Mol Breeding (2020) 40:19
https://doi.org/10.1007/s11032-020-1104-2

Markers for selection of three alleles of *ClSUN25-26-27a (Cla011257)* associated with fruit shape in watermelon

Reeve Legendre · Jesse Kuzy · Cecilia McGregor

Received: 18 October 2019 / Accepted: 19 January 2020

Abstract Fruit shape is an economically important trait in many horticultural crops. The genes and mechanisms underlying fruit shape variation are poorly understood in watermelon. A deletion (DEL) in *ClSUN25-26-27a (Cla011257)*, a member of the *SUN* gene family, has been shown to be associated with elongated fruit. In the current study, we sequenced *ClSUN25-26-27a* in three watermelon cultivars with different fruit shape indices (FSIs). A novel allele containing a non-synonymous point mutation in exon 3 was identified in Klondike Black Seeded. KASP assays (ClSUN-1 and ClSUN-2) were developed to distinguish three *ClSUN25-26-27a* alleles (WT, KBS, and DEL), and their effects on ovary and fruit shape traits were determined in a cultivar panel and segregating populations. The different alleles showed a significant association with ovary and fruit length and shape as well as the angles of the distal and proximal ends. These effects were primarily pre-anthesis. The KASP assays can serve as tools for marker-assisted selection (MAS) of fruit length and shape in watermelon. This study highlights the allelic variation in *ClSUN25-26-27a* and its association with fruit shape in watermelon.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11032-020-1104-2) contains supplementary material, which is available to authorized users.

R. Legendre · J. Kuzy · C. McGregor (✉)
Department of Horticulture, University of Georgia, 1111 Miller Plant Sciences, Athens, GA 30602, USA
e-mail: cmcgre1@uga.edu

Keywords *Citrullus lanatus* · Watermelon · Fruit · Length · Shape · KASP · *Cla011257* · *ClSUN25-26-27a* · Digital phenotyping

Introduction

Watermelon (*Citrullus lanatus*) is a major horticultural crop from the Cucurbitaceae family grown throughout the temperate regions of the world. Worldwide production of watermelon in 2016 was valued over 33 billion US dollars. The majority of crop production is dominated by China, which produces nearly 68% of the world's watermelon. The USA produces only 1.6% of the total global yield but is still the seventh leading producer worldwide behind Turkey, Iran, Brazil, Uzbekistan, and Algeria (FAOSTAT 2016). US production is centered in Texas, Florida, Georgia, and California, with two of the top four states located in the southeastern USA. Because of watermelon's economic importance, development of new cultivars with desirable agronomic and fruit characteristics is a high priority for watermelon breeding.

Fruit shape is an important trait in watermelon breeding due to the role it plays in consumer preference, packaging, and transportation logistics. At the point of sale, consumers often prefer one shape of watermelon over others, with current trends moving away from large, elongated watermelons toward smaller, blocky, or round watermelon. Watermelon fruits develop from enlarged, inferior ovaries that grow into mature fruit after pollination. The mature fruit shape is highly

Published online: 29 January 2020

correlated with ovary shape at anthesis (Tanaka et al. 1995; Weetman 1937). Fruit phenotypes have been associated with the *O* gene, with round fruit recessive to elongate fruit and heterozygotes producing oval fruit (Poole and Grimball 1945; Weetman 1937). Other genes that have been associated with fruit shape include the *andromonoecious* (*a*) and *pistillate* (*p*) genes, with hermaphroditic flowers having rounder fruit shape and pistillate flowers having more elongated fruit shapes (Loy 2005; Poole and Grimball 1945; Rosa 1928). Fruit shape in other cucurbits is also mainly determined pre-anthesis, although more variation in shape has been observed in some species (Pan et al. 2017; Périn et al. 2002; Stepansky et al. 1999; Weng et al. 2015).

Quantitative trait loci (QTL) mapping in watermelon has revealed several QTL that control fruit shape. Since fruit shape is generally described in terms of fruit shape index (FSI = fruit length [FL]/fruit diameter [FD]), QTL for FSI often overlap with QTL for FL and FD. A stable major effect QTL for FSI has been identified on chromosome 3 ($Qfsi3^M$) in multiple genetic backgrounds (Dou et al. 2018; Kim et al. 2015; Liu et al. 2016; Sandlin et al. 2012). A QTL associated with FSI (*Qfsi2*) was also identified on chromosome 2 in two different genetic backgrounds (Cheng et al. 2016; Sandlin et al. 2012). The FSI QTL on chromosome 3 is associated with variation in FL, while the FSI QTL on chromosome 2 is associated with variation in FD. *Cla011257* (*ClSUN25-26-27a*) was identified as a putative candidate gene co-localizing with *ClFSI3.1* in several recent studies (Dou et al. 2018; Kim et al. 2015; Pan et al. 2017; Pan et al. 2019). *Cla011257* is a member of the SUN gene family in tomato which encodes a member (*ClSUN25-26-27a*; Pan et al. 2019) of the IQD family of calmodulin-binding proteins. An increase in expression of the *SUN* gene causes an increase in cell numbers in the proximal-distal direction in tomato fruit which is accompanied by a decrease in cell numbers in the lateral direction, resulting in elongated fruit. In tomato, these cell patterns are established pre-anthesis, but the effect is more obvious shortly after pollination (Wu et al. 2011). In contrast to tomato, where the variation in fruit shape associated with *SUN* is due to a duplication event leading to higher expression, the effect of *ClSUN25-26-27a* in watermelon is associated with allelic variation in the 3rd exon of the gene. Four alleles have been described to date, including 3 single nucleotide polymorphisms (SNPs) (Kim et al. 2015) and a 159 bp deletion that was associated with increased expression of *ClSUN25-26-27a*, and an elongated fruit phenotype as opposed to a round or spherical fruit phenotype (Dou et al. 2018).

In order to exploit the allelic variation of *ClSUN25-26-27a* in watermelon breeding, high-throughput genotyping assays are needed for marker-assisted selection of desired phenotypes. The availability of molecular marker technologies like Kompetitive Allele Specific PCR (KASP) assays makes genotyping of individuals quick and cost-efficient compared to older marker technologies. Additionally, KASP assays can differentiate genotypes using small amounts of DNA extracted using inexpensive extraction methods and eliminates the need for gel electrophoresis.

Fruit shape is an important quality trait in watermelon, and *ClSUN25-26-27a* is associated with elongated fruit. The aim of the current study was to further understand the genetic effects of allelic variation in *ClSUN25-26-27a* on fruit and ovary shape in watermelon and to develop high-throughput markers for selection of these traits. First, a search was conducted for allelic variations between the parents of a population segregating for fruit shape. High-throughput KASP markers for the alleles were developed, and the association of the genotype with ovary and fruit phenotypic variation in selected watermelon cultivars and lines was determined. Lastly, the marker trait association was validated in segregating populations with different genetic backgrounds.

