNPs. To reduce the incidence of false-positive signals, a unified mixed linear model that controls population structure and familial relatedness was used (6PC + K (for 6 Principal Components and Kinship); Yu et al. 2006) to test associations between fruit shape index and 421,101 of the 4,661,625 SNPs with minor allele frequency  $> 5\%$ .

The difference of tenfold was calculated through the analysis of natural variation of watermelon fruit shape across the diversity panel coupled with a 85% (broad-sense) heritability, revealing that the observed natural variation is largely dictated by genetic variation across the population rather than environmental factors (Harjes et al. 2008; Chandler et al. 2013). GWAS profiles among 315 varieties identified the major locus of 26.32–27.94 Mb designated on watermelon chromosome 3 which might control watermelon fruit shape (Fig. 3). All the SNPs information in this interval is listed in Table S4. The strongly associated SNPs were located at 26,783,440 and 26,847,336 bp and the two SNPs were within the coding region of *Cla011249* and *Cla011257*, respectively (Fig. 3; Table S4).



**Fig. 3** Locus of watermelon fruit shape was identified through GWAS. Manhattan plot of genome-wide association for watermelon fruit shape showed that the region of 26.32–27.94 Mb on watermelon chromosome 3 controlled the fruit shape

### *CIFS1* gene located on chromosome 3 through BSA-seq

We constructed  $F_2$  population by crossing elongate watermelon ‘Duan125’ and spherical watermelon ‘Zhengzhouzigua’ (Fig. 1). A total of 30 extreme elongate individuals (E-pool) and 30 extreme spherical individuals (S-pool) from the  $F_2$  population in 2016 winter were selected and mixed to generate two pools. We sequenced the two pools on the Illumina HiSeq™ PE150 platform. A total of 30.1 GB of raw data were generated for both pools, with approximately  $30 \times$  depth and more than 99% coverage for each. The high-throughput sequencing resulted in 102,832,531 and 108,242,182 short reads from E-pool and S-pool, respectively. Data were aligned to ‘97103’ watermelon reference genome (<http://www.icugi.org>) and 304,565 SNPs were identified between two pools. Each identified SNP was used to compute an SNP-index. Using 1 kb sliding window, an average SNP-index was calculated in a 1 Mb interval. SNP-index graphs of E-pool and S-pool were generated by plotting the average SNP-index against the position of each sliding window in the ‘97103’ watermelon genome assembly. Graph for  $\Delta(\text{SNP-index})$  was plotted and computed against

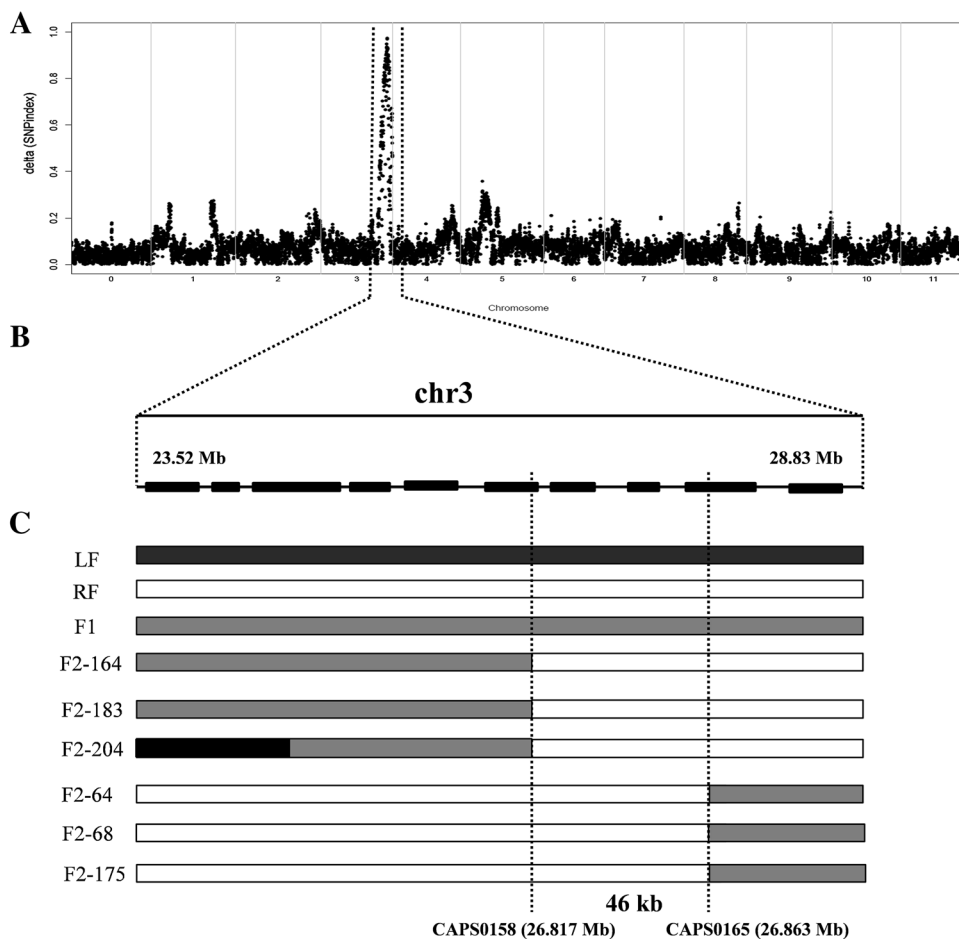
the genome positions by combining SNP-index of E-pool and S-pool (Fig. 4a).

