Theoretical and Applied Genetics (2026) 139:87
https://doi.org/10.1007/s00122-026-05177-x

ORIGINAL ARTICLE

Fruit weight regulation by a paralog of *Cell Size Regulator* (*CSR*) in tomato and other crops

Qian Feng[1] · Lara Pereira[2] · Manoj Sapkota[1,5] · Krishna Sai Karnatam[3] · Yanbing Wang[2] · Padma Nimmakayala[3] · Umesh K. Reddy[3] · Esther van der Knaap[1,2,4]

Received: 19 November 2025 / Accepted: 1 February 2026 / Published online: 5 March 2026
© The Author(s) 2026

Abstract

Key message **A paralog of *Cell Size Regulator* (*CSR*), *CSR-like1*, underlies the novel *fw6.2* QTL in tomato. The gene and locus regulate fruit weight by increasing pericarp cell size and its function on fruit weight appears to be conserved in other crops.**

Abstract Fruit weight is a quantitative trait that was under strong selection during the domestication of fruit and vegetable crops such as tomato (*Solanum lycopersicum*). While numerous fruit weight QTLs have been identified, only three tomato fruit weight genes have been cloned. In this study, we utilized a genetically diverse tomato panel, the Varitome collection, to identify additional genetic loci that control fruit weight. We mapped and fine mapped two fruit weight QTLs on chromosome 6, *fw6.1* and *fw6.2*, by using Genome Wide Association studies (GWAS) and linkage mapping in bi-parental populations. We identified a member of the *Cell Size Regulator* family, *CSR-like1*, as the likely candidate underlying *fw6.2*. The near isogenic lines (NILs) carrying the derived allele of *fw6.2* produced heavier fruits with larger fruit pericarp cells than lines with wildtype (WT) allele. Transgenic downregulation of *CSR-like1* led to a decrease in fruit weight and pericarp cells, supporting the role of this gene at the *fw6.2* locus. The haplotype analysis implied that the *CSR-like1*-Derived (*CSR-like1*-D) allele was selected in the transition from the fully wild *S. pimpinellifolium* to the earliest *S. lycopersicum cerasiforme* accessions. Four single nucleotide polymorphisms (SNPs) were identified in the regulatory region of *CSR-like1* that were conserved in the accessions carrying *CSR-like1*-WT and were significantly associated with lower fruit weight and pericarp cell size at the locus. Moreover, a pepper GWAS identified a *CSR-like1* ortholog that was associated with fruit weight. Together, our findings established *CSR-like1* as a novel fruit weight gene likely conserved in other crops in the Solanaceae family.

Introduction

Plant domestication has led to the transformation from wild species into high-yielding and nutritionally valuable crops. Genetic modifications through repeated selection and cultivation have reshaped the plant morphology, physiology, and development to suit human needs (Doebley et al. 2006). Increases in fruit or grain size is one of the key traits that differentiate the domesticated crops from their wild progenitors, collectively referred to as "domestication syndrome" (Doebley et al. 2006; Gross and Olsen 2010). For tomato (*Solanum lycopersicum*), huge variation in fruit size and shape is evident in the cherry-sized type *S. lycopersicum* var *cerasiforme* (SLC) and fully domesticated type *S. lycopersicum* var. *lycopersicum* (SLL) compared to its wild relative, *S. pimpinellifolium* (SP) (Tanksley 2004).

Several studies have revealed insights regarding the complex domestication history of tomato resulting in different

Communicated by Sanwen Huang.

✉ Esther van der Knaap
esthervanderknaap@uga.edu

[1] Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Athens, GA, USA

[2] Center for Applied Genetic Technologies, University of Georgia, Athens, GA, USA

[3] Gus R. Douglass Institute and Department of Biology, West Virginia State University, Institute, WV, USA

[4] Department of Horticulture, University of Georgia, Athens, GA, USA

[5] Department of Horticulture, University of Kentucky, Lexington, KY, USA

models. One model illustrated a two-step process (Razifard et al. 2020): SP first originated in the Andean region of South America and gave rise to the semi-domesticated SLC. The more domesticated South American SLC groups then migrated northwards when some domestication traits, including fruit size, were lost. A second round of selection took place in Mesoamerica before the emergence of the fully domesticated SLL. A recent study suggested a more complex trajectory (Blanca et al. 2022): SP expanded much further northwards into Mesoamerica where it evolved into wild SLC progenitor species. Some SLCs subsequently migrated back to South America and admixed with SP in Ecuador and Peru. Ecuadorian and Peruvian SLC then underwent another northward migration to further evolve into fully domesticated SLL in Mexico.

Fruit weight is a character controlled by multiple quantitative trait loci (QTL) (Paran and van der Knaap 2007). Though over twenty QTL have been mapped in tomato (Causse et al. 2002; Grandillo et al. 1999; Illa-Berenguer et al. 2015; Pereira et al. 2021b), only three fruit weight genes have been cloned: *fw2.2*, *fw3.2* and *fw11.3*. *Fw2.2* underlies *Cell Number Regulator* (*CNR*) (Frary et al. 2000), a negative cell division regulator with a change in transcription dynamic that is associated with lower cell division rate (Cong et al. 2002). The putative casual mutation is thought to lie in the promoter region since the coding region is highly conserved in small-fruit allele and large-fruit allele (Frary et al. 2000). *Fw3.2* (*SlKLUH*) encodes a P450 enzyme of the CYP78A subfamily that regulates cell number in the fruit (Chakrabarti et al. 2013). A tandem duplication of around 50 kb containing *SlKLUH* leads to higher expression of *SlKLUH* and an increase in fruit weight. (Alonge et al. 2020). *Fw11.3* encodes *Cell Size Regulator* (*CSR*) and is the only fruit weight gene known to affect cell size (Mu et al. 2017). The derived allele (*CSR-D*) has a 3′ 1.4 kb deletion that affects the coding region as well as higher expression throughout fruit development compared to the wildtype allele (*CSR-WT*) (Mu et al. 2017). In addition to the fruit weight loci, two locule number loci, *lc* (Muños et al. 2011) and *fas* (Cong et al. 2008; Xu et al. 2015), can affect the fruit weight in tomato through regulating the number of locules in the fruit.

Here, we report the fine-mapping and identification of two linked fruit weight QTL in tomato, *fw6.1* and *fw6.2*, and the cloning and characterization of a new fruit weight gene *CSR-like1* at the second locus. Similar to its paralog *CSR*, *CSR-like1* also controls the fruit mass through regulating the pericarp cell size. Yet different from *CSR*, *CSR-like1* underwent selection in rather the early stages of domestication and was much higher expressed in fruit tissues. Transgenic plants that downregulated *CSR-like1* expression showed decreased fruit weight that developed late during fruit growth, as well as effects on other traits such as ripening. We also present genetic evidence on the function of a pepper ortholog of *CSR-like1* on fruit weight. The findings imply a conserved function of *CSR-like1* in fruit size regulation in Solanaceae crops.

Materials and methods

Plant materials

The 166 Varitome collection accessions have been described previously (Mata-Nicolás et al. 2020; Razifard et al. 2020). The plants for the GWAS and mapping populations were grown at University of Georgia's Horticulture Research Farm (Athens, GA, USA), Vidalia Onion and Vegetable Research Center (Lyons, GA, USA), Georgia Mountain Research and Education Center (Blairsville, GA, USA) and University of Florida's North Florida Research and Education Center-Suwannee Valley (Live Oak, FL, USA). Fresh fruit weight of GWAS population was measured as described in Razifard et al. (2020). Three replicates were grown for each accession and approximately 20 fruits were weighed to calculate the average fruit weight for each replicate. Population development including crossing, generating F_1 and the subsequent recombinant screening and progeny testing populations were conducted in the University of Georgia (Athens, GA, USA) greenhouses where plants were grown under standard conditions. Phenotypic characterizations including cellular measurement, days to ripening and plant height measurement were performed in greenhouse as well.

Variant calling and genome-wide association study (GWAS)

Short-read sequencing data for the Varitome tomato accessions were retrieved from NCBI (BioProject PRJNA454805) and aligned to the SL4.0 reference genome. The raw reads were aligned and processed for Variant calling for SNPs and INDELs (insertions and deletions) using the pipeline as previously described (Pereira et al. 2021a). Variants were filtered to retain only biallelic sites with missingness $< 5\%$, minor allele frequency (MAF) $\geq 5\%$, and mapping quality ≥ 30 using bcftools (Danecek et al. 2021). After filtering, 659,002 INDELs were retained. For computational efficiency, SNPs were further thinned to one SNP per kilobase using VCFtools (Danecek et al. 2011), resulting in a final set of 609,034 SNPs. SNP genotypes were exported from VCF files using the R/vcfR package (Knaus and Grünwald 2017). INDELs were encoded as a presence/absence matrix and treated as pseudo-SNPs, with the reference allele representing absence and the alternate allele representing presence of the INDEL. Association mapping was performed using GWASpoly (https://jendelman.github.io/GWASp

oly/GWASpoly.html) (Rosyara et al. 2016) under a diploid setting (ploidy = 2). The raw fruit weight data is shown in Supplementary Dataset 1 and the normalized fruit weight data was used for association mapping. The diplo-additive model was used as the primary framework, where genotypes are encoded as 0/1/2 copies with heterozygotes intermediate between homozygotes. This model was chosen for its simplicity, statistical power, and biological interpretability. Although other models were tested, results were broadly consistent; therefore, only the diploid additive model is presented. Population structure was controlled using the first three principal components, and kinship was modeled with a leave-one-chromosome-out (LOCO) approach (Cheng et al. 2013). Genome-wide significance thresholds were established using the Bonferroni correction at $\alpha = 0.05$.