Materials and methods

Sequencing of *ClSUN25-26-27a*

Three different cultivars (Klondike Black Seeded [KBS], New Hampshire Midget [NHM], and Charleston Gray [CG]), exhibiting three distinct phenotypes, were sequenced using Phusion® High Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA USA). Sequencing primers (Table S1) were designed using Primer3Plus™ (www.primer3plus.com) for *Cla011257* in the 97103 reference genome sequence (Guo et al. 2013). Two different plants of each cultivar were sequenced. A master mix for each primer pair was prepared using 12.6 μL of sterile water, 0.6 μL of DMSO, 4.0 μL of 5x Phusion® High-Fidelity buffer, 0.4 μL of dNTP mix, 0.1 μL of forward and reverse primers, and 0.2 μL of Phusion® High-Fidelity Polymerase. The master mix was combined with 2 μL of DNA with a concentration of 50–150 μg/μL. PCR amplification was

carried out in a Bio-Rad S1000™ Thermocyler (Bio-Rad, Hercules, CA USA) using the following PCR conditions: 98 °C for 30 s, followed by 35 cycles of 98 °C for 5 s, 65 °C for 20 s, and 72 °C for 30 s and finishing at 72 °C for 10 min. A total of 3 μL of the amplified product was separated on 1.5% agarose gels in order to confirm specific amplification. PCR products were purified using the E.Z.N.A.® Cycle Pure Kit (Bio-Rad, Hercules, CA) following the manufacturer's instructions and sent to Eurofins Genomics LLC (Louisville, KY) for Sanger sequencing. Sequences were trimmed for quality and aligned to the 97103 (Guo et al. 2013) and Charleston Gray (Wu et al. 2019) genomes using Geneious™ version 8.1.9 (Kearse et al. 2012).

Development of KASP assays and genotyping

KASP™ (LGC Genomics, Teddington, UK) primers were developed using Primer3Plus™ (www.primer3plus.com). The ClSUN-1 assay was designed to differentiate the NHM (wild type, WT) and CG (deletion, DEL) alleles, while ClSUN-2 differentiated the NHM (WT) and KBS point mutation alleles (Table 1, Fig. S1). KASP assay reactions contained 1.94 μL of 2 × KASP Master Mix (LGC Genomics), 0.06 μL of KASP Primer Mix, and 2 μL of DNA (approximately 200 ng/μl). The primer mix consisted of 12 μL of each of the forward primers (100 μM), 30 μL of reverse primer (100 μM), and 46 μL of autoclaved distilled water. PCR conditions were as follows: 95 °C for 15 min; followed by 10 cycles of touchdown PCR at 95 °C for 20s, 67 °C with 1 °C decrease each cycle for 25 s, and 72 °C for 15 s; and 35 cycles at 95 °C for 10s, 58 °C for 60s, and 72 °C for 15 s. The assay fluorescence was quantified using an Infinite M200 Pro plate reader (Tecan, Männedorf, Switzerland) and genotypic calls were made using KlusterCaller™ (LGC Genomics).

Plant material for 2016 and 2017 field phenotyping

In order to obtain ovary and mature fruit phenotypes, the following 10 genotypes, including cultivars, selections, and plant introductions (PI), were grown in a randomized complete block design with five replicates at the UGA Durham Horticulture Farm (Watkinsville, GA): CG, NHM, KBS, Allsweet (AS), AU-Producer (AUP), Calhoun Gray (CALG), Crimson Sweet (CS), Mickylee (MICK), UGA147, and Egusi (PI 560023) (Table S2). The within-row spacing was 1.2 m with 3.7 m between rows in 2016 and 1.8 m between rows in 2017.

Phenotyping of plant material in 2016 and 2017 field studies

Ovary measurements were taken on the day of anthesis using a digital caliper. Ovary length (oFL) was measured from the proximal end to the distal end, and ovary diameter (oFD) was measured at the widest part of the ovary. Ovaries were tagged and numbered to enable mature fruit measurements to be taken on the same samples. Mature fruits were cut in half lengthwise, and length was measured from the proximal to distal end, while diameter was measured at the widest part of the fruit using a ruler. Data was analyzed using JMP® Pro, version 13.2.1 (SAS Institute Inc. 1989–2007). Correlations were calculated using pairwise Pearson's correlation coefficients.

Table 1 KASP Assay Primers for ClSUN-1 and ClSUN-2

Assay name	Label/primer	Sequence (5′-3′)	*ClSUN25-26-27A (Cla011257)* allele detected
ClSUN1	FAM	GAAGGTGACCAAGTTCATGCTCGGTACACAGCAGAAGACTCCC	DEL
	VIC	GAAGGTCGGAGTCAACGGATTCACCAAACTCTCTCGTCTCCACTC	WT
	Reverse	GATCCACCGCCAGAGTTCAC	
ClSUN-2	FAM	GAAGGTGACCAAGTTCATGCTCCCCAAAATCGTGGAAATCA	KBS
	VIC	GAAGGTCGGAGTCAACGGATTCCCCAAAATCGTGGAAATCG	WT
	Reverse	GCACCAGTCAGAGTCATGGA	

ClSUN-1 is used to differentiate between the DEL and WT alleles, while ClSUN-2 is used to differentiate between the KBS and WT alleles

Plant material in 2018 parental panel

In 2018, eight parental cultivars were grown at the UGA Durham Horticulture farm in a randomized complete block design with three replicates of 12 plants per replicate with a 1.2 m in-row plant spacing and 1.8 m between-row spacing. The parental cultivars were NHM, Jade Star (JS), SB, KBS, Fiesta, CG, and CALG. UGA147 was not included due to lack of seed.

Validation populations

The KBS x NHM recombinant inbred line (RIL) population used by Sandlin et al. (2012) to map *Qfsi3*^{M} was used for validation of the ClSUN-2 KASP assay (n = 136). Two other F_2 populations, Sugar Baby x Calhoun Gray (SB x CALG; n = 125) and CG x UGA147 (n = 115) were previously phenotyped (Meru and McGregor 2016a, b) in the greenhouse (Athens, GA USA). Plants were grown vertically and hand-pollinated. Fruits were bagged in mesh bags and harvested at maturity. Fruit measurements were taken as described for the 2016 and 2017 Field Studies. Both populations were used for validation of KASP assay ClSUN-1.

Two additional F_2 populations were grown at the UGA Durham Horticulture Research farm in 2018. Jade Star x Fiesta (JS x Fiesta; n = 42) was grown in order to validate the ClSUN-2 assay, while NHM x Calhoun Gray (NHM x CALG; n = 52) was grown for validation of the ClSUN-1 assay.