The region on chromosome 3 from 23.52 to 28.83 Mb displayed a higher average than 0.5  $\Delta(\text{SNP-index})$  (Fig. 4a, b) and was significantly different from 0 of  $\Delta(\text{SNP-index})$  value at 95% significance level. The results showed that there was a candidate gene controlling watermelon fruit shape in that region.

### Analysis of the CAPS markers narrowed down *CIFS1* to a 46 kb interval

To confirm the watermelon fruit shape locus detected by BSA-seq and GWAS, 768  $F_2$  individuals from the 2016 winter and 2017 spring (Table S1) were analyzed. We developed 298 CAPS markers from chromosome 3 and screened the  $F_2$  segregating population for polymorphic analysis (Table S2). Only six recombinant individuals were obtained between markers CAPS0158 and CAPS0165. Since no reliable SNPs were found in the region between the two markers, further fine mapping was unfeasible. Therefore, the *CIFS1* gene was delimited in the region between markers CAPS0158

**Fig. 4** Genetic mapping of the fruit shape gene *CIFS1* in watermelon. **a**  $\Delta(\text{SNP-index})$  graph of BSA-seq analysis. **b** Locus at the interval of 23.52–28.83 Mb on chromosome 3 was identified to control watermelon fruit shape. **c** Examination of recombinants in  $F_2$  refined the location of *CIFS1* in an interval defined by two markers CAPS0158 and CAPS0165, which was a 46 kb interval



and CAPS0165, corresponding to 26.817 and 26.863 Mb (46 kb interval) on chromosome 3 (Fig. 4c).

### Identification of *ClFS1* gene

According to the watermelon genome database (<http://www.icugi.org/>), four putative genes were annotated in 46 kb interval (Table S5). DNA sequences of these four genes were obtained from the elongate fruit ‘Duan125’ and spherical fruit ‘Zhengzhouzigua’ parents. The CDS of *Cla011257* represented a 159 bp indel (530–688 bp) between the two parents (Fig. 5a, b), while no sequence changes were found in the other three genes between the two parents. The analysis of *Cla011257* protein sequence of ‘Duan125’ and ‘Zhengzhouzigua’ showed that 53 amino acids were deleted in elongate watermelon ‘Duan125’ (Fig. 5b). We proposed that *Cla011257* is the candidate gene for fruit shape in watermelon.

To verify this hypothesis, we analyzed the nucleotide polymorphism of *Cla011257* between two watermelon reference genomes spherical fruit ‘97103’ and elongate fruit ‘Charleston Gray’ (<http://www.icugi.org/>). A 159 bp deletion was also checked in ‘Charleston Gray’ compared with ‘97103’ (Fig. 6a), which was the same as above results. Furthermore, deletion of 53 amino acids was found in ‘Charleston Gray’ in the same region as above. The blast alignment to *tair* (<https://www.arabidopsis.org/>) showed that *Cla011257* is highly homologous to *AT3G16490* in Arabidopsis genome which encodes IQD26 proteins (Fig. 6). Network analysis of IQD26 protein in *Arabidopsis thaliana*

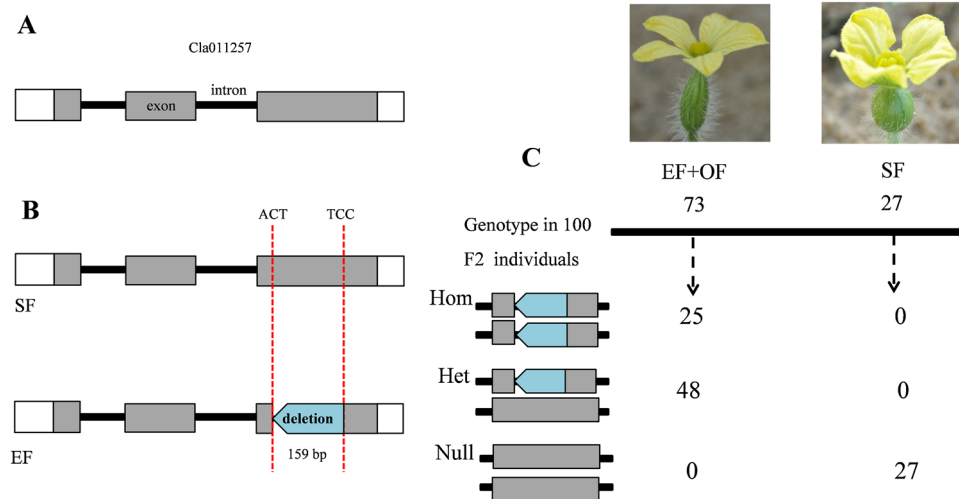
(<https://string-db.org/>) showed that many IQD family members were associated with microtubule arrays and  $\text{Ca}^{2+}$ -CaM signaling modules (Burstenbinder et al. 2017) and also had interaction with each other (Fig. 6b). Therefore, it further suggested that the *Cla011257* might be the candidate gene controlling watermelon fruit shape.