Analysis of orthologous *CSR-like1* gene underlying fruit weight variation in the *Capsicum* Genus

A total of 594 accessions of *Capsicum* spp., comprising 82 *C. annuum*, 35 *C. annuum var. glabriusculum*, 243 *C. baccatum*, and 234 *C. chinense*, were obtained from the USDA–ARS Germplasm Resources Information Network (GRIN), Plant Genetic Resources Conservation Unit (Griffin, GA, USA). These accessions represented a wide range of geographical origins and fruit weight variation. Plants were cultivated for the three seasons in a randomized block design with three replications, and phenotypic data were recorded from five plants per accession in each replication. Genomic DNA was extracted from young leaf tissue, and genotyping-by-sequencing (GBS) was performed following the protocol described by Elshire et al. (2011) using the Illumina HiSeq 2500 platform. The resulting sequence reads were aligned to the *Capsicum annuum* cv. CM334 v1.55 reference genome (Kim et al. 2014) (http://peppergenome.snu.ac.kr/). SNP calling and filtering were performed according to Reddy et al. (2025), retaining 45,987 high-quality SNPs with a minor allele frequency (MAF) ≥ 0.05 and call rate $\geq 90\%$. Population structure was assessed through principal component analysis (PCA) using GAPIT 3.0 (Wang and Zhang 2021), and admixture proportions were inferred with ADMIXTURE v1.3 (Alexander et al. 2015) for K = 1–10. Genome-wide association studies (GWAS) for fruit weight were conducted in GAPIT 3.0 (Wang and Zhang 2021) using both the mixed linear model (MLM) and the Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) model to ensure robustness of detection.

Biparental mapping population development, finemapping and progeny testing

Five F_2 biparental populations were generated to validate *fw6.1* and *fw6.2* (Supplementary Table S1). The populations were fixed for the other known fruit weight and shape genes. To fine map the *fw6.1* locus, recombinant plants with one recombination event between marker 18EP962 and 19EP629 (Supplementary Table S2) were evaluated for fruit weight in an F_4 population 20S156 (N = 202). Interval mapping results was plotted using R/qtl package (map.funciton = kosambi, model = normal) (Broman et al. 2003). Quantile normalization was performed for fruit weight data using qqnorm function in R. The logarithm of odds (LOD) threshold of $\alpha = 0.01$ was calculated using permutation tests with 1000 permutations. To fine map *fw6.1* and *fw6.2* separately, we generated F_6 and F_7 recombinant population from a cross between large-fruited BGV06232 and small-fruited BGV008225. Population 22S33 (N = 137) segregated for *fw6.1* and population 22S30 (N = 88) segregated for *fw6.2*. Family 22S33 was fixed at BGV006232 allele between marker 21EP44 and 18EP327 for *fw6.2*. Family 22S30 was fixed at BGV008225 allele between marker 18EP962 and 21EP41 for *fw6.1*. Nine recombinant plants from different generations were progeny tested. Pedigree information for fine mapping *fw6.1* and *fw6.2* is shown in Supplementary Fig. S1.

To genotype the plants, Kompetitive Allele Specific PCR (KASP) markers were designed using the Primer Express® Software version 3.0.1 (Applied Biosystems, Carlsbad, CA) as described in Topcu et al. (2021) and derived cleaved amplified polymorphic sequences (dCAPS) markers were designed using the tool indCAPS (Hodgens et al. 2017). Primer information is shown in Supplementary Table S2.

Knockout of *CSR-like1*

To utilize CRISPR-*Cas9* editing to create knockout mutations of *CSR-like1*, a single guide RNA was designed using CRISPR-P (http://crispr.hzau.edu.cn/cgi-bin/CRISPR/CRISPR) (Supplementary Table S2) and cloned into p201N vector (Addgene plasmid #59,175). The construct was transformed into BGV007931 (SLC) from the Varitome collection and a breeding line Fla.8059 (SLL) (Scott et al. 2008) using standard transformation procedures (Van Eck et al. 2019). The transformation yielded two independent T_0s in BGV007931, and three independent T_0s in Fla. 8059 background (Supplementary Table S3). Primers for genotyping the presence of transgene and the mutation are listed in Supplementary Table S2.

Downregulation of *CSR-like1*

One artificial microRNA (amiRNA) targeting the coding region of *CSR-like1* was designed using WMD3 Web MicroRNA Designer (http://wmd3.weigelworld.org/). Target search with maximum 5 mismatches of the selected amiRNA sequence (TCTTGAGTCGAGTTGCGTCAT) showed *CSR-like1* as the only target. The amiRNA sequence

and its pair amiRNA* (ATAACGCAACTCGTCTCAAGT) were cloned into the Arabidopsis miR319a precursor backbone. The amiRNA precursor was synthesized by Azenta Life Sciences (South Plainfield, NJ, USA) and cloned into pKYLX71 vector at the *XhoI* and *SacI* restriction enzyme sites. The construct was transformed into the large-fruited mapping parent BGV006232. Four T_0s were obtained and verified for the presence of transgene using primers 20EP672/20EP673 (Supplementary Table S2). The BC_1F_1 families 24S156, 25S158, 25S159 and 25S160 were used for phenotypic evaluations. Pedigree information is shown in Supplementary Fig. S2.

Phenotypic evaluations

The mapped interval of *fw6.2* spanned 374 kb between marker 21EP56 and 24EP416. We generated a *fw6.2* NIL family 24S155 in which *fw6.2*-D plants carry the large-fruited derived allele from BGV006232 and *fw6.2*-WT plants carry the wildtype allele from BGV008225. For *fw6.2* NIL family 24S155 and the amiRNA transgenic family 24S156, 6 to 10 ovaries per plant were collected and scanned using a flatbed scanner. Ovary size was measured using Fiji software (Schindelin et al. 2012). Flowers were hand-pollinated and tagged for additional fruit phenotyping. Per plant, five to 10 of the largest fruits at turning or red ripe stage were cut at the equatorial plane along the medio-lateral axis and scanned. The images were analyzed using Tomato Analyzer 4.0 (Rodríguez et al. 2010) to obtain the total cross-section area, pericarp area, pericarp area ratio, columella area and columella area ratio. Fruit weight at various developmental stages was also measured in grams: 3, 5, 10, 15, 20, 25 DPA with three to four replicates per plant; mature green and red ripe stage consisted of five to 10 replicates per plant. Pericarp thickness, cell layer and mesocarp cell size were measured following the previously established protocols (Mu et al. 2017). Mesocarp cell size was calculated as the mean of the five largest cells in a pericarp slice. The number of days between anthesis to red ripe was recorded for the amiRNA transgenic families 24S156, 25S158, 25S159 and 25S160. Plant height at 55 days after sowing was measured for transgenic families 25S158, 25S159 and 25S160.

Haplotype analysis of *CSR-like1*

The genetic diversity of *CSR-like1* in the Varitome collection was analyzed following the previously described protocol (Pereira et al. 2021a). Briefly, SNPs and INDELs were extracted from within *CSR-like1* as well as 3 kb upstream of the start site and 1 kb downstream of the termination site with VCFTools (Danecek et al. 2011). The locations and functions of the variants were annotated using SnpEff (Cingolani et al. 2012) with a reference SL4.0 tomato reference genome (https://solgenomics.net/). The haplotype heatmap was generated using the R package "pheatmap" (Kolde 2019), accompanied with the phylogeny of the accessions (Razifard et al. 2020). The fruit weight of each accession was classified into three groups: small (below 3 g), medium (3–20 g) and big (above 20 g). The *p*-values of pairwise comparison among haplotype clusters were conducted using the R package "emmeans" (Lenth 2022) and the Tukey test with a significance threshold at 0.05 was used. Single variant association with fruit weight and pericarp cell size was calculated by Kruskal–Wallis test and the *p*-values were adjusted by the Benjamini–Hochberg method to control the False Discovery Rate (FDR). The haplotype distribution of *CSR-like1* was visualized in a geographic map as described in Sapkota et al. (2023) using R packages "rnaturalearth" (https://CRAN.R-project.org/package=rnaturalearth) and "ggspatial" (https://CRAN.R-project.org/package=ggspatial). The latitude and longitude of collection sites were retrieved from Razifard et al. (2020).

The haplotypes of *CSR-like1* in pepper were identified using five GBS SNPs (S06_227195491, S06_227195529, S06_227195579, S06_227195589, and S06_227195619) and the pairwise comparison of fruit weight among haplotypes were performed using the same method as for tomato haplotype analysis.

Results

Identification and fine mapping of *fw6.1* and *fw6.2* in tomato

To identify novel genes regulating fruit weight in tomato, a GWAS analysis was performed using the Varitome collection which included 27 wild SP, 121 semi domesticated SLC and 18 ancestral landrace SLL accessions. A total of 15 fruit weight GWAS loci using SNPs and INDELs were identified in this study, which led to six unique fruit weight QTLs on chromosomes 2, 3, 6, 11, and 12, and another six QTLs that were scattered on chromosome 1 (Fig. 1a,b; Supplementary Table S4). The *GWAS_fw2.2* locus is likely associated with the known *fw2.2*, whereas *GWAS_fw11.1* is likely associated with the linked loci *fw11.3* and *fas*. *GWAS_fw3.2* is likely associated with *fw3.2*, however, the variant was 5 Mb from the known fruit weight gene. Among the novel GWAS loci, *GWAS_fw6.1* has two linked INDEL markers associated with fruit weight (ISL4.0ch06_44805731 and ISL4.0ch06_45074906) and was further characterized in this study.

To validate the newly identified *fw6.1* locus, we evaluated a total of five F_2 biparental populations using flanking markers approximately 1–2 Mb upstream and downstream of the significant GWAS loci. The parental selection