Phenotyping of plant material in 2018 cultivar panel and validation populations

In 2018, phenotyping was performed digitally. For the cultivar panel, ovaries were harvested on the day of anthesis from each cultivar and placed into a petri dish with a moist paper towel to prevent desiccation. Ovaries were cut in half from the proximal end to the distal end and scanned using an HP® DeskJet F4280 (Palo Alto, CA) at a pixel density of 600 pixels per inch (PPI). Mature fruit were also digitally imaged using a mobile photography setup comprised of a Nikon™ Rebel XSi camera (Tokyo, JP) mounted to a PVC frame. Watermelons were cut in half from the proximal to distal end and placed on a black background with an X-Rite ColorChecker® color card (Grand Rapids, Michigan USA) and ruler for size reference. In order to remove background noise and isolate the fruit for digital phenotyping, the fruit were extracted from the images using a custom MATLAB R2018a script (Mathworks Inc., Natick, MA USA). The script extracted the fruit by transforming RGB images to HSV, binarizing the value channel, and creating a mask from the binarized map. The mask was filled to remove gaps created by seeds and roughly cropped to remove problematic background features in the perimeter of images, after which the largest object remaining in the mask was determined to be the fruit. A list of fruit pixels was created from the fruit mask, and the fruit image was reassembled on a black background and consolidated into images suitable for analysis. Consolidated images were analyzed using the software program Tomato Analyzer version 3.0 (Gonzalo et al. 2009; Brewer et al. 2006).

Tomato Analyzer was used to measure the maximum length and maximum diameter (FD) of each fruit, which were in turn used to calculate the FSI (Fig. S2). Additionally, the distal end angle (DEA) and proximal end angle (PEA) were measured after manually setting the proximal and distal end points in Tomato Analyzer. Best fit lines were drawn through 5% of the total perimeter on either side of the fruit from the proximal and distal end points to establish the angle (Fig. S2).

DNA extraction

Genomic DNA for KASP analysis of all plants was extracted from leaf material frozen at − 80 °C using a modified SDS-NaCL DNA extraction method (Fall et al. 2018), adapted from King et al. 2014. DNA concentrations were measured using a Tecan Infinite M200 Pro plate reader and diluted to approximately 100 ng/μL in sterile distilled water.

Results

Sequencing of *ClSUN25-26-27a* reveals three distinct alleles

Alignment of sequences to the 97103 genome revealed three alleles (Fig. 1a). The NHM sequence was identical to the 97103 reference genome (Fig. 1a). We designate this allele as the wild type (WT). The KBS sequence had a non-synonymous point mutation at bp 597 in exon 3 (Fig. 1a). The base change from G to A causes an amino acid change from aspartic acid to asparagine (Fig. 1b).

We designated this allele as KBS. Compared to the WT sequence, the CG sequence had a 159 bp deletion in the coding region of exon 3 from bp 534 to 693, and a point mutation in the coding region of exon 2 at bp 302 (Fig. 1a). This deletion was also reported by Dou et al. (2018). The point mutation of G to C causes a non-synonymous substitution resulting in an amino acid change of alanine to threonine (Fig. 1b). This point mutation was not present in the CG reference genome. The 159 bp deletion present in CG sequenced in the present study and the CG reference genome was designated as the DEL allele.

ClSUN-1 and ClSUN-2 KASP assays differentiate allelic variation

KASP assays were developed based on the sequencing information for the different polymorphisms. KASP assay ClSUN-1 was designed to distinguish the WT and KBS alleles from the 159 bp deletion found in the

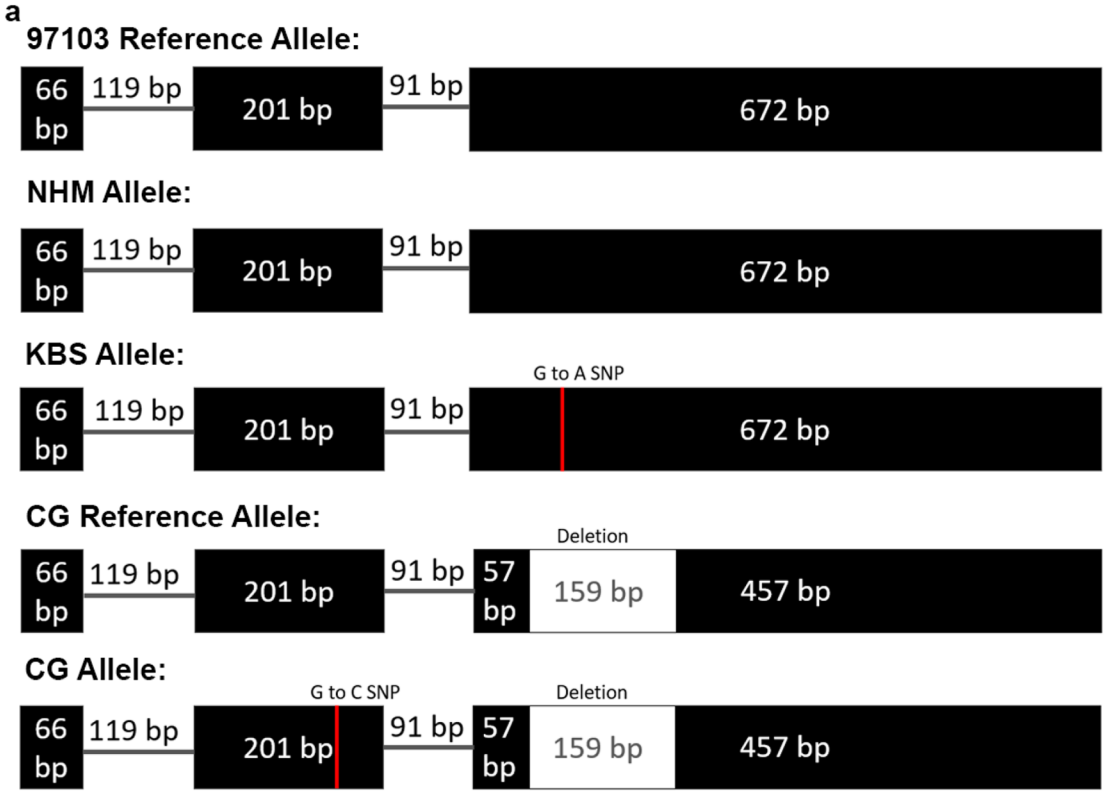

Fig. 1 (**a**) NHM, CG, and KBS alleles of *ClSUN25-26-27a* (*Cla011257)* compared to the 97103 and CG reference genome alleles. (**b**) Amino acid alignments of NHM, CG, and KBS alleles

DEL allele (Fig. S1a). Assay ClSUN-2 was developed to distinguish between the KBS allele point mutation and the WT allele (Fig. S1b). No amplification of the DEL allele will occur with this assay, since the labeled primers bind within the deletion (Fig. S1b).

Multiple assays were designed to genotype the SNP found in Exon 2 (302 bp) in CG. However, none of the assays were able to differentiate between the genotypes due to non-amplification or monomorphic results. A BLAST search of the CG genome revealed a partial duplication of the coding sequences of *ClSUN25-26-27a* on chromosome 5 from bp 4,747,878 to 4,747,999 that may compromise specificity of the primers designed in this region.