### Expression analysis of *Cla011257* during watermelon fruit development

The expression pattern of *Cla011257* was investigated using RT-PCR at flowering stage in different tissues that includes root, stem, young leaf, male flower, and flower bud; we also determined the expression level in ovary, pericarp, and flesh at different developmental stages of fruit in elongate and spherical plant parents. The primers for *Cla011257* and reference gene *Actin* are listed in Table S3. *Cla011257* had the highest expression level in ovaries formation stage in two parents. The expression level of *Cla011257* in elongate watermelon was almost threefold more than spherical watermelon at this stage (Fig. 7). During different developmental stages of fruit, transcript level of *Cla011257* was significantly lower in other tissues as compared to ovaries in the formation stage (Fig. 7).

### An indel marker was developed to check watermelon fruit shape

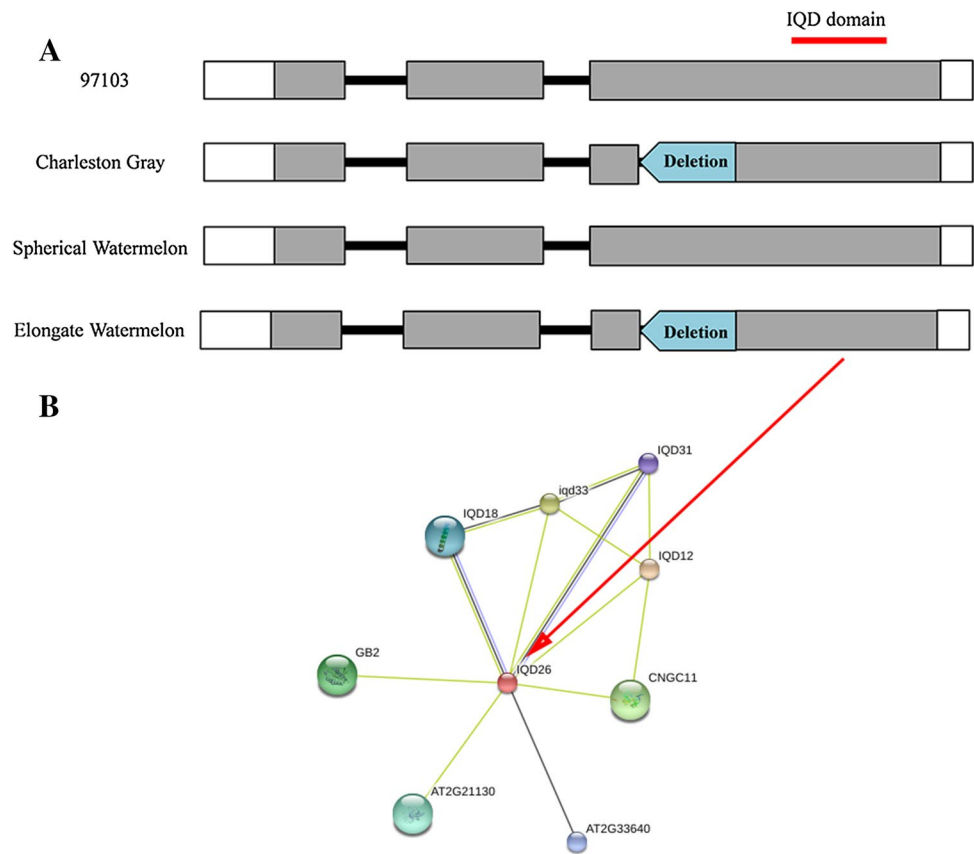
An indel marker was developed according to the 159 bp sequence of *Cla011257*. The primer for this marker is listed



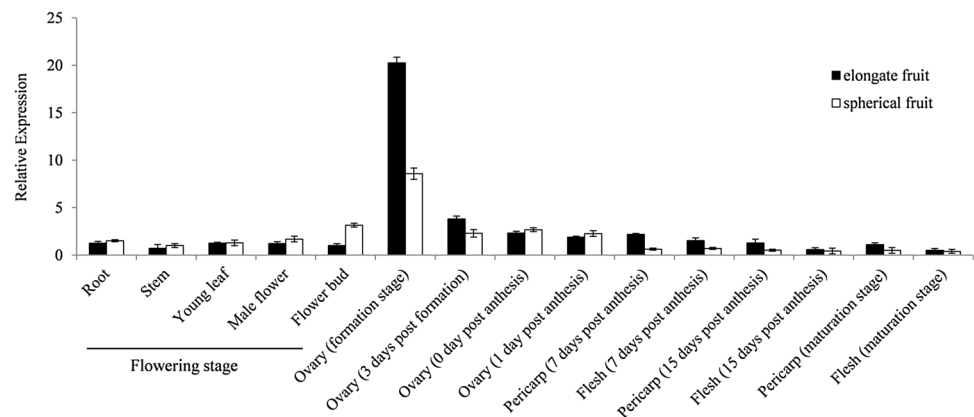
**Fig. 5** Identification of watermelon fruit shape gene *Cla011257*. **a** Structure of *Cla011257* gene. Open and gray boxes represent untranslated regions (UTRs) and exons, respectively, while lines denote introns. **b** Confirmation of the deletion in elongate watermelon by sequencing. The red dotted line indicates the 159 bp deletion from ACT to TCC. **c** Co-segregation of the fruit shapes phenotype and the