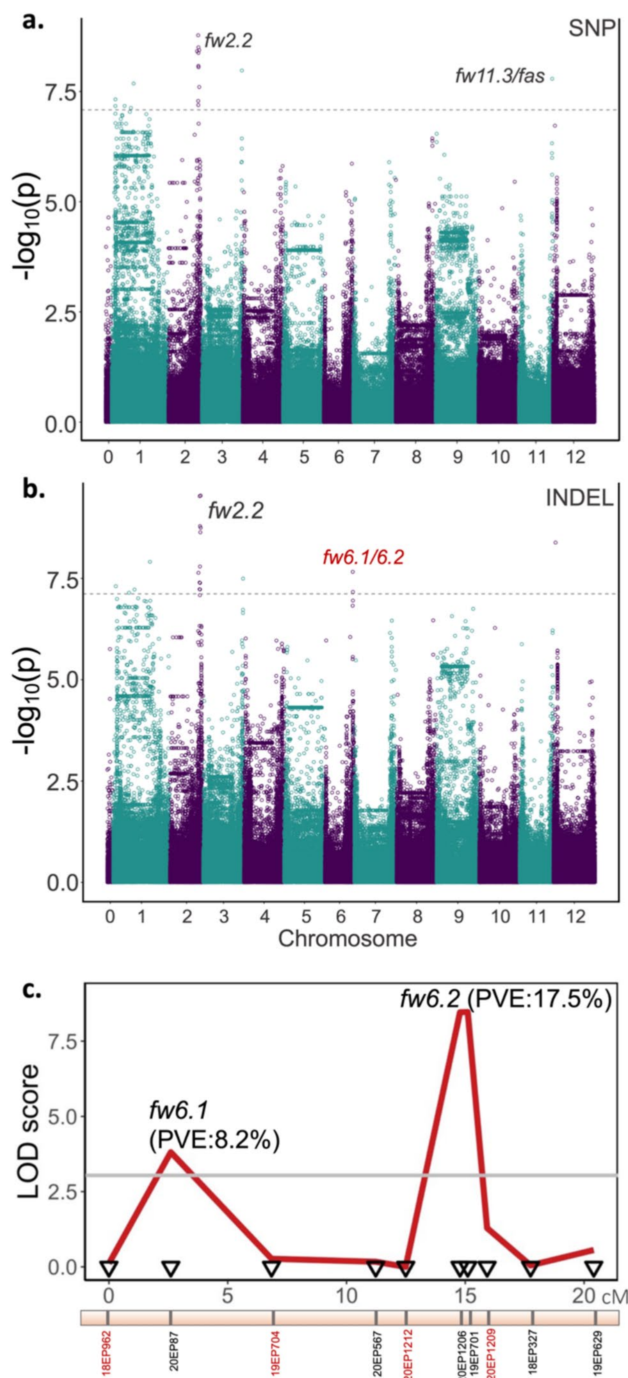

Fig. 1 Identification of *fw6.1* and *fw6.2*. **a** GWAS for fruit weight in the Varitome collection based on SNPs. **b** GWAS for fruit weight based on INDELs. The x-axis shows the chromosome number and y-axis shows the -log10(*p*-value) of the variants. The Bonferroni-corrected Genome-wide significance thresholds at α = 0.05 are indicated by the dotted lines. **c** Composite interval mapping of *fw6.1* and *fw6.2* in an F_4 population 20S156. The markers used to genotype recombinant plants in 20S156 are shown below the QTL map. The markers flanking the two intervals are highlighted in red. The LOD threshold of α = 0.01 is indicated in a gray line

of the F_2 populations was based on the predicted variants from the GWAS and the estimated effect of the QTL on fruit weight. For all F_2 populations, the two parents of each population were fixed at the known tomato fruit weight and fruit shape genes (Supplementary Table S1). *GWAS_fw6.1* was confirmed in the four F_2 populations that segregated at the *GWAS_fw6.1* locus, namely 18S33, 18S39,18S40 and 18S133 (Supplementary Table S5). The QTL showed a significant additive effect with a small, nonsignificant dominance deviation, indicating primarily additive gene action with partial dominance (Supplementary Table S5). The marker association was most significant in 18S39, and therefore this population was used to fine map the locus. Family 18S39 was derived from a cross between an Ecuadorian accession (BGV006232) and a Peruvian SLC accession (BGV008225), and the two parental accessions exhibited a fourfold difference in fruit weight (Supplementary Table S1).

To narrow the interval of *fw6.1,* we screened for recombinant plants using self-pollinated progeny from the 18S39 F_2 population. Linkage mapping indicated the presence of two linked QTLs that were significantly associated with fruit weight on chromosome 6 (Fig. 1c). *fw6.1* was flanked by marker 18EP962 (SL4.0ch06:39,930,114) and 20EP567 (SL4.0ch06:42,662,777*),* spanning a 2.73 Mb interval, whereas *fw6.2* was flanked by marker 20EP1256 (SL4.0ch06:43,033,158) and 19EP629 (SL4.0ch06:44,319,744), spanning a 1.29 Mb interval. BGV006232 carried the large-fruited derived allele at both *fw6.1* and *fw6.2*, as shown by the allelic effects of the highest associated markers at each QTL (Supplementary Fig. S3).

To map each locus further, we selected recombinant plants that were segregating at one locus while fixed at the other. By evaluating two recombinant populations (Fig. 2a,b) and nine progeny testing families (Fig. 2c; Supplementary Table S6), we narrowed *fw6.1* to a 1.03 Mb region flanked by marker 22EP373 (SL4.0ch06:40,416,852) and 22EP37 (SL4.0ch06:41,451,757), and *fw6.2* to a 374 kb region between marker 21EP56 (SL4.0ch06:43,262,334) and 24EP416 (SL4.0ch06:43,636,481).

To carefully analyze fruit phenotypes that are controlled by *fw6.2* and gain insights into its mechanism to regulate fruit weight, we used NILs from an $F_{7:8}$ plant 22S310-30 in which *fw6.1* was fixed for the derived BGV006232 allele. At *fw6.2*, *fw.6.2-D* NILs carried the derived BGV006232 allele and *fw6.2-WT* NILs carried the wildtype BGV008225 allele. Fruit weight was significantly different between the two genotypes (Fig. 3a; Table 1), whereas ovary size was not affected (Table 1). The difference in fruit weight started to manifest itself during the development of the fruit 25 days after pollination (Fig. 3a). We conducted additional morphological and cytological measurements (Table 2). *Fw6.2*-D NILs showed a larger total fruit and pericarp area

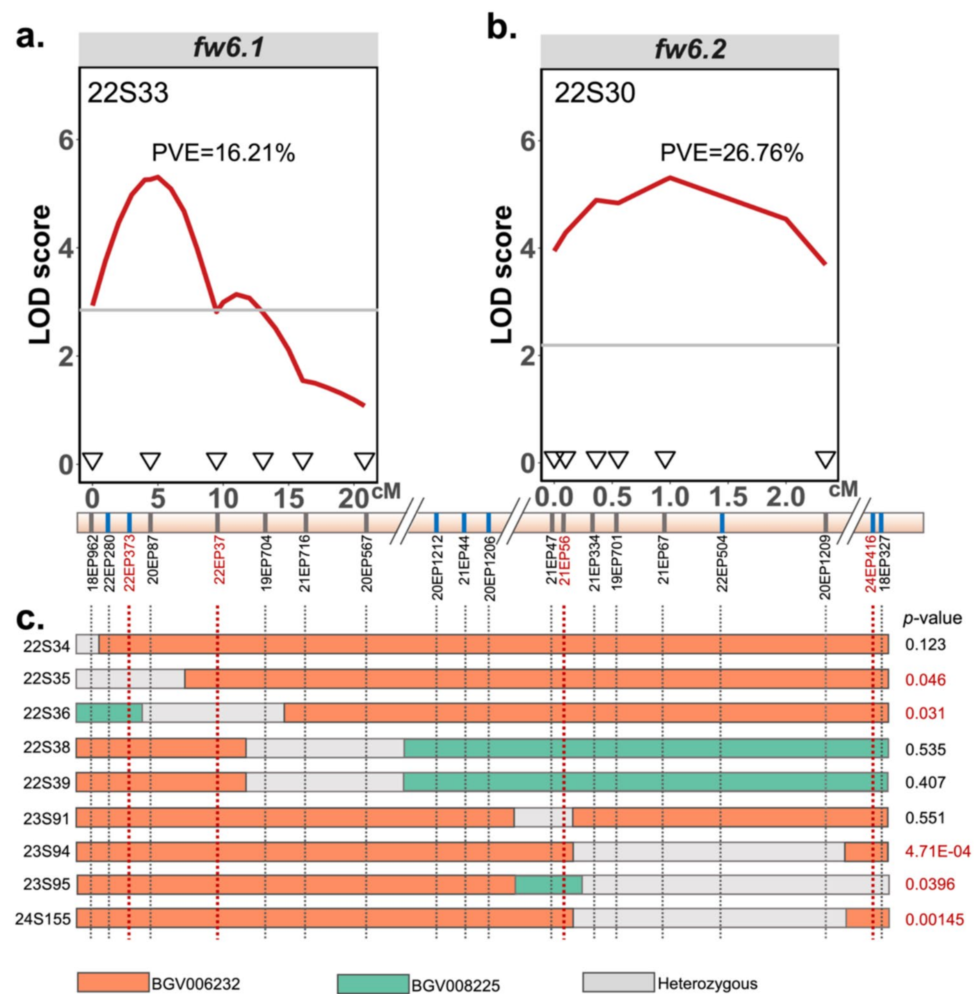

Fig. 2 Fine mapping of *fw6.1* and *fw6.2*. **a** Interval mapping of *fw6.1* in an F_6 population 22S33. **b** Interval mapping of *fw6.2* in an F_7 population 22S30. The markers used to genotype recombinant plants in 22S30 and 22S33 are shown with grey bars and additional markers used to genotype the progeny testing families are shown with blue bars. For a and b, the LOD threshold of $\alpha = 0.01$ is indicated in a gray line. **c** Progeny testing of *fw6.1* and *fw6.2*. The colored boxes indicate the parental genotype of each progeny testing family. All plants were grown under field condition except greenhouse-grown 24S155. The *p*-values were calculated using Student's t-test, and *p*-values < 0.05 are highlighted in red. The markers flanking *fw6.1* and *fw6.2* are highlighted with red dotted lines

than *fw6.2*-WT NILs. On the other hand, there was no significant segregation for pericarp area ratio, columella area, and columella area ratio. This suggests that the fruit weight increase in the *fw6.2*-D NILs was primarily driven by the enlargement of the pericarp. Moreover, the pericarp cell size was significantly larger in the *fw6.2-D* NILs, while the two genotypes showed a similar number of pericarp cell layers (Fig. 3b; Table 2). Therefore, *fw6.2* is likely to affect fruit weight by regulating the cell size in the pericarp.

Candidate gene at *fw6.2*: *CSR-like1*

Since *fw6.2* showed the highest PVE and smallest introgression, we evaluated the genes in the 374 kb interval in the tomato reference SL4.0 genome (Supplementary Table S7). Among the 51 genes, *Solyc06g073940* stood out as the paralog of the known fruit weight gene *CSR* controlling pericarp cell size, *CSR-like1* (Mu et al. 2017). The sequence analysis of *CSR-like1* showed that the BGV006232 and BGV008225 alleles carried no nucleotide polymorphisms in the protein coding region, and instead three SNPs in the 5' untranslated region (UTR), one SNP in the 3'UTR, and 16 SNPs in the putative 3 kb regulatory region upstream of *CSR-like1*.