2016 and 2017 field fruit shape index is associated with *ClSUN25-26-27a* alleles

In order to investigate the effects of the different alleles, 10 cultivars/lines were genotyped for the WT, KBS, and DEL alleles and phenotyped for fruit traits. The cultivars/lines with the WT/WT genotype (AUP, CS, Egusi, MICK, NHM, and UGA147) had an ovary fruit shape index (oFSI) of 1.17 to 1.56 and FSI of 1.02 to 1.24 (Table S3). The only cultivar with the KBS/KBS genotype was KBS which had an oFSI of 2.16 and FSI of 1.47. The cultivars with the DEL/DEL genotype (AS, CG and CALG) had the largest oFSI, ranging from 2.48 to 3.24 and FSI between 1.79 and 2.22. For all genotypes, the shape index was larger for ovaries than for mature fruit, indicating that elongation happens primarily pre-anthesis.

The length of WT/WT genotype ovaries was 1.03 to 1.76 cm long, with mature fruit spanning from 13.3 to 25.71 cm. KBS/KBS ovaries had an oFL of 1.94 cm and FL of 28.34 cm. The ovary length of the DEL/DEL cultivars ranged from 1.89 to 2.59 cm, while the FL was between 35.6 and 43.1 cm (Table S3).

The diameter of the WT/WT ovaries ranged from 0.80 to 1.22 cm wide, with mature WT/WT fruit measured from 11.58 to 23.29 cm wide. The KBS/KBS ovaries were 0.90 cm wide, and mature fruit were 19.23 cm wide. The diameter of the DEL/DEL ovaries varied from 0.78 to 0.98 cm, with mature DEL/DEL fruit from 19.30 to 21.04 cm wide (Table S3). No data was collected for KBS and NHM genotypes in 2016 due to disease pressure.

oFL is weakly correlated (Table S4) to oFD ($r = 0.28$), while oFSI was strongly positively correlated to ovary length ($r = 0.78$) and negatively correlated to diameter ($r = -0.36$). There was a significant positive correlation between FL and FD ($r = 0.50$). Mature FSI was strongly positively correlated with FL ($r = 0.89$), but not significantly correlated with fruit diameter ($r = 0.06$). The correlations between ovary and mature fruit attributes were also examined (Table S4) to better understand the mechanisms of fruit development. Ovary and fruit length ($r = 0.69$) and ovary and fruit shape ($r = 0.91$) were strongly correlated, while oFD and FD were not significantly correlated ($r = 0.14$).

Ovary and fruit shape indexes were significantly different ($p < 0.0001$) based on pooled genotypes (Fig. S3). Average oFSI and FSI were lowest for the WT/WT plants (oFSI = 1.40; FSI = 1.08) while KBS/KBS plants (oFSI = 2.15; FSI = 1.47) were intermediate. oFSI and FSI were highest in the DEL/DEL genotypes (oFSI = 2.77; FSI = 1.96; Fig. S3a). The same pattern was observed for oFL and FL, with WT/WT genotypes being the shortest (oFL = 1.45 cm; FL = 20.70 cm), KBS/KBS genotypes being intermediate (oFL = 1.94 cm; FL = 28.34 cm) and DEL/DEL genotypes being the longest (oFL = 2.36 cm; FL = 38.91 cm; Fig. S3b).

The diameter of the ovaries did not follow the same pattern. The WT/WT genotypes had the widest ovaries (oFD = 1.04 cm), and DEL/DEL genotypes had the narrowest ovaries (oFD = 0.85 cm). The WT/WT and DEL/DEL ovaries were statistically different from one another; however, KBS/KBS ovaries (oFD = 0.90 cm) were not statistically different from either the DEL/DEL or WT/WT phenotypes ($p < 0.0001$; Fig. S3c). FD was not significantly different for any of the genotypes (WT/WT = 19.12 cm; KBS/KBS = 19.23 cm; DEL/DEL = 19.94 cm; Fig. S3c).

2018 parental panel fruit shape index is associated with ClSUN25-26-27a alleles

Digital phenotyping methods were employed in 2018 to better quantify the effect of the allele on the parents of validation populations (Fig. 2). In addition to providing precise measurements for the FL, FD, and FSI traits, it also allowed for quantification of the end shape of the watermelons by calculating the DEA and the PEA. The WT/WT cultivars in the parental panel were JS, NHM, and SB, which displayed the shortest and roundest ovaries and fruits on average (Fig. 2; Table S5). Fiesta and KBS had the KBS/KBS genotype, with ovaries and

mature fruit measurements being intermediate between WT/WT and DEL/DEL genotypes. The CG and CALG parental cultivars (DEL/DEL) showed the most elongated ovaries and fruits with the largest oFL and FSI. FSI and oFSI were significantly different based on genotype ($p < 0.0001$). The WT/WT cultivars were rounder (oFSI = 1.42; FSI = 1.07), followed by KBS/KBS cultivars (oFSI = 2.13; FSI = 1.60), with DEL/DEL cultivars being the longest (oFSI = 2.79; FSI = 1.98; Fig. 3a). As observed in the 2016 and 2017 data, the oFSI for genotypes were larger than the FSI (Fig. 2; Table S5).

Fruit and ovary lengths followed the same significant trend ($p < 0.0001$) as shape. Length was shortest for WT/WT genotypes (oFL = 1.15 cm; FL = 22.60 cm), KBS/KBS genotypes had intermediate lengths (oFL = 1.76 cm; FL = 36.47 cm), and the DEL/DEL genotypes had the longest ovaries and fruits (oFL = 2.08 cm; FL = 47.26 cm) (Fig. 3b).

KBS/KBS genotypes had the widest ovaries at oFD = 0.83 cm, significantly different ($p = 0.0352$) from the DEL/DEL genotypes, which had the narrowest ovaries (oFD = 0.76 cm) (Fig. 3c). There was no significant difference between WT/WT ovary diameter and the other genotypes. There was also no significant difference among any of the genotypes for fruit diameter (Fig. 3c).

The DEA and PEA of ovaries and fruit were significantly different ($p < 0.0001$) for all three genotypes. DEA for WT/WT genotypes was widest (oDEA = 134.7°; DEA = 152.4°), while KBS/KBS genotype DEA (oDEA = 93.8°; DEA = 121.0°) were intermediate, and DEL/DEL was the narrowest (oDEA = 68.3°; DEA = 100.4°) (Fig. 3d). PEA followed the same pattern ($p < 0.0001$) with WT/WT genotypes having the widest angle (oPEA = 138.7°; PEA = 153.9°), DEL/DEL genotypes the narrowest (oPEA = 68.8°; PEA = 108.3°), and KBS/KBS intermediate (oPEA = 94.3°; PEA = 124.2°) (Fig. 3c and d).