159 bp deletion of *Cla011257* in  $F_2$  population. Genotyping by PCR revealed that 25 elongate fruit individuals were dominant homozygous (Hom) for the deletion and the 48 oval fruit individuals were heterozygous (Het), whereas 27 spherical fruit individuals were recessive homozygous (Null) (color figure online)

**Fig. 6** Analysis of nucleotide polymorphisms and protein function prediction. **a** Nucleotide polymorphisms of Cla01257 among four watermelon genomes including two watermelon reference genomes ‘97103’ (spherical fruit) and ‘Charleston Gray’ (elongate fruit). The 159 bp deletion was also found in two reference genomes. **b** Network analysis of IQD26 protein which was blasted to high sequence consistency with Cla01257 in *A. thaliana* (<https://string-db.org/>) showed that many IQD family members had interaction each other



**Fig. 7** Relative expression level of Cla01257 during the different tissues and different development stages of ovary between elongate and spherical watermelons



in Table S3. To verify this marker, 100 individuals including the six recombinant individuals were selected in  $F_2$  population to check the polymorphism. The result confirmed that 25 elongate individuals were homozygous dominant and 48 oval individuals were heterozygous, whereas 27 spherical individuals were homozygous recessive, consistent with the phenotype (Fig. 5c). This marker was used to screen 105 watermelon accessions, which contained 10 elongate and 95 spherical fruits (Table S6). As expected, the genotypes perfectly matched the phenotype in spherical and elongate watermelon (Table S6). These results indicated that the *Cla01257* might control watermelon spherical fruit shape

and the deletion of 159 bp in *Cla01257* may result in the fruit elongation during evolutionary process.

## Discussion

The conventional QTL mapping method is a laborious and time-consuming process that the molecular markers are developed and every individual is genotyped and phenotyped in a mapping population. NGS-assisted BSA is less laborious, much cheaper, and has no population size limitation for genotyping work, because it provides an effective



and simple method to identify molecular markers linked to target genes/QTLs controlling the traits by genotyping only a pair of bulked DNA samples from two sets of individuals with distinct or opposite extreme phenotypes (Takagi et al. 2013; Yang et al. 2013). In this study, we employed this method together with GWAS to identify a candidate gene for fruit shape in watermelon.

As an important cucurbit crop throughout the world, watermelon presents different fruit shapes including elongate, oval, and spherical (Guo et al. 2013). However, the genetic pattern and the developmental mechanisms of watermelon fruit shape are not justified yet. Previously, the inheritance of the watermelon fruit shapes had been controversial, some researchers claimed that it was controlled by a single gene (*O*), and others thought that it was determined by a number of QTLs together (Tanaka et al. 1995; Guner and Wehner 2004; Sandlin et al. 2012; Zhang et al. 2012; Ren et al. 2014; Cheng et al. 2016). In our 2 year study, we found that watermelon fruit shape was controlled by a single, incompletely dominant gene resulting in fruit that are elongate (*OO*), oval (*Oo*), and spherical (*oo*).

Limited number of QTLs was identified in watermelon for fruit shape in the previous studies (Sandlin et al. 2012; Zhang et al. 2012; Ren et al. 2014; Cheng et al. 2016). Assuming that the QTLs belonged to the same QTL loci, the common genetic mechanisms could be shared underlying the fruit shape. Ren et al. (2014) identified three consensus QTLs for fruit length, fruit width, and fruit shape index on chromosome 3 of watermelon genome by constructing an integrated genetic map based on four mapping populations. Similarly, there were six QTLs for fruit length, fruit width, and fruit shape index on a genetic linkage map (Cheng et al. 2016). In this study, watermelon fruit shapes were divided into three distinct categories (elongate, oval, and spherical) by measuring fruit shape index and gene was located to a single locus on chromosome 3. To identify *CIFSI* gene, we performed a combinatorial approach by integrating BSA-seq in an F<sub>2</sub> population and GWAS in 315 watermelon accessions. The same locus on chromosomes 3 in watermelon genome was mapped using the GWAS- and NGS-assisted BSA approach (Figs. 3, 4). The BSA-seq was utilized for genome-wide identification of SNPs between two bulked pools, which had been used to develop molecular markers in gene mapping (Lee et al. 2014; Ramirez-Gonzalez et al. 2015). The *CIFSI* gene was delimited in the region of 26.817–26.863 Mb on chromosome 3 using 298 CAPS markers (Fig. 3c; Table S2). Sequence annotation analysis showed that there are four putative genes in this region (Table S5). Sequences alignment between elongate plant and spherical plant showed that the CDS of *Cla011257* represented a 159 bp deletion in elongate watermelon (Fig. 5a, b). The same 159 bp deletion was also checked in two watermelon reference genomes ‘Charleston Gray’ (elongate)

and ‘97103’ (spherical) (Fig. 6a). Based on the deletion, an indel marker was developed to test the polymorphism among 105 watermelon accessions. Results identified that *Cla011257* was the candidate gene for watermelon fruit shape (Table S6).