To validate the function of *CSR-like1* in fruit weight, we first sought to create knockout mutants using the CRISPR-*Cas9* gene editing approach in two distinct accessions, a small-fruited SLC BGV007931 and a large-fruited SLL Fla. 8095 accession (Scott et al. 2008). We recovered two independent T_0s in BGV007931 and three independent T_0s in Fla. 8095 backgrounds, respectively, but all five T_0s carried in frame deletions of 6 bp or 12 bp (Supplementary Table S3). Despite the CRISPR-edited alleles were in-frame and not nulls, we evaluated the fruit weight variation in these lines (Supplementary Fig. S4). As expected, the in-frame edited alleles did not show consistent fruit weight differences compared to wildtype. We concluded that *CSR-like1* may play an essential role in plant development such that a complete knockout is likely to be lethal.

We next sought to downregulate the expression of *CSR-like1* using the artificial microRNA (amiRNA) technology, and recovered four T_0 in the large-fruited BGV006232

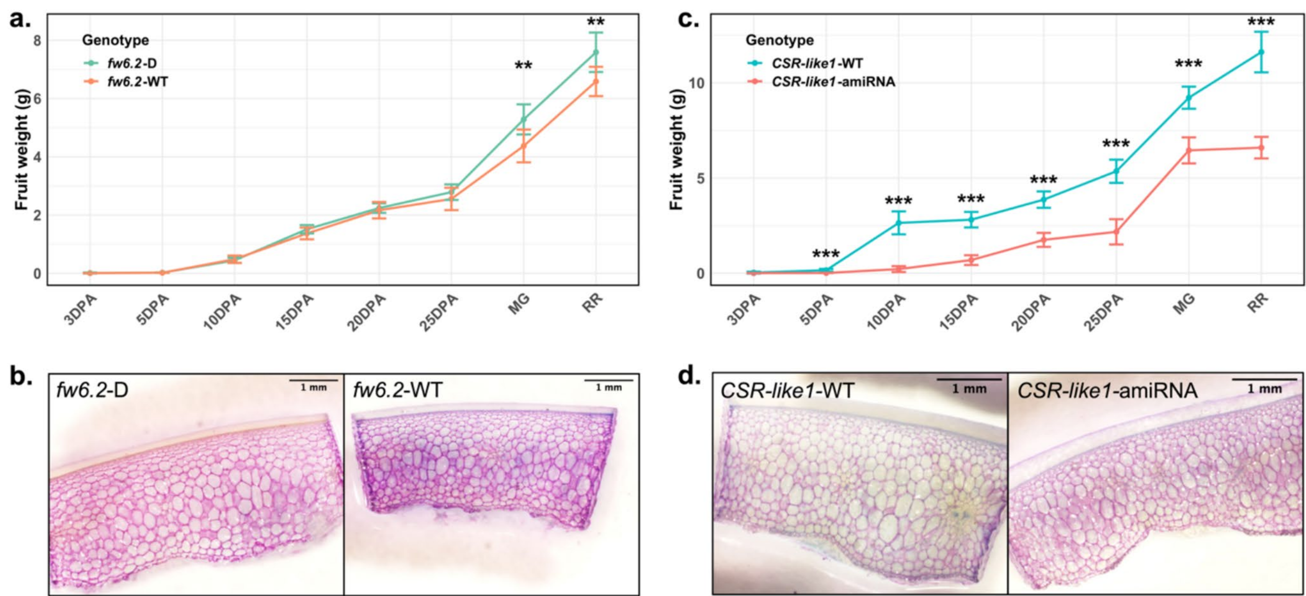

Fig. 3 Fruit development in *fw6.2* NILs and *CSR-like1* transgenic family. **a** Fruit weight at various fruit development stages (3, 5 10, 15, 20 and 25 DPA, mature green (MG) and red ripe (RR)) for *fw6.2* NILs. **b** Pericarp section of a representative mature green fruit of *fw6.2-D* and *fw6.2-WT* stained with toluidine blue. **c** Fruit weight at various fruit development stages for amiRNA transgenic family. **d** Pericarp section of a representative mature green fruit of *CSR-like1*-WT and *CSR-like1*-amiRNA. Student's t-test is used to calculate the *p*-values in a and c. Asterisks denote levels of statistical significance: $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***)

Table 1 Fruit weight and ovary size of *fw6.2* NILs (24S155) and *CSR-like1* amiRNA transgenic plants (24S156)

Population	24S155 (*fw6.2* NILs)			24S156 (*CSR-like1* amiRNA BC_1F_1)		
Genotype	*fw6.2*-D (n = 10)	*fw6.2*-WT (n = 10)	*P-value*[b]	WT (n = 10)	amiRNA (n = 10)	*P-value*
Average red ripe fruit weight (g)[a]	7.59 ± 0.68	6.59 ± 0.50	0.00145	11.62 ± 1.07	6.60 ± 0.57	1.13E-10
Ovary size (mm^2)	1.26 ± 0.096	1.23 ± 0.11	0.498	1.50 ± 0.104	1.35 ± 0.0826	2.74E-03

[a]All measurements in the table are given as mean ± standard deviation (SD)

[b]The *p*-values were calculated by Student's t-test

Table 2 Weight attribute evaluations of fruits and cellular evaluations of fruit pericarp from the *fw6.2* NILs (24S155) and *CSR-like1* amiRNA transgenic plants (24S156)

Population	24S155 (*fw6.2* NIL)			24S156 (*CSR-like1*amiRNA BC_1F_1)		
Genotype	*fw6.2*-D (n = 10)	*fw6.2*-WT (n = 10)	*P-value*[b]	WT (n = 10)	amiRNA (n = 10)	*P-value*
Perimeter (cm)[a]	8.25 ± 0.28	7.81 ± 0.18	4.29E-04	9.08 ± 0.21	8.41 ± 0.23	9.22E-06
Total area (cm^2)	4.89 ± 0.33	4.37 ± 0.20	4.45E-04	5.85 ± 0.25	5.05 ± 0.27	8.88E-06
Pericarp area (cm^2)	2.12 ± 0.15	1.90 ± 0.11	0.00109	2.35 ± 0.12	1.98 ± 0.12	9.31E-06
Columella area (mm^2)	6.50 ± 0.37	5.88 ± 0.52	0.00713	6.24 ± 0.59	5.54 ± 0.31	0.00743
Pericarp area ratio	0.433 ± 0.008	0.433 ± 0.008	0.892	0.401 ± 0.012	0.393 ± 0.011	0.121
Columella area ratio	0.013 ± 0.00070	0.013 ± 0.0011	0.741	0.011 ± 0.00089	0.011 ± 0.00047	0.430
Pericarp thickness (mm)	2.11 ± 0.13	1.88 ± 0.13	0.00112	2.42 ± 0.16	2.00 ± 0.09	8.02E-06
Pericarp cell size (mm^2)	0.0810 ± 0.0072	0.0671 ± 0.0077	7.74E-04	0.0974 ± 0.0117	0.0728 ± 0.0061	6.26E-05
Pericarp cell layer	12.54 ± 0.65	12.38 ± 0.80	0.642	11.53 ± 0.49	11.20 ± 0.30	0.105

[a]All measurements in the table are given as mean ± SD

[b]The *p*-values were calculated by Student's t-test

background. The fruit weight of *CSR-like1*-amiRNA plants was between 27 to 43% lower than controls under greenhouse conditions in the transformed lines, including a significant decrease in ovary size (Table 1; Supplementary Table S8). The pericarp area was significantly decreased in *CSR-like1*-amiRNA plants resulting from reduced cell size (Fig. 3d; Table 2). We also observed slower fruit development in *CSR-like1*-amiRNA plants, starting as early as 5DPA (Fig. 3c), and these plants required approximately 8–12 days more to reach the red ripe stage compared to *CSR-like1*-WT plants (Table 3). No vegetative effects were observed as the plants looked normal (Supplementary Table S8; Supplementary Fig. S5).

Evolution of *CSR-like1* in tomato domestication

Fruit weight was an important trait for selection during the domestication of most vegetable crops. To investigate the evolution of *CSR-like1* during tomato domestication, we explored the genetic diversity at the locus by generating a heatmap containing six haplotype clusters in the Varitome collection (Fig. 4a). A total of 88 variants were identified at the locus, of which 12 were INDELs ranging in size from 1 to 9 bp in addition to 76 SNPs (Supplementary Table S9). Most of the variants were located in the regulatory region and the UTRs. Similar to the parents in the mapping population, the coding region of *CSR-like1* was highly conserved in the Varitome collection with only 2 SNPs in distantly related accessions. SNP_ SL4.0ch06:43,344,537 resulted in a missense mutation from proline to leucine at amino acid position 355, while SNP_ SL4.0ch06:43,344,932 led to a synonymous mutation of glutamine at amino acid position 223. The clustering of haplotype groups was also associated with the phylogenetic groups of the accessions. Most of the SLC and all SLL accessions were found in Cluster I, II, and III. All of the wild tomato SP accessions were in Cluster IV, V and VI, carrying more mutations in the regulatory region and the UTRs. The missense mutation was only found in four accessions (1 SP and 3 SLC) in Cluster VI. We also surveyed an additional accession panel (SLL_CUL) including 73 modern cultivars, landraces and heirlooms (Tieman et al. 2017), and assigned their haplotypes using conserved SNPs across Cluster III to VI (Supplementary Table S10). These cultivated accessions also have either haplotype I or II, similar to the ancestral SLL in the Varitome collection (Supplementary Table S10). We will now refer haplotype I and II allele together as *CSR-like1-D* for derived which was most similar to the Heinz 1706 reference genotype, and haplotype III, IV, V, and VI together as *CSR-like1-WT* for wildtype.

For the two mapping parents of *fw6.2*, the large-fruited parent BGV006232 carries the *CSR-like1-D* allele (haplotype II) while the small-fruited parent BGV0008225 carries the *CSR-like1-WT* allele (haplotype IV). To further investigate the association between *CSR-like1* haplotypes in the entire Varitome collection and fruit traits, we selected only accessions carrying the wildtype *fw11.3/CSR* allele since that gene has a large effect on fruit weight and pericarp cell size. The resulting 108 SLC and 27 SP accessions showed that fruit weight and pericarp cell size were larger in the accessions carrying *CSR-like1-D* alleles (Fig. 4b,c). Similarly, when fixing for *CNR* (*fw2.2*), *KLUH* (*fw3.2*) and all three known fruit weight genes, the *CSR-like1* haplotype I and II were always associated with larger fruits and pericarp cells (Supplemental Figure S6), further supporting the notion that this gene is critical in the regulation of fruit weight via increased cell sizes. We also identified four SNPs in the upstream and downstream of *CSR-like1* that are exclusively found in *CSR-like1-WT* alleles (red asterisks in Fig. 4a). These SNPs are the four most significant variants associated with fruit weight and pericarp cell size (Supplementary Table S11).