Allelic association verified in validation populations

To test the usefulness of the ClSUN-1 and ClSUN-2 KASP assays for marker-assisted selection of fruit length and fruit shape in multiple genetic backgrounds, the assays were tested on five segregating validation populations. The JS × Fiesta ($n = 42$) and KBS × NHM ($n = 136$) populations were segregating for the WT and KBS alleles and were genotyped using the ClSUN-2 assay. There were significant differences among the FSIs of the three genotypes in both populations ($p < 0.0001$). In both populations, the WT/WT

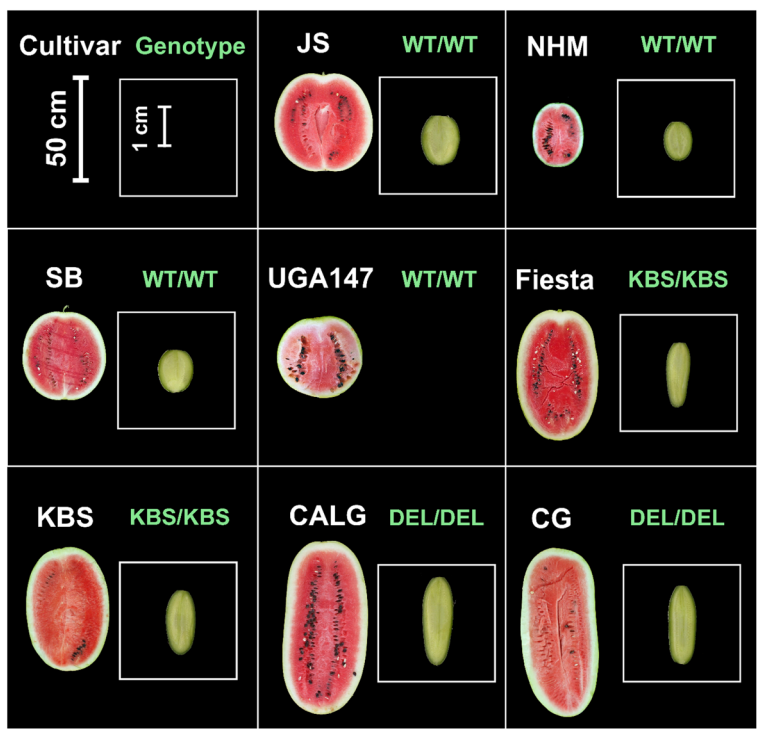

Fig. 2 Representative parental phenotypes and genotypes of validation population parents grown in 2018. UGA147 was not phenotyped in 2018, so no digital data was collected. A picture of UGA147 is provided for visual comparison only

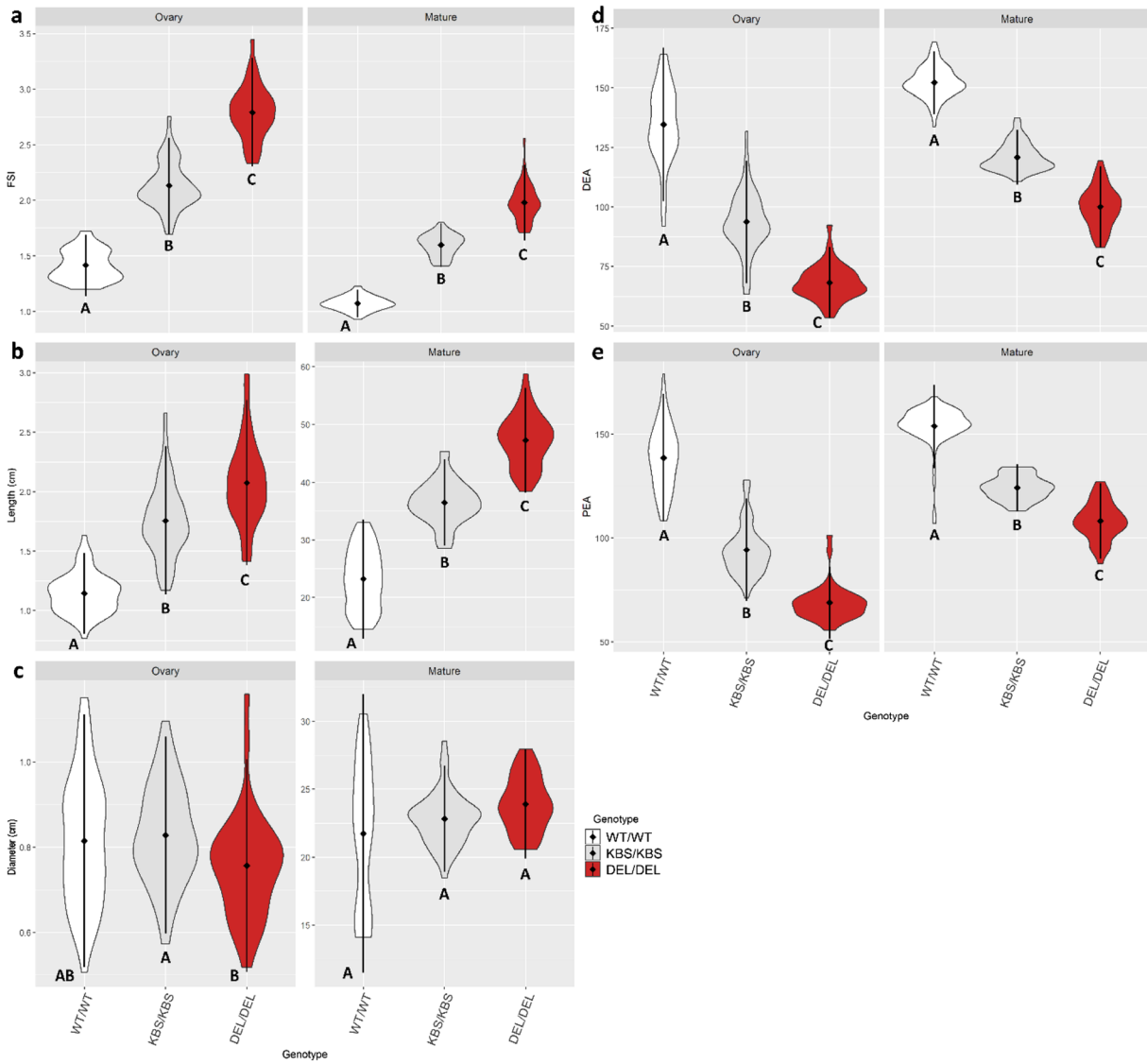

Fig. 3 Effect of *ClSUN25-26-27a* alleles on (**a**) ovary and mature FSI, (**b**) ovary and mature fruit length, (**c**) ovary and mature fruit diameter, (**d**) ovary and mature DEA, and (**e**) ovary and mature PEA. Data taken from 2018 parental cultivar panel show the distributions of the different genotypes. Black diamonds and lines show the means and standard deviations, respectively. Different letters show significant differences among genotypes within maturity stages ($p < 0.0001$)

genotypes had the roundest fruit with the lowest FSI. The WT/KBS heterozygotes had intermediate FSIs and KBS/KBS homozygotes had the most elongated fruit with the highest FSIs (Fig. 4a).