The critical role of cell division and cell enlargement in fruit shape has been demonstrated (Ando and Grumet 2010; Wang et al. 2017). In cucumber, the cell number and cell size in short cucumber fruit was lower than that of length, possibly due to abnormalities in auxin signaling in short fruit (Wang et al. 2017). Fruit development in many horticultural crops can be divided into three phases: ovary development, cell division, and cell expansion (Marcelis and Hofman-Eijer 1993). Cell division typically occurs about 3 day post-pollination and cell enlargement persists throughout the development of fruit, whereas ovary development occurred at pre-pollination (Marcelis and Hofman-Eijer 1993). In melon, it had been confirmed that the mature fruit shape exhibited a high correlation with the ovary and the fruit shape could be determined before pollination of ovary (Perin et al. 2002; Eduardo et al. 2007; Wei et al. 2016). In the present study, we noticed the similar mechanism that watermelon fruit shape can be predicted at the stage of ovary formation. The highest expression pattern of *CIFSI* had been detected at the stage of ovary formation, and the expression level of elongate fruit was almost threefold higher than that of spherical fruit (Fig. 7), suggesting that the fruit shape was predominantly determined by ovary and the high expression of *CIFSI* at the stage of ovary formation resulted in the watermelon fruit shape difference.

Genetic variants and genes mutation started from rare ones of very low frequency that occurred in one or few individuals in a given population, and these rare variants were turned into common ones through evolutionary forces such as selection, migration, and genetic drift. Even though there has been a great progress in several crops and model plants in generating comprehensive maps of genome variation (Lam et al. 2010; Chia et al. 2012; Qi et al. 2013; Lin et al. 2014). It had been reported that the watermelon genome speciation event occurred 15–23 million years ago (Guo et al. 2013). In the existing watermelon germplasm resources, most of the wild watermelons are spherical. In our study, the *Cla011257* had the same sequence as the reference genome of ‘97103’ bearing spherical fruit and a 159 bp deletion was detected in another reference genome of ‘Charleston Gray’ which has elongate fruit. Analysis of 105 watermelon accessions showed that the consistency between genotype and phenotype of *Cla011257* was 100% in spherical watermelon (Table S6). The most likely reason was that the *CIFSI* genetic variant resulted in fruit shape variation during the process of watermelon evolution. When watermelon genome occurred, the *CIFSI* determined spherical fruit and was recessively inherited.

The 159 bp deletion of *CIFSI* led to elongated fruit; inheritance was dominant and easy in descendant.

Phylogenetic analysis showed that *Cla011257* designated as *CIFSI* have high sequence identity with *AT3G16490* in Arabidopsis genome which belongs to IQD protein family. Network analysis of IQD26 protein in *A. thaliana* (<https://string-db.org/>) showed that many IQD family members which have been proved to be associated with microtubule arrays and  $\text{Ca}^{2+}$ -CaM signaling modules (Burstenbinder et al. 2017) have close interaction with each other (Fig. 6b). In tomato, the *SUN* controlling elongate fruit encodes a member of the IQD family of calmodulin-binding proteins which interacts with both calmodulin/calmodulin-like proteins (CaM/CMLs) and kinesin light-chain-related protein-1 (KLCR1); the IQD proteins can be recruited to microtubules through CaM/CMLs and KLCR1 (Abel et al. 2005; Xiao et al. 2008; Burstenbinder et al. 2013). The *SUN* gene led to elongate fruit by decreasing cell number in septum mediolateral and increasing cell number in proximal–distal direction of pericarp in tomato (Wu et al. 2011). Research in Arabidopsis showed that IQD domain families regulated cell shape and growth through the integrating CaM-dependent  $\text{Ca}^{2+}$  signaling and possibly other signal transduction pathways during plant development (Burstenbinder et al. 2017). IQD family may provide an assortment of versatile platform proteins that facilitate and specify CaM/CML dynamics during  $\text{Ca}^{2+}$  signaling at the cell periphery, on the cytoskeleton, and in the cell nucleus (Charpentier and Oldroyd 2013). Auxin could regulate the expression of IQDs, suggesting that plant growth and development regulation is the function of cellular auxin and calcium signaling (Cai et al. 2016). In this study, the different watermelon fruit shapes may be due to the *CIFSI* that provide a bridge of  $\text{Ca}^{2+}$  signaling, resulting in the difference of auxin level in cells.

Discovery of watermelon fruit shape gene provides a good entry point to explain the genetic mechanisms of fruit development as well as providing fundamental insights into the domestication and selection history of watermelon. The results of this study demonstrated the candidate gene of watermelon fruit shape (*CIFSI*) *Cla011257*, an indel marker was developed which can be used to identify the fruit shape at the seedling stage. In addition, the current results will be useful in marker-assisted breeding for selection of desired traits at early stage of seedling.

**Author contribution statement** HK and WL conceived the research and designed the experiments. XL and NH developed the plants population. LZ and SZ analyzed BSA and GWAS data. AA checked the manuscript. JD performed most of the experiment and wrote the manuscript. All authors reviewed and approved this submission.

**Acknowledgements** This work was supported by the Agricultural Science and Technology Innovation Program (ASTIP) (CAAS-ASTIP-2017-ZFR1), the China Agriculture Research System (CARS-25-03), and the Central Public-interest Scientific Institution Basal Research Fund, China (1610192016301 and 1610192016209).