Table 3 The number of days from anthesis to red ripe for *CSR-like1*-WT and *CSR-like1*-amiRNA fruits in four amiRNA transgenic families

Population	Genotype	Anthesis to red ripe (days)[a]
24S156 (BC_1F_1)	WT (n = 10)	40.56 ± 1.76
	amiRNA (n = 10)	52.28 ± 2.31
	P-value[b]	1.88E-10
25S158 (BC_1F_1)	WT (n = 11)	34.97 ± 1.40
	amiRNA (n = 9)	44.82 ± 4.35
	P-value	1.86E-06
25S159 (BC_1F_1)	WT (n = 10)	36.19 ± 1.03
	amiRNA (n = 10)	44.46 ± 2.60
	P-value	2.46E-08
25S160 (BC_1F_1)	WT (n = 10)	34.95 ± 1.40
	amiRNA (n = 10)	42.92 ± 2.46
	P-value	5.23E-08

[a]All measurements in the table are given as mean ± SD

[b]The *p*-values were calculated by Student's t-test

CSR-like1 associated with fruit weight in pepper

Orthologs of tomato fruit weight genes have been reported as likely candidate genes underlying fruit weight QTLs in other Solanaceae crops including pepper and eggplant (Rinaldi et al. 2016; Toppino et al. 2016; Zygier et al. 2005) as well as some Cucurbitaceae crops including melon, cucumber and watermelon (Monforte et al. 2014; Pan et al. 2020). To explore the potential conserved function of *CSR-like1* in regulating fruit weight in pepper, we conducted GWAS using multi-year phenotypic data collected from four *Capsicum*

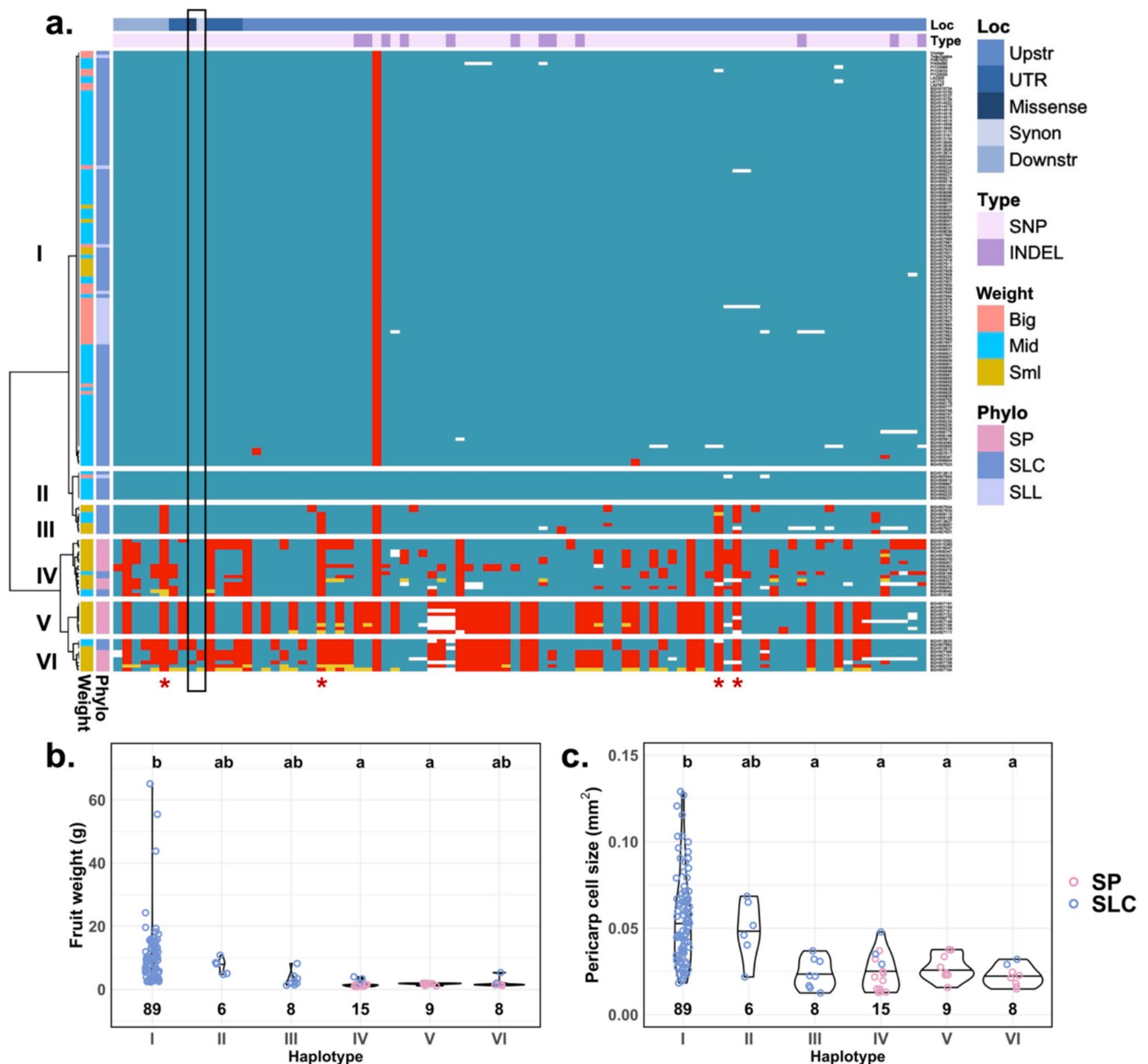

Fig. 4 Haplotype analysis of *CSR-like1* in the Varitome collection. **a** Heatmap representing the genotypes of accessions (rows) at the *CSR-like1* locus. Reference genotypes are represented in blue, alternate in red, heterozygous in yellow and missing data in white. The first left column indicates the fruit weight category (big, medium, or small) of each accession. The second left column shows the phylogenic group of each accession (SP, SLC, or SLL). The first top bar shows the location of the variant: Upstream (Upstr), 5' or 3' untranslated region (UTR), protein coding region (missense or synonymous (Synon) mutation), and downstream (Downstr). The second top bar indicates the type of variant (SNP or INDEL). The only two SNPs in the coding region are highlighted in the black box. Red asterisks indicate the four variants that are conserved in most of the accessions carrying *CSR-like1*-WT allele (Supplementary Table S11). **b,c** Violin plots of fruit weight and pericarp cell size of the Varitome accessions fixed with *CSR-WT* allele, grouped by haplotype clusters

species (*C. annuum*, *C. annuum* var. *glabriusculum*, *C. baccatum*, and *C. chinense*) (Supplementary Table S12). We identified three significant SNPs (S06_227195491, S06_227195529, and S06_227195619) in the mixed linear model (MLM), and one SNP (S06_227195491) in the BLINK model, all located within the coding region of *CaCSR-like1* (*CA06g22610*) (Fig. 5a,b). The *Capsicum* ortholog CaCSR-like1 exhibited conserved motifs and domain structures characteristic of the CSR/FAF-like proteins (Mu et al. 2017), suggesting a conserved molecular mechanism in regulating cell size and fruit mass.

To further characterize allelic diversity of *CaCSR-like1*, we identified a total of five SNPs (S06_227195491, S06_227195529, S06_227195579, S06_227195589, and

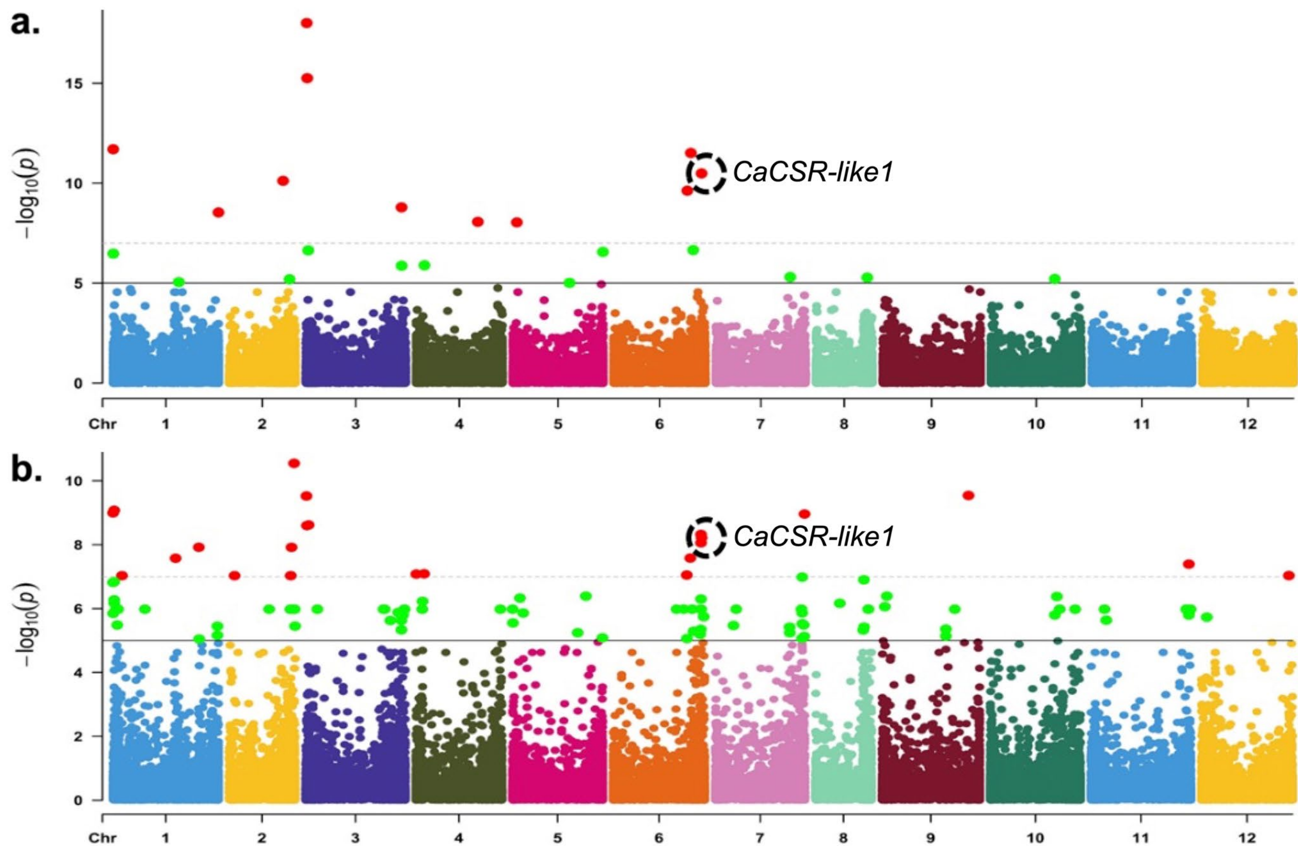

Fig. 5 GWAS for the fruit weight in pepper using BLINK (**a**) and MLM (**b**) models. The horizontal dotted line indicates the Bonferroni-corrected threshold ($p < 0.05$), and the dotted circle highlights significant SNPs linked to the *CaCSR-like1* gene