In the KBS × NHM population, the WT/KBS and KBS/KBS genotypes were not significantly different from one another for FL, but the WT/WT homozygotes were significantly shorter ($p < 0.0001$; Fig. 4b) than the other genotypes. The FL of the JS × Fiesta population was not significantly different ($p = 0.15$) among the genotypes.

The ClSUN-1 assay was used to genotype the populations segregating for the WT and DEL alleles: NHM × CALG ($n = 52$), CG × UGA147 ($n = 115$), and SB × CALG ($n = 125$). The FSI of every population was significantly associated with the genotype, with all genotypes being distinct within populations ($p < 0.0001$). In all populations, FSI was lowest for WT/WT individuals, intermediate in the WT/DEL heterozygotes, and highest in DEL/DEL homozygotes (Fig. 4a).

The lengths for each genotype of the NHM × CALG population were statistically different from each other

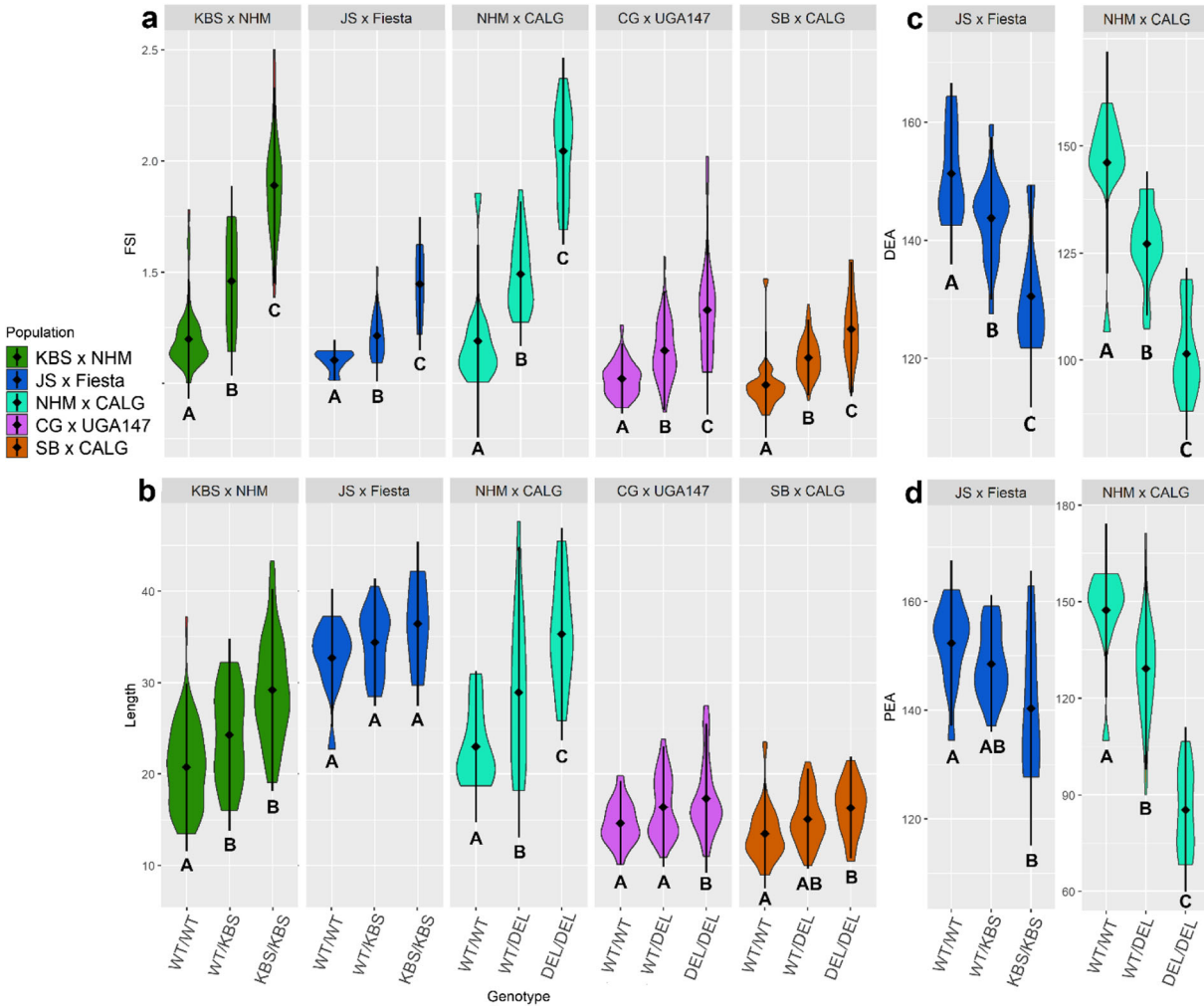

Fig. 4 Effect of *ClSUN25-26-27a* alleles on (**a**) FSI, (**b**) length (cm), (**c**) DEA, and (**d**) PEA in segregating validation populations. Violin plots are separated by population. Different letters show significant differences among genotypes within populations and traits ($p < 0.05$). Black diamonds and lines show the means and standard deviations, respectively. The NHM × CALG ($n = 52$) and JS × Fiesta ($n = 42$) populations were phenotyped using TomatoAnalyzer 4.0, while the CG × UGA147 ($n = 115$) and SB × CALG ($n = 125$) populations were phenotyped in the greenhouse in 2015 and 2016. KBS × NHM ($n = 136$) were previously phenotyped by Sandlin et al. (2012)

($p < 0.0001$). In the CG × UGA147 and SB × CALG populations, the WT/WT and DEL/DEL genotypes were statistically significantly different ($p < 0.0001$), but the WT/DEL heterozygotes in the CG × UGA147 population were not statistically different from the WT/WT genotypes, while in the SB × CALG populations, the heterozygotes were not statistically different from either of the homozygotes (Fig. 4b).

The DEA and PEA were only measured in the JS × Fiesta and NHM × CALG F_2 populations that were digitally phenotyped in 2018. The DEA for the JS × Fiesta F_2 genotypes were significantly different from one another ($p < 0.0001$), with the WT/WT homozygotes having the widest DEA (151.7°). The heterozygotes had a narrower DEA at 144.0°, and KBS/KBS DEA was the narrowest with an average DEA of 130.6° (Fig. 4c). The PEA for this population showed significant differences ($p = 0.0087$) between the WT/WT (PEA = 152.4°) and KBS/KBS genotypes (PEA = 140.4°), but the heterozygote genotypes (PEA = 148.6°) were not significantly different from the homozygotes (Fig. 4d).