## Compliance with ethical standards

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

**Ethical standard** The experiments in this study comply with the current laws of China.

## References

- Abe A, Kosugi S, Yoshida K, Natsume S, Takagi H, Kanzaki H, Matsumura H, Yoshida K, Mitsuoka C, Tamiru M, Innan H, Cano L, Kamoun S, Terauchi R (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat Biotechnol* 30:174–178
- Abel S, Savchenko T, Levy M (2005) Genome-wide comparative analysis of the IQD gene families in *Arabidopsis thaliana* and *Oryza sativa*. *BMC Evol Biol* 5:72
- Ando K, Grumet R (2010) Transcriptional profiling of rapidly growing cucumber fruit by 454-pyrosequencing analysis. *J Am Soc Hortic Sci* 135:291–302
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J R Stat Soc B Methods* 57:289–300
- Bo KL, Ma Z, Chen JF, Weng YQ (2015) Molecular mapping reveals structural rearrangements and quantitative trait loci underlying traits with local adaptation in semi-wild Xishuangbanna cucumber (*Cucumis sativus* L. var. *xishuangbannanensis* Qi et Yuan). *Theor Appl Genet* 128:25–39
- Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin H, Sari MA, Collin F, Flowers JM, Pitrat M, Purugganan MD, Dogimont C, Bendahmane A (2008) A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. *Science* 321:836–838
- Burstenbinder K, Savchenko T, Muller 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
- Burstenbinder K, Moller B, Plotner 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 RH, Zhang CS, Zhao Y, Zhu KJ, Wang YF, Jiang HY, Xiang Y, Cheng BJ (2016) Genome-wide analysis of the IQD gene family in maize. *Mol Genet Genom* 291:543–558
- Chandler K, Lipka AE, Owens BF, Li HH, Buckler ES, Rocheford T, Gore MA (2013) Genetic analysis of visually scored orange kernel color in maize. *Crop Sci* 53:189–200
- Charpentier M, Oldroyd GED (2013) Nuclear calcium signaling in plants. *Plant Physiol* 163:496–503
- Cheng Y, Luan FS, Wang XZ, Gao P, Zhu ZC, Liu S, Baloch AM, Zhang YS (2016) Construction of a genetic linkage map of watermelon (*Citrullus lanatus*) using CAPS and SSR markers and QTL analysis for fruit quality traits. *Sci Hortic* 202:25–31