S06_227195619) in the coding region of *CaCSR-like1* among the *Capsicum* GWAS panel using the GBS markers (Fig. 6a). Six haplotypes (haplotype I to VI) were assigned using the five segregating sites, and only SNP S06_227195619, S06_227195579 and S06_22719549 led to non-synonymous changes at amino acid position 10, 24 and 53 (Fig. 6a). These non-synonymous mutations were outside of the FAF domain (Pfam accession: PF11250) which is located between amino acid position 230 and 278. The wild pepper *C. annuum* var. *glabriusculum* in this study only carried haplotype II or III, while the cultivated pepper *C. annuum*, *C. baccatum*, and *C. chinense* exhibited all six haplotypes across different accessions (Fig. 6b; Supplementary Table S13). There is also clear association of haplotype VI with higher fruit weight within each cultivated pepper species (Fig. 6b).

Discussion

Fruit weight is a critical component in modern agriculture as it is linked to yield and the price of produce. Hence a key focus in many vegetable breeding programs is to maintain or increase fruit weight. Understanding the genetic basis underlying this highly quantitative trait can provide us valuable insights into fruit development and help accelerate breeding programs. In this study, we successfully identified five unique fruit weight QTL from a GWAS with diverse SP, SLC and SLL accessions. Of these loci, three were likely to correspond to known fruit weight genes while two were novel. We subsequently fine mapped two QTLs on chromosome 6 using association mapping and progeny testing. Of these*, fw6.2* harbors the smaller interval of 374 kb. Based on changes in pericarp cell size in *fw6.2* NILs, *CSR-like1* was identified as the most likely candidate gene at the locus. *CSR-like1* is a paralog of the only known fruit weight gene *CSR* that regulates cell enlargement.

An additional 50 genes are found at the *fw6.2* locus. Therefore, we can't exclude the possibility of other genes contributing to fruit weight at this locus. Even so, we successfully excluded *SP* (*SELF-PRUNING*; *Solyc06g074350*) that had recently been shown to control fruit weight (Ning et al. 2023). *SP* is located ~ 291 kb downstream of *CSR-like1*, and is primarily known to modulate plant growth habit by regulating the vegetative-reproductive switch in shoot growth (Pnueli et al. 1998). *fw6.3* was mapped to a

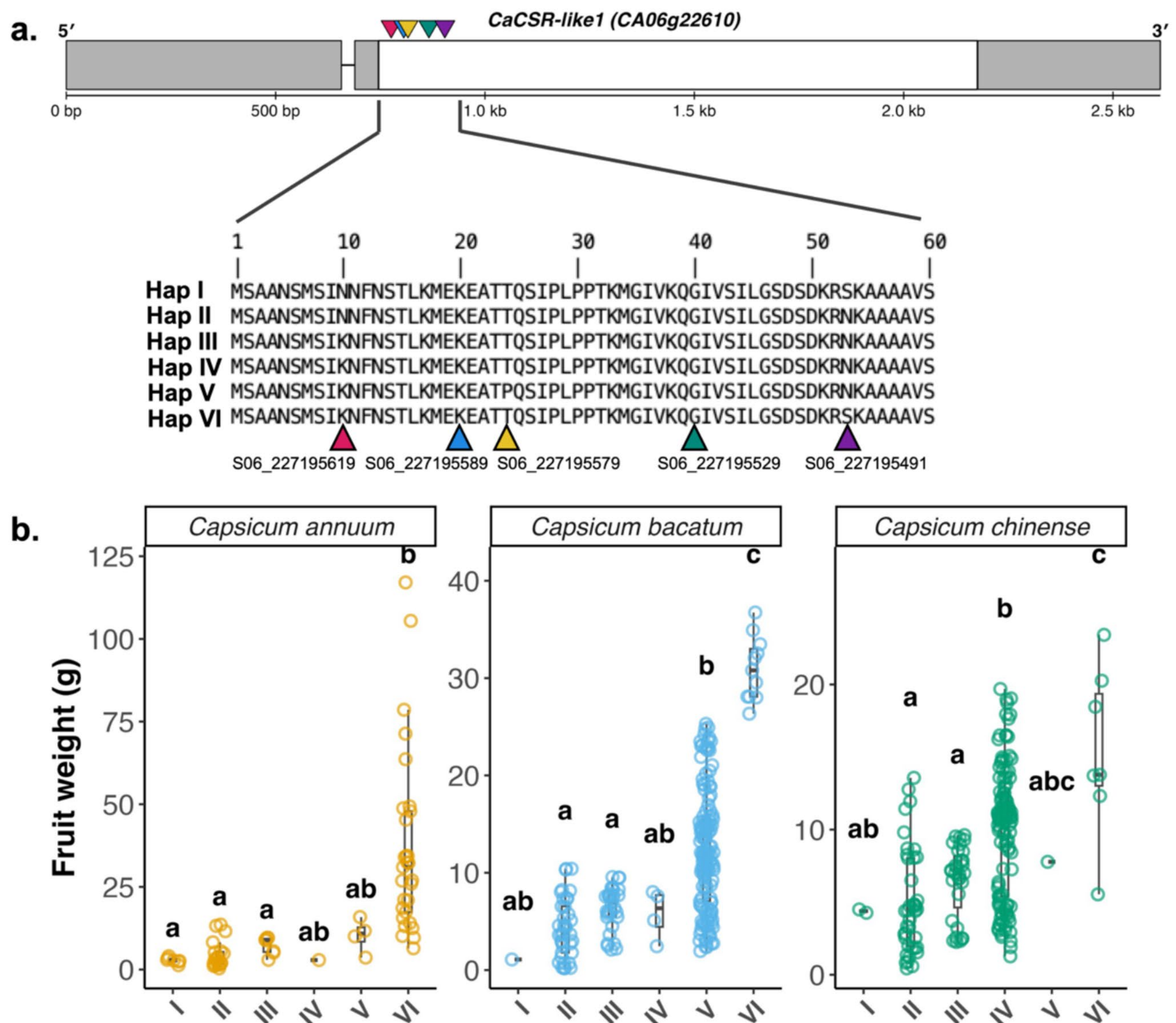

Fig. 6 Haplotype analysis of *CaCSR-like1*. **a** Six haplotypes identified in the pepper GWAS panel. The gene structure of *CaCSR-like1* is shown on the top in which the grey boxes represent the 5′ and 3′ UTR, the white box represents the coding sequence, and the straight line represents the intron. Colored triangles represent the five segregating SNPs. Part of the amino acid sequence (position 1 to 60) containing the SNPs is shown for each haplotype. **b** Haplotype association with fruit weight in cultivated species *C. annuum*, *C. baccatum*, and *C. chinense*

62.6 kb interval that concluded *SELF-PRUNING* (*SP*) as the most likely candidate gene (Ning et al. 2023). We designed a marker 24EP416, locating in the third intron of SP, and were able to exclude *SP* from *fw6.2* using progeny testing of families 23S94 and 24S155 (Fig. 2).

The haplotype analyses of *CSR-like1* showed that smaller pericarp cell size was associated with accessions carrying *CSR-like1-WT* alleles in backgrounds where the other cloned fruit weight genes were fixed (Fig. 4c and Supplementary Figure S6). Moreover, among the F_2 populations that were used to validate *GWAS_fw6.1*, four showed significant association with fruit weight at the locus (Supplementary Table S5). Of those that segregated at the GWAS locus, the large-fruited parent of each of the four populations carries *CSR-like1-D* allele while the other parent carries *CSR-like1-WT* allele except for 18S40 where both parents carry the same *CSR-like1-D* allele. The small effect on fruit weight at *GWAS_fw6.1* may be the result of the other segregating fruit weight QTLs in the population and imply that other genes at the *GWAS_fw6.1* locus affect the trait as well (Supplementary Tables S1 and S5). The fifth F_2 population, 18S32, was neither segregating for polymorphisms at *GWAS_fw6.1*, nor for the *CSR-like1* allele. Therefore, as expected, the locus was not associated with fruit weight in the 18S32 population.

Together, genetic mapping and haplotype association analysis demonstrate that *CSR-like1* is a likely candidate for fruit weight at the *fw6.2* locus. In addition, the haplotype analyses revealed four most significant SNPs associated with fruit weight that could be utilized in future breeding programs for selecting desired fruit weight in tomato (Supplementary Table S11).

CSR-like1 was located 19 kb away from one of the selective sweep windows in the transition from SP to SLC (Razifard et al. 2020), suggesting that the gene may have been selected in the early stage of tomato domestication. This agrees with the distribution of *CSR-like1* haplotypes in the genetically distinct groups classified in Razifard et al. (2020) (Supplementary Fig. S7). All SP accessions in the Varitome collection carried haplotype IV, V or VI, which were associated with smaller fruit weight. Based on the tomato domestication hypothesis proposed by Razifard et al. (2020), Ecuadorian SLC (SLC_ECU) accessions would be the most ancestral SLC group, and thus would be the first group where larger fruit-associated haplotype I and II allele of *CSR-like1* appeared. The frequency of haplotype I and II allele would increase with further domestication. Haplotype III allele is associated with smaller fruit weight compared to haplotype I and II, and is mostly found in a group of SLC accessions from Mexico, Central America, and northern South America (together annotated as SLC_CA). The rise of haplotype III allele would coincide with the transition of fruit weight back to more wild-like tomatoes in SLC_CA, and we would expect to see a reselection of Haplotype I and II of *CSR-like1* in the fully cultivated SLL from SLC_Mexico. However, such re-selection was not detected in the study of Razifard et al. (2020) at the *CSR-like1* locus. In fact, the haplotype distribution of *CSR-like1* could demonstrate another scenario in the evolution of tomato such as described in Blanca et al. (2022) (Supplementary Fig S8). SLC_CA was likely to be the intermediate group after the northward migration of SP towards Mesoamerica. Later SLC_CA migrated back to South America and admixed with Ecuadorian and Peruvian SP to give rise to Ecuadorian and Peruvian SLC in which haplotype I and II allele of *CSR-like1* were further distributed. Some Peruvian SLC then migrated back to Mexico where it evolved fully to SLL.