Similar results were observed in the NHM × CALG population segregating for the DEL and WT alleles. The

WT/WT had the widest DEA (146.1°), the WT/DEL heterozygote had an intermediary DEA of 127.2°, and the DEL/DEL homozygotes had the narrowest DEA (101.5°; Fig. 4c). The PEA for the NHM × CALG population was also significantly different for all genotypes ($p < 0.0001$). The WT/WT genotypes had the highest PEA (147.4°), the WT/DEL genotypes averaged a PEA of 129.3°, and the DEL/DEL genotypes had the narrowest PEA at 85.4° (Fig. 4d).

Discussion

ClSUN25-26-27a is a candidate gene for fruit shape in watermelon (Dou et al. 2018; Jin et al. 2018; Kim et al. 2015; Sandlin et al. 2012). *ClSUN25-26-27a* is a member of the *SUN* family, which has been shown to control fruit length and shape in other crops, like tomato, melon, and cucumber (Monforte et al. 2014; Pan et al. 2017; Perpiñá et al. 2016; Wu et al. 2011). The duplication of the tomato *SUN* gene leads to increased expression with an associated increase in cell division in the proximal and distal directions before anthesis (Wu et al. 2011). In melon, *CmSUN2* (*CmSUN-14)* is a *SUN* homolog believed to help determine fruit elongation and radial growth during fruit development (Monforte et al. 2014; Perpiñá et al. 2016). *CsSUN25-26-27a* (*CsGy1G026840.1*, *Csa1G575000*; Pan et al. 2019) is a candidate gene proposed to change fruit size, fruit length, and fruit diameter in cucumber (Pan et al. 2017). Based on the evidence supporting the hypothesis that this gene family influences fruit length and shape in a range of crops, the effect of this homolog on fruit shape was investigated in watermelon.

It is interesting that in tomato, the SUN-associated fruit elongation is associated with a duplication event that leads to higher expression of the gene (Xiao et al. 2008; Xiao et al. 2009) while in watermelon a deletion in exon 3 of *ClSUN25-26-27a* is associated with higher expression during ovary formation (Dou et al. 2018) and ovary and fruit elongation. In contrast, a deletion in the first exon of *CsSUN25-26-27a* in cucumber is associated with rounder fruit and reduced expression of the gene (Pan et al. 2017). Further research is needed on the allelic variation and expression patterns of *SUN* family genes in Cucurbitaceae to elucidate their role in fruit shape in this economically important crop family.

We discovered a novel allele (KBS) with a non-synonymous point mutation in the same region as the deletion previously found in the DEL allele in Charleston Gray (Dou et al. 2018). Compared to the WT allele, the SNP in the KBS allele results in a functional mutation, changing aspartic acid to asparagine. This point mutation is associated with intermediary phenotypes as compared to DEL and WT. Future research will include gene expression studies to determine whether the phenotype is associated with changes in expression.

In addition to the previously described 159 bp deletion (Dou et al. 2018), we also discovered a non-synonymous functional point mutation in Charleston Gray. This mutation is not present in the Charleston Gray reference genome (Wu et al. 2019). It is possible that this SNP was not detected in the CG reference genome due to a sequencing error, but this SNP has also not been detected in any other genotype sequenced in this study or other studies to date. Since we sequenced two different CG plants and both had the SNP, it seems unlikely that it is a sequencing error in the present study. Most likely it is a mutation in the specific Charleston Gray seed we obtained. Further comparison of the phenotypes of the two CG alleles might shed more light on the effect, if any, of this mutation on fruit shape.

A higher-quality draft (version 2) of the 97103 genome was recently published by Guo et al. (2019). The *Cla011257* gene in version 1 corresponds to the *Cla97C03G066390* gene in version 2. Exon 1 in *Cla011257* is only 66 bp, whereas exon 1 in *Cla97C03G066390* includes an additional 339 bp upstream of the version 1 start codon, resulting in 135 amino acids rather than 22 (v. 1) in the first codon. There are no sequence differences in exons 2 or 3 between the two versions. The sequencing in our current study did not include the additional 339 bp upstream region; therefor, the present study does not elucidate any potential effect of sequence differences in this region on fruit shape traits.

Two KASP assays were developed to differentiate the WT and DEL (ClSUN-1) and WT and KBS (ClSUN-2) alleles. In the preliminary field studies from 2016 and 2017, a trend emerged that was evident throughout the rest of our studies. The WT allele resulted in the shortest ovaries and fruit, with the lowest FSI, the DEL allele had the longest ovaries and fruit with the highest FSI, and the KBS allele was associated with intermediary phenotypes.

The high correlations between oFL and FL and oFSI and FSI confirm that watermelon fruit length and shape are predominantly determined pre-anthesis. This is

consistent with what is found in melon (Périn et al. 2002) and cucumber (Pan et al. 2017). *SUN*, along with other loci, has been found to control ovary and fruit elongation in tomato both pre-anthesis and post-anthesis (Wu et al. 2015). Our results indicate that in watermelon, *ClSUN25-26-27a* is responsible for controlling ovary and fruit length and shape. *ClSUN25-26-27a* primarily works pre-anthesis by creating a framework of cells that provide a basis for the mature fruit shape throughout fruit development. The low correlations of ovary diameter with fruit diameter suggest that diameter is controlled post-anthesis. The lack of association between the *ClSUN25-26-27a* genotypes and fruit diameter also confirms that diameter is controlled by a different locus. The variation in oFSI and FSI between cultivars with the same *ClSUN25-26-27a* genotypes also confirms that other loci affect shape index in watermelon.

We implemented digital phenotyping in 2018 in order to better quantify the effect of the *ClSUN25-26-27a* alleles. In addition to confirming the association of the *ClSUN25-26-27a* genotypes with fruit and ovary length and shape, we also observed significant associations with DEA and PEA. These measurements approximate the end shape of the ovary and fruit. In tomato fruit, the *SUN* allele (gain of function duplication) caused a more acute distal tip shape than the wild-type allele (Wu et al. 2015). While diameter, length, and shape index are useful to describe fruit characteristics, our results underscore the usefulness of digital phenotyping to describe additional shape traits.

The usefulness of the KASP assays for MAS for the three *ClSUN25-26-27a* alleles were validated using five segregating populations. The FSI of all validation populations showed highly significant associations with the genotypes. The validation populations confirm that KASP assays ClSUN-1 and ClSUN-2 can effectively be used to select for FSI, DEA, and PEA in multiple genotypic backgrounds. Only two validation populations (JS × Fiesta and NHM × CALG) were digitally phenotyped, allowing for measurement of the DEA and PEA. Significance for DEA and PEA was found in both populations. While not as important as length and FSI, it is important to note that selection using ClSUN-1 and ClSUN-2 may impact more fruit shape traits than just the FSI and length.

The effectiveness of the KASP assays for the selection of length is not as pronounced as for FSI. Significant differences in fruit length were observed in four of the validation populations (KBS × NHM, NHM × CALG, CG × UGA147, and SB × CALG), but not in the JS × Fiesta population. The lack of significant differences could be due to the small population size (n = 42) or additional loci contributing to fruit characteristics in this population.