- Chia JM, Song C, Bradbury PJ, Costich D, de Leon N, Doebley J, Elshire RJ, Gaut B, Geller L, Glaubitz JC, Gore M, Guill KE, Holland J, Hufford MB, Lai JS, Li M, Liu X, Lu YL, McCombie R, Nelson R, Poland J, Prasanna BM, Pyhajarvi T, Rong TZ, Sekhon RS, Sun Q, Tenaillon MI, Tian F, Wang J, Xu X, Zhang ZW, Kaeppeler SM, Ross-Ibarra J, McMullen MD, Buckler ES, Zhang GY, Xu YB, Ware D (2012) Maize HapMap2 identifies extant variation from a genome in flux. *Nat Genet* 44:803–U238
- Chusreeaom K, Ariizumi T, Asamizu E, Okabe Y, Shirasawa K, Ezura H (2014) A novel tomato mutant, *Solanum lycopersicum* elongated fruit1 (Slef1), exhibits an elongated fruit shape caused by increased cell layers in the proximal region of the ovary. *Mol Genet Genom* 289:399–409
- Eduardo I, Arus P, Monforte AJ, Obando J, Fernandez-Trujillo JP, Martinez JA, Alarcon AL, Alvarez JM, van der Knaap E (2007) Estimating the genetic architecture of fruit quality traits in melon using a genomic library of near isogenic lines. *J Am Soc Hortic Sci* 132:80–89
- Fernandez-Silva I, Moreno E, Essafi A, Fergany M, Garcia-Mas J, Martin-Hernandez AM, Alvarez JM, Monforte AJ (2010) Shaping melons: agronomic and genetic characterization of QTLs that modify melon fruit morphology. *Theor Appl Genet* 121:931–940
- Garcia-Mas J, Benjak A, Sanseverino W, Bourgeois M, Mir G, Gonzalez VM, Henaff E, Camara F, Cozzuto L, Lowy E, Alioto T, Capella-Gutierrez S, Blanca J, Canizares J, Ziarso P, Gonzalez-Ibeas D, Rodriguez-Moreno L, Droegge M, Du L, Alvarez-Tejado M, Lorente-Galdos B, Mele M, Yang LM, Weng YQ, Navarro A, Marques-Bonet T, Aranda MA, Nuez F, Pico B, Gabaldon T, Roma G, Guigo R, Casacuberta JM, Arus P, Puigdomenech P (2012) The genome of melon (*Cucumis melo* L.). *Proc Natl Acad Sci USA* 109:11872–11877
- Guner N, Wehner TC (2004) The genes of watermelon. *HortScience* 39:1175–1182
- Guo SG, Zhang JG, Sun HH, Salse J, Lucas WJ, Zhang HY, Zheng Y, Mao LY, Ren Y, Wang ZW, Min JM, Guo XS, Murat F, Ham BK, Zhang ZL, Gao S, Huang MY, Xu YM, Zhong SL, Bombarely A, Mueller LA, Zhao H, He HJ, Zhang Y, Zhang ZH, Huang SW, Tan T, Pang EL, Lin K, Hu Q, Kuang HH, Ni PX, Wang B, Liu JA, Kou QH, Hou WJ, Zou XH, Jiang J, Gong GY, Klee K, Schoof H, Huang Y, Hu XS, Dong SS, Liang DQ, Wang J, Wu K, Xia Y, Zhao X, Zheng ZQ, Xing M, Liang XM, Huang BQ, Lv T, Wang JY, Yin Y, Yi HP, Li RQ, Wu MZ, Levi A, Zhang XP, Giovannoni JJ, Wang J, Li YF, Fei ZJ, Xu Y (2013) The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat Genet* 45:51
- Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan JB, Buckler ES (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319:330–333
- Huang SW, Li RQ, Zhang ZH, Li L, Gu XF, Fan W, Lucas WJ, Wang XW, Xie BY, Ni PX, Ren YY, Zhu HM, Li J, Lin K, Jin WW, Fei ZJ, Li GC, Staub J, Kilian A, van der Vossen EAG, Wu Y, Guo J, He J, Jia ZQ, Ren Y, Tian G, Lu Y, Ruan J, Qian WB, Wang MW, Huang QF, Li B, Xuan ZL, Cao JJ, Asan WuZG, Zhang JB, Cai QL, Bai YQ, Zhao BW, Han YH, Li Y, Li XF, Wang SH, Shi QX, Liu SQ, Cho WK, Kim JY, Xu Y, Heller-Uszynska K, Miao H, Cheng ZC, Zhang SP, Wu J, Yang YH, Kang HX, Li M, Liang HQ, Ren XL, Shi ZB, Wen M, Jian M, Yang HL, Zhang GJ, Yang ZT, Chen R, Liu SF, Li JW, Ma LJ, Liu H, Zhou Y, Zhao J, Fang XD, Li GQ, Fang L, Li YR, Liu DY, Zheng HK, Zhang Y, Qin N, Li Z, Yang GH, Yang S, Bolund L, Kristiansen K, Zheng HC, Li SC, Zhang XQ, Yang HM, Wang J, Sun RF, Zhang BX, Jiang SZ, Wang J, Du YC, Li SG (2009) The genome of the cucumber, *Cucumis sativus* L. *Nat Genet* 41:1275–U1229
- Jiang L, Yan SS, Yang WC, Li YQ, Xia MX, Chen ZJ, Wang Q, Yan LY, Song XF, Liu RY, Zhang XL (2015) Transcriptomic analysis reveals the roles of microtubule-related genes and transcription factors in fruit length regulation in cucumber (*Cucumis sativus* L.). *Sci Rep* 5:8031
- Kong QS, Yuan JX, Gao LY, Zhao LQ, Cheng F, Huang Y, Bie ZL (2015) Evaluation of appropriate reference genes for gene expression normalization during watermelon fruit development. *PLoS One* 10:e0130865
- Lam HM, Xu X, Liu X, Chen WB, Yang GH, Wong FL, Li MW, He WM, Qin N, Wang B, Li J, Jian M, Wang JA, Shao GH, Wang J, Sun SSM, Zhang GY (2010) Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat Genet* 42:1053–U1041
- Lee YP, Cho Y, Kim S (2014) A high-resolution linkage map of the Rfd1, a restorer-of-fertility locus for cytoplasmic male sterility in radish (*Raphanus sativus* L.) produced by a combination of bulked segregant analysis and RNA-Seq. *Theor Appl Genet* 127:2243–2252
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc GPD (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079
- Lin T, Zhu GT, Zhang JH, Xu XY, Yu QH, Zheng Z, Zhang ZH, Lun YY, Li S, Wang XX, Huang ZJ, Li JM, Zhang CZ, Wang TT, Zhang YY, Wang AX, Zhang YC, Lin K, Li CY, Xiong GS, Xue YB, Mazzucato A, Causse M, Fei ZJ, Giovannoni JJ, Chetelat RT, Zamir D, Stadler T, Li JF, Ye ZB, Du YC, Huang SW (2014) Genomic analyses provide insights into the history of tomato breeding. *Nat Genet* 46:1220–1226
- Lipka AE, Tian F, Wang QS, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang ZW (2012) GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28:2397–2399
- Liu JP, Van Eck J, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc Natl Acad Sci USA* 99:13302–13306
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(–Delta Delta C) method. *Methods* 25:402–408
- Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* 82:1420–1425
- Lou LL, Wehner TC (2016) Qualitative inheritance of external fruit traits in watermelon. *HortScience* 51:487–496
- Marcelis L, Hofman-Eijer B (1993) Cell division and expansion in the cucumber fruit. *J Hortic Sci* 68:665–671
- Miao H, Gu XF, Zhang SP, Zhang ZH, Huang SW, Wang Y, Cheng ZC, Zhang RW, Mu S, Li M, Zhang ZX, Fang ZY (2011) Mapping QTLs for fruit-associated traits in *Cucumis sativus* L. *Sci Agric Sin* 44:5031–5040
- Munos S, Ranc N, Botton E, Berard A, Rolland S, Duffe P, Carretero Y, Le Paslier MC, Delalande C, Bouzayen M, Brunel D, Causse M (2011) Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. *Plant Physiol* 156:2244–2254
- Pan YP, Liang XJ, Gao ML, Liu HQ, Meng HW, Weng YQ, Cheng ZH (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
- Paran I, van der Knaap E (2007) Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *J Exp Bot* 58:3841–3852
- Paris MK, Zalapa JE, McCreight JD, Staub JE (2008) Genetic dissection of fruit quality components in melon (*Cucumis melo* L.) using