We functionally validated the role of *CSR-like1* in controlling fruit weight and pericarp cell size by downregulating its expression in transgenic plants. In the amiRNA lines, the effect of reduced expression of *CSR-like1* on fruit development started to manifest itself in ovaries one day prior to anthesis. *CSR-like1* is the only member of the *CSR/FAF-like* family in tomato that has relatively high expression in young flower buds and flowers at anthesis (Mu et al. 2017). This suggests that *CSR-like1* might regulate cell differentiation and enlargement in floral development as well even though the NILs did not show a difference at that time. Similar differences between lines carrying natural alleles and lines carrying transgenes has been observed in *fw3.2/KLUH* (Chakrabarti et al. 2013). Compared to *fw3.2* NILs, the downregulated transgenic lines showed additional phenotypic defects across the entire plant, including plant height, leaves and leaflets, seed number and side shoot number (Chakrabarti et al. 2013). Future experiments are needed to determine the developmental time frame for when and how *CSR-like1* acts in floral development.

Tomato serves as an important model crop for studying fruit development, with many genes displaying conserved functions across plant species. Recently, two association mapping studies in *C. annuum* and *C. chinense*, respectively, reported SNPs in the orthologs of *CSR-like1* to be significantly associated with fruit weight (Nimmakayala et al. 2016, 2021). Here we are adding another piece of evidence of a pepper *CSR-like1* regulating fruit weight using a more expanded *Capsicum* collection with multi-model GWAS identifying significant SNPs within the coding region of *CaCSR-like1*. The presence of the distinct haplotypes associated with fruit weight further indicates that allelic diversification of *CaCSR-like1* contributed to the phenotypic spectrum observed in cultivated *Capsicum* species. Haplotype VI of *CaCSR-like1* was enriched among large-fruited accessions, likely represents a derived allele favored during domestication or selection for increased fruit mass.

There is evidence of *CSR-like1* being relevant in other crops. In watermelon and cucumber, QTL mapping using recombinant inbred lines (RILs) has identified fruit weight/size QTLs that harbor the respective orthologs of *CSR-like1* (*ClCG02G022450* in watermelon; *CsGy6G022740* in cucumber) (Guo et al. 2024; Weng et al. 2015). A recent GWAS utilizing the CucCAP cucumber core collection also identified a SNP which is ~ 250 kb upstream of *CsGy6G022740* to be significantly associated with fruit size (Lin et al. 2024). In eggplant, though there have not been any reported fruit weight QTLs mapped close to *CSR-like1*, a fruit weight QTL of around 10 cM on chromosome 12 was mapped in an F_2 population (*S. melongena* 305E40 x *S. melongena* 67/3), with *SMEL4.1_12g014140.1*, the ortholog of *CSR*, residing at the locus (Gaccione et al. 2023; Portis et al. 2014). Direct genetic modification in other crops could be applied to broaden our understanding on the potential conserved function of *CSR-like1* and its related genes in regulating fruit size and other species-specific functions.

While the function on fruit size and cell size by *CSR-like1* and *CSR* in tomato is apparent, its molecular function is less clear. *CSR-like1* is predicted to encode a protein that contains a FANTASTIC FOUR (FAF) domain, and is an ortholog of *FAF-like* (*AT5G22090*) in *Arabidopsis thaliana* (Mu et al. 2017). *FAF-like* and its related family *FAF* were first reported in *Arabidopsis*, which carries four *FAF* members and one *FAF-like* (Wahl et al. 2010).

Phylogenetic analysis in monocotyledonous and dicotyledonous plants suggest that the *FAF* genes have evolved from a *FAF-like* gene to become a dicotyledonous-specific gene family after gene duplication (Wahl et al. 2010). Many of the plant species in the Rosids and Asterids clades carry only one copy of *FAF-like* (Mu et al. 2017; Wahl et al. 2010). Interestingly, an expansion of *FAF-like* is observed in the Solanaceae family. Tomato, potato (*S. tuberosum*) and chili pepper (*Capsicum annuum*) carry four *FAF-like* genes whereas eggplant (*S. melongena*) carries three members (Mu et al. 2017). In Arabidopsis, *FAF-like* (*AT5G22090*) encodes a protein described as EAR1 (ENHANCER of ABA CO-RECEPTOR1), which can enhance the activity of clade A type 2C protein phosphatases (PP2Cs) by binding to their N termini, causing the inhibition of Snf1-related kinases2 (SnRK2s) (Wang et al. 2018). Meanwhile, the expression of many downstream targets of ABA signaling rely on phosphorylation by SnRK2s (Hasan et al. 2022). Therefore, EAR1 is a negative regulator of ABA signaling and shown to affect seed germination, primary root growth and drought tolerance (Wang et al. 2018). Interestingly, in pepper, *CaCSR-like1* (*CA06g22610*; or *CaFAF1*) regulates the ABA signaling pathway but in a different manner than *EAR1* (Lim et al. 2022; Wang et al. 2018). CaCSR-like1 does not interact with the known pepper PP2Ps in yeast two-hybrid assays, and unlike *EAR1*, *CaCSR-like1* does not affect seed germination and primary root growth when overexpressed in Arabidopsis. Instead, *CaCSR-like1* plays a positive role in drought stress and a negative role in salt stress. Therefore, *CaCSR-like1's* function differs under certain abiotic stresses (Lim et al. 2022).

In this study, fruit ripening is delayed in the transgenically downregulated *CSR-like1* lines. It is a complex process controlled by many regulators through phytohormone and environmental signals (Kou et al. 2021). ABA is known to be a key signaling hormone to regulate the ripening process in tomato through crosstalk with ethylene biosynthesis and metabolism (Kou et al. 2021; Mou et al. 2016; Zhang et al. 2009). As discussed earlier, EAR1 in Arabidopsis is a negative regulator of ABA signaling pathway (Wang et al. 2018). Silencing one of the tomato *PP2C*s, *SlPP2C3* (*Solyc06g076400*), was shown to accelerate the ripening process (Liang et al. 2021). Therefore, if CSR-like1 were to have a similar function as EAR1 and interact with PP2Cs like PP2C3, we would expect to see a shortened ripening process in our amiRNA transgenic plants; instead, we observed the opposite result. This suggests that CSR-like1 might have a divergent function from EAR1. As mentioned earlier, a similar case has been made in pepper where CaCSR-like1 was demonstrated to affect ABA signaling through a different mechanism under abiotic stresses (Lim et al. 2022). The expansion of CSR/FAF-like family especially in the Solanaceae crops also suggests CSR/FAF-like proteins might have evolved to have novel functions evolution.

In summary, our data support a critical role for the tomato *FAF-like* genes, *CSR* and *CSR-like 1* on fruit weight in this species and potentially other crops, especially in the Solanaceae family. Further understanding its molecular function should lead to insights into fundamental plant processes while at the same time enabling breeders to implement the knowledge presented in their breeding programs.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-026-05177-x.

Acknowledgements We want to thank Dr. Ana Caicedo for helpful suggestions about the evolution of the locus; Katherine Hardigree for plant care and field experiments; Neda Keyhaninejad for tomato transformations; and all members of the Van der Knaap lab for field harvest and fruit weight evaluations.

Author contribution EvdK conceived the study and supervised the research. QF, LP, and MS performed all experiments and data analyses on tomato. YW created the *CSR-like1* amiRNA construct and genotyped the T_0 tomato lines. KSK, PN and UR performed all experiments and data analyses on pepper. QF and EvdK drafted the original manuscript. All authors reviewed, provided comments and agreed to the published version of the manuscript.

Funding This work was supported by NSF IOS 1564366 to EvdK.

Data availability The raw DNA sequence data for tomato is available in NCBI (https://www.ncbi.nlm.nih.gov/; SRA: SRP150040, SRP045767, SRP094624, and PRJNA353161). The GBS data for pepper is available in NCBI (https://www.ncbi.nlm.nih.gov/; PRJNA1305095).

Declarations

Conflict of interest The corresponding author, Esther van der Knaap, is a member of the Editorial Board of Theoretical and Applied Genetics. As required by journal policy, they will not be involved in the editorial handling or peer-review process of this manuscript. Another Editor with no competing interests will be assigned to oversee the review. All authors declare that no other competing interests exist.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alexander DH, Shringarpure SS, Novembre J and Lange K (2015) Admixture 1.3 software manual. Los Angeles: UCLA Human Genetics Software Distribution

Alonge M, Wang X, Benoit M et al (2020) Major impacts of widespread structural variation on gene expression and crop improvement in tomato. Cell 182:145-161.e23. https://doi.org/10.1016/j.cell.2020.05.021

Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890. https://doi.org/10.1093/bioinformatics/btg112

Blanca J, Sanchez-Matarredona D, Ziarsolo P, Montero-Pau J, van der Knaap E, Díez MJ and Cañizares J (2022)Haplotype analyses reveal novel insights into tomato history and domestication driven by long-distance migrations and latitudinal adaptations. Horticulture Research 9. https://doi.org/10.1093/hr/uhac030

Causse M, Saliba-Colombani V, Lecomte L, Duffé P, Rousselle P, Buret M (2002) QTL analysis of fruit quality in fresh market tomato: a few chromosome regions control the variation of sensory and instrumental traits. J Exp Bot 53:2089–2098. https://doi.org/10.1093/jxb/erf058

Chakrabarti M, Zhang N, Sauvage C et al (2013) A cytochrome P450 regulates a domestication trait in cultivated tomato. Proc Natl Acad Sci U S A 110:17125–17130. https://doi.org/10.1073/pnas.1307313110