Conclusion

Fruit shape is an important quality trait for many crops, including watermelon. The discovery of the novel KBS allele of *ClSUN25-26-27a* brings the total number of known alleles to five. The full extent of the phenotypic effects of these alleles remains to be determined.

The two KASP assays (ClSUN-1 and ClSUN-2) enable marker-assisted selection for fruit shape index and end angles in different genetic backgrounds. This will enable MAS of fruit shape from seed or seedlings, without the need to grow plants to the flowering or fruiting stage.

Funding information This work was in part supported by the United States Department of Agriculture (GEO00759) and the United States Department of Agriculture Specialty Crop Research Initiative (Award No. 2014-51181-22471) and the UGA Department of Horticulture.

Compliance with ethical standards

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

References

Brewer, MT, Lang, L, Fujimura, K, Dujmovic, N, Gray, S, van der Knaap, E (2006) Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other plant species. Plant Physiology 141 (1):15-25

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. https://doi.org/10.1016/j.scienta.2016.01.004

Dou J, Zhao S, Lu X, He N, Zhang L, Ali A, Kuang H, Liu W (2018) Genetic mapping reveals a candidate gene (*ClFS1*) for fruit shape in watermelon (*Citrullus lanatus* L.). Theor Appl Genet 131:947–958

Fall LA, Clevenger J, McGregor C (2018) Assay development and marker validation for marker assisted selection of *Fusarium oxysporum* f. sp. *niveum* race 1 in watermelon. Mol Breed 38: 130

FAOSTAT (2016) FAOSTAT statistical database

Gonzalo, MJ, Brewer, MT, Anderson, C, Sullivan, D, Gray, S, and van der Knaap, E (2009) Tomato fruit shape analysis using morphometric and morphology attributes implemented in Tomato Analyzer software program. J. Am. Soc. of Hort. Sci. 134: 77-87

Guo S, Zhang J, Sun H, Salse J, Lucas WJ, Zhang H, Zheng Y, Mao L, Ren Y, Wang Z (2013) The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. Nat Genet 45:51

Guo S, Zhao S, Sun H, Wang X, Wu S, Lin T, Ren Y, Gao L, Deng Y, Zhang J, Lu X, Zhang H, Shang J, Gong G, Wen C, He N, Tian S, Li M, Liu J, Wang Y, Zhu Y, Jarret R, Levi A, Zhang X, Huang S, Fei Z, Liu W, Xu Y (2019) Resequencing of 414 cultivated and wild watermelon accessions identifies selection for fruit quality traits. Nat Genet 51:1616–1623

Jin B, Kim J, Jung J, Kim D, Park Y (2018) Characterization of IQ domain gene homologs as common candidate genes for elongated fruit shape in cucurbits. Hortic Sci Technol 36: 85–97

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649

Kim KH, Hwang JH, Han DY, Park M, Kim S, Choi D, Kim Y, Lee GP, Kim ST, Park YH (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 (*Citrullus lanatus* var. *lanatus*). PLoS One 10:e0145665. https://doi.org/10.1371/journal.pone.0145665

King Z, Serrano J, Roger Boerma H, Li Z (2014) Non-toxic and efficient DNA extractions for soybean leaf and seed chips for high-throughput and large-scale genotyping. Biotechnol Lett 36:1875–1879. https://doi.org/10.1007/s10529-014-1548-8

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. And Nakai) constructed using whole-genome re-sequencing data. Breed Sci 66:244–259

Loy JB (2005) Fruit size in melon in monoecious and andromonoecious isolines. Cucurbit Genet Coop Rpt 28-29: 12-13

Meru G, McGregor C (2016a) Genotyping by sequencing for SNP discovery and genetic mapping of resistance to race 1 of *Fusarium oxysporum* in watermelon. Sci Hortic 209:31–40

Meru G, McGregor CE (2016b) A genetic locus associated with resistance to *Fusarium oxysporum* f. sp. *niveum* race 2 in *Citrullus lanatus*-type watermelon. J Am Soc Hortic Sci 141: 617–622

Monforte AJ, Diaz A, Caño-Delgado A, Van Der Knaap E (2014) The genetic basis of fruit morphology in horticultural crops: lessons from tomato and melon. J Exp Bot 65:4625–4637

Pan Y, Liang X, Gao M, Liu H, Meng H, Weng Y, Cheng Z (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

Pan Y, Wang Y, McGregor C, Liu S, Luan F, Guo M, Weng Y (2019) Genetic architecture of fruit size and shape variation in cucurbits; a comparative perspective. Theor Appl Genet 133:1–21. https://doi.org/10.1007/s00122-019-03481-3

Périn C, Hagen L, Giovinazzo N, Besombes D, Dogimont C, Pitrat M (2002) Genetic control of fruit shape acts prior to anthesis in melon (*Cucumis melo* L.). Mol Gen Genomics 266:933–941

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

Poole CF, Grimball PC (1945) Interaction of sex, shape, and weight genes in watermelon. J Agric Res 71:533

Rosa J (1928) The inheritance of flower types in *Cucumis* and *Citrullus*. Hilgardia 3:233–250

Sandlin K, Prothro J, Heesacker A, Khalilian N, Okashah R, Xiang W, Bachlava E, Caldwell DG, Taylor CA, Seymour DK (2012) Comparative mapping in watermelon [*Citrullus lanatus* (Thunb.) Matsum. Et Nakai]. Theor Appl Genet 125:1603–1618

Stepansky A, Kovalski I, Perl-Treves R (1999) Intraspecific classification of melons (*Cucumis melo* L.) in view of their phenotypic and molecular variation. Plant Syst Evol 217: 313–332

Tanaka T, Wimol S, Mizutani T (1995) Inheritance of fruit shape and seed size of watermelon. J Jpn Soc Hortic Sci 64:543–548

Weetman LM (1937) Inheritance and correlation of shape, size and color in the watermelon, *Citrullus vulgaris* Schrad. Iowa Agric Home Econ 20:1

Weng Y, Colle M, Wang Y, Yang L, 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. https://doi.org/10.1104/pp.111.181065

Wu S, Clevenger JP, Sun L, Visa S, Kamiya Y, Jikumaru Y, Blakeslee J, van der Knaap E (2015) The control of tomato fruit elongation orchestrated by sun, *ovate* and *fs8.1* in a wild relative of tomato. Plant Sci 238:95–104

Wu S, Wang X, Reddy U, Sun H, Bao K, Gao L, Mao L, Patel T, Ortiz C, Abburi VL (2019) Genome of 'Charleston Gray', the principal American watermelon cultivar, and genetic characterization of 1,365 accessions in the US National Plant Germplasm System watermelon collection. Plant Biotechnol J:1-13 doi:https://doi.org/10.1111/pbi.13136

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, Hsu 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