- a RIL population derived from exotic x elite US Western Shipping germplasm. *Mol Breed* 22:405–419
- Perin C, Hagen LS, Giovinazzo N, Besombes D, Dogimont C, Pitrat M (2002) Genetic control of fruit shape acts prior to anthesis in melon (*Cucumis melo* L.). *Mol Genet Genom* 266:933–941
- Poole CF, Grimboll PC (1945) Interaction of sex, shape, and weight genes in watermelon. *J Agric Res* 71:533–552
- 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
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909
- Qi JJ, Liu X, Shen D, Miao H, Xie BY, Li XX, Zeng P, Wang SH, Shang Y, Gu XF, Du YC, Li Y, Lin T, Yuan JH, Yang XY, Chen JF, Chen HM, Xiong XY, Huang K, Fei ZJ, Mao LY, Tian L, Stadler T, Renner SS, Kamoun S, Lucas WJ, Zhang ZH, Huang SW (2013) A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nat Genet* 45:1510–U1149
- Ramirez-Gonzalez RH, Segovia V, Bird N, Fenwick 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
- Ren Y, McGregor C, Zhang Y, Gong GY, Zhang HY, Guo SG, Sun HH, Cai WT, Zhang J, Xu Y (2014) An integrated genetic map based on four mapping populations and quantitative trait loci associated with economically important traits in watermelon (*Citrullus lanatus*). *BMC Plant Biol* 14:33
- Rodriguez GR, Munos S, Anderson C, Sim SC, Michel A, Causse M, Gardener BBM, Francis D, van der Knaap E (2011) Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. *Plant Physiol* 156:275–285
- Rodriguez GR, Kim HJ, van der Knaap E (2013) Mapping of two suppressors of OVATE (*sov*) loci in tomato. *Heredity* 111:256–264
- Sandlin K, Prothro J, Heesacker A, Khalilian N, Okashah R, Xiang WW, 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
- Schwarz G (1978) Estimating the dimension of a model. *Ann Stat* 6:461–464
- Sun G, Zhu C, Kramer MH, Yang SS, Song W, Piepho HP, Yu J (2010) Variation explained in mixed-model association mapping. *Heredity* 105:333–340
- Sun L, Rodriguez GR, Clevenger JP, Illa-Berenguier E, Lin JS, Blakeslee JJ, Liu WL, Fei ZJ, Wijeratne A, Meulia T, van der Knaap E (2015) Candidate gene selection and detailed morphological evaluations of fs8.1, a quantitative trait locus controlling tomato fruit shape. *J Exp Bot* 66:6471–6482
- Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S, Innan H, Cano LM, Kamoun S, Terauchi R (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
- Tanaka T, Wimol S, Mizutani T (1995) Inheritance of fruit shape and seed size of watermelon. *J Jpn Soc Hortic Sci* 64:543–548
- Tomason Y, Nimmakayala P, Levi A, Reddy UK (2013) Map-based molecular diversity, linkage disequilibrium and association mapping of fruit traits in melon. *Mol Breed* 31:829–841
- van der Knaap E, Chakrabarti M, Chu YH, Clevenger JP, Illa-Berenguier 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* 5:1–13
- Wang LN, Cao CX, Zheng SS, Zhang HY, Liu PJ, Ge Q, Li JR, Ren ZH (2017) Transcriptomic analysis of short-fruit 1 (*sf1*) reveals new insights into the variation of fruit-related traits in *Cucumis sativus*. *Sci Rep* 7:2950
- Weetman LM (1937) Inheritance and correlation of shape, size and color in the watermelon, *Citrullus vulgaris* Schrad. *Res Bull* 20:1
- Wei QZ, Fu WY, Wang YZ, Qin XD, Wang J, Li J, Lou QF, Chen JF (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 YQ, Colle M, Wang YH, Yang LM, Rubinstein M, Sherman A, Ophir R, 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
- Yang XY, Wang Y, Jiang WJ, Liu XL, Zhang XM, Yu HJ, Huang SW, Liu GQ (2013) Characterization and expression profiling of cucumber kinesin genes during early fruit development: revealing the roles of kinesins in exponential cell production and enlargement in cucumber fruit. *J Exp Bot* 64:4541–4557
- Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38:203–208
- Zhang ZW, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu JM, Arnett DK, Ordoñas JM, Buckler ES (2010) Mixed linear model approach adapted for genome-wide association studies. *Nat Genet* 42:355–U118
- Zhang HY, Wang H, Guo SG, Ren Y, Gong GY, Weng YQ, Xu Y (2012) Identification and validation of a core set of microsatellite markers for genetic diversity analysis in watermelon, *Citrullus lanatus* Thunb, Matsum, & Nakai. *Euphytica* 186:329–342