Cheng R, Parker CC, Abney M, Palmer AA (2013) Practical considerations regarding the use of genotype and pedigree data to model relatedness in the context of genome-wide association studies. G3 Genes|genomes|genetics 3:1861–1867. https://doi.org/10.1534/g3.113.007948

Cingolani P, Adrian P, Lily WL, Melissa C, Tung N, Luan W, J. LS, Xiangyi L, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly 6:80–92. https://doi.org/10.4161/fly.19695

Cong B, Liu J, Tanksley SD (2002) Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. Proc Natl Acad Sci U S A 99:13606–13611. https://doi.org/10.1073/pnas.172520999

Cong B, Barrero LS, Tanksley SD (2008) Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nat Genet 40:800–804. https://doi.org/10.1038/ng.144

Danecek P, Auton A, Abecasis G et al (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM and Li H (2021) Twelve years of SAMtools and BCFtools. Gigascience 10. https://doi.org/10.1093/gigascience/giab008

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127:1309–1321. https://doi.org/10.1016/j.cell.2006.12.006

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379. https://doi.org/10.1371/journal.pone.0019379

Frary A, Nesbitt TC, Frary A, Grandillo S, Van Der Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB (2000) Fw2. 2: a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88. https://doi.org/10.1126/science.289.5476.85

Gaccione L, Martina M, Barchi L, Portis E (2023) A compendium for novel marker-based breeding strategies in eggplant. Plants 12:1016. https://doi.org/10.3390/plants12051016

Grandillo S, Ku HM, Tanksley SD (1999) Identifying the loci responsible for natural variation in fruit size and shape in tomato. Theor Appl Genet 99:978–987. https://doi.org/10.1007/s001220051405

Gross BL, Olsen KM (2010) Genetic perspectives on crop domestication. Trends Plant Sci 15:529–537. https://doi.org/10.1016/j.tplants.2010.05.008

Guo S, Tian M, Du H, Liu S, Yu R, Shen H (2024) Quantitative trait loci mapping and comparative transcriptome analysis of fruit weight (FW) in watermelon (*Citrullus lanatus* L.). Genes. https://doi.org/10.3390/genes15070933

Hasan MM, Liu XD, Waseem M, Guang-Qian Y, Alabdallah NM, Jahan MS, Fang XW (2022) ABA activated SnRK2 kinases: an emerging role in plant growth and physiology. Plant Signal Behav 17:2071024. https://doi.org/10.1080/15592324.2022.2071024

Hodgens C, Nimchuk ZL, Kieber JJ (2017) IndCAPS: a tool for designing screening primers for CRISPR/Cas9 mutagenesis events. PLoS ONE 12:e0188406. https://doi.org/10.1371/journal.pone.0188406

Illa-Berenguer E, Van Houten J, Huang Z, van der Knaap E (2015) Rapid and reliable identification of tomato fruit weight and locule number loci by QTL-seq. Theor Appl Genet 128:1329–1342. https://doi.org/10.1007/s00122-015-2509-x

Kim S, Park M, Yeom SI et al (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. Nat Genet 46:270–278. https://doi.org/10.1038/ng.2877

Knaus BJ, Grünwald NJ (2017) Vcfr: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour 17:44–53. https://doi.org/10.1111/1755-0998.12549

Kou X, Zhou J, Wu CE, Yang S, Liu Y, Chai L, Xue Z (2021) The interplay between ABA/ethylene and NAC TFs in tomato fruit ripening: a review. Plant Mol Biol 106:223–238. https://doi.org/10.1007/s11103-021-01128-w

Liang B, Sun Y, Wang J, Zheng Y, Zhang W, Xu Y, Li Q, Leng P (2021) Tomato protein phosphatase 2C influences the onset of fruit ripening and fruit glossiness. J Exp Bot 72:2403–2418. https://doi.org/10.1093/jxb/eraa593

Lim CW, Bae Y, Lee SC (2022) Differential role of *Capsicum annuum* FANTASTIC FOUR-like gene CaFAF1 on drought and salt stress responses. Environ Exp Bot 199:104887. https://doi.org/10.1016/j.envexpbot.2022.104887

Lin YC, Weng Y, Fei Z, Grumet R (2024) Mining the cucumber core collection: phenotypic and genetic characterization of morphological diversity for fruit quality characteristics. Horticult Res. https://doi.org/10.1093/hr/uhae340

Mata-Nicolás E, Montero-Pau J, Gimeno-Paez E, Garcia-Carpintero V, Ziarsolo P, Menda N, Mueller LA, Blanca J, Cañizares J, Van Der Knaap E, Díez MJ (2020) Exploiting the diversity of tomato: the development of a phenotypically and genetically detailed germplasm collection. Horticult Res. https://doi.org/10.1038/s41438-020-0291-7

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. https://doi.org/10.1093/jxb/eru017

Mou W, Li D, Bu J, Jiang Y, Khan ZU, Luo Z, Mao L, Ying T (2016) Comprehensive analysis of ABA effects on ethylene biosynthesis and signaling during tomato fruit ripening. PLoS ONE 11:e0154072. https://doi.org/10.1371/journal.pone.0154072

Mu Q, Huang Z, Chakrabarti M, Illa-Berenguer E, Liu X, Wang Y, Ramos A, van der Knaap E (2017) Fruit weight is controlled by Cell Size Regulator encoding a novel protein that is expressed in maturing tomato fruits. PLoS Genet 13:e1006930. https://doi.org/10.1371/journal.pgen.1006930

Muños S, Ranc N, Botton E, Bérard A, Rolland S, Duffé P, Carretero Y, Le Paslier M-C, Delalande C, Bouzayen M (2011) Increase in

tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiol 156:2244–2254. https://doi.org/10.1104/pp.111.173997

Nimmakayala P, Abburi VL, Saminathan T et al (2016) Genome-wide diversity and association mapping for capsaicinoids and fruit weight in *Capsicum annuum* L. Sci Rep 6:38081. https://doi.org/10.1038/srep38081

Nimmakayala P, Lopez-Ortiz C, Shahi B, Abburi VL, Natarajan P, Kshetry AO, Shinde S, Davenport B, Stommel J, Reddy UK (2021) Exploration into natural variation for genes associated with fruit shape and size among *Capsicum chinense* collections. Genomics 113:3002–3014. https://doi.org/10.1016/j.ygeno.2021.06.041

Ning Y, Wei K, Li S et al (2023) Fine mapping of fw6.3, a major-effect quantitative trait locus that controls fruit weight in tomato. Plants 12:2065

Pan Y, Wang Y, McGregor C, Liu S, Luan F, Gao M, Weng Y (2020) 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

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. https://doi.org/10.1093/jxb/erm257

Pereira L, Sapkota M, Alonge M et al (2021a) Natural genetic diversity in tomato flavor genes. Front Plant Sci. https://doi.org/10.3389/fpls.2021.642828

Pereira L, Zhang L, Sapkota M, Ramos A, Razifard H, Caicedo AL, Van der Knaap E (2021b) Unraveling the genetics of tomato fruit weight during crop domestication and diversification. Theor Appl Genet 134:3363–3378. https://doi.org/10.1007/s00122-021-03902-2

Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, Ganal M, Zamir D, Lifschitz E (1998) The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. Development 125:1979–1989. https://doi.org/10.1242/dev.125.11.1979

Portis E, Barchi L, Toppino L, Lanteri S, Acciarri N, Felicioni N, Fusari F, Barbierato V, Cericola F, Valè G, Rotino GL (2014) QTL mapping in eggplant reveals clusters of yield-related loci and orthology with the tomato genome. PLoS ONE 9:e89499. https://doi.org/10.1371/journal.pone.0089499

Razifard H, Ramos A, Della Valle AL et al (2020) Genomic evidence for complex domestication history of the cultivated tomato in Latin America. Mol Biol Evol 37:1118–1132. https://doi.org/10.1093/molbev/msz297

Reddy UK, Karnatam KS, Talavera-Caro A et al (2025) Uncovering the genetic architecture of pungency, carotenoids, and flavor in *Capsicum chinense* via TWAS-mGWAS integration and spatial transcriptomics. Hortic Res. https://doi.org/10.1093/hr/uhaf243

Rinaldi R, Van Deynze A, Portis E, Rotino GL, Toppino L, Hill T, Ashrafi H, Barchi L, Lanteri S (2016) New insights on eggplant/tomato/pepper synteny and identification of eggplant and pepper orthologous QTL. Front Plant Sci 7:1031. https://doi.org/10.3389/fpls.2016.01031

Rodríguez GR, Moyseenko JB, Robbins MD, Huarachi Morejón N, Francis DM, Van Der Knaap E (2010) Tomato analyzer: a useful software application to collect accurate and detailed morphological and colorimetric data from two-dimensional objects. J vis Exp. https://doi.org/10.3791/1856

Rosyara UR, De Jong WS, Douches DS, Endelman JB (2016) Software for genome-wide association studies in autopolyploids and its application to potato. Plant Genome. https://doi.org/10.3835/plantgenome2015.08.0073

Sapkota M, Pereira L, Wang Y, Zhang L, Topcu Y, Tieman D, van der Knaap E (2023) Structural variation underlies functional diversity at methyl salicylate loci in tomato. PLoS Genet 19:e1010751. https://doi.org/10.1371/journal.pgen.1010751

Schindelin J, Arganda-Carreras I, Frise E et al (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods 9:676–682. https://doi.org/10.1038/nmeth.2019

Scott JW, Baldwin EA, Klee HJ, Brecht JK, Olson SM, Bartz JA, Sims CA (2008) Fla. 8153 hybrid tomato; Fla. 8059 and Fla. 7907 breeding lines. HortScience 43:2228–2230. https://doi.org/10.21273/hortsci.43.7.2228

Tanksley SD (2004) The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. Plant Cell 16:S181–S189. https://doi.org/10.1105/tpc.018119

Tieman D, Zhu G, Resende MF, Lin T, Nguyen C, Bies D, Rambla JL, Beltran KSO, Taylor M, Zhang B (2017) A chemical genetic roadmap to improved tomato flavor. Science 355:391–394. https://doi.org/10.1126/science.aal1556

Topcu Y, Sapkota M, Illa-Berenguer E, Nambeesan SU, Van der Knaap E (2021) Identification of blossom-end rot loci using joint QTL-seq and linkage-based QTL mapping in tomato. Theor Appl Genet 134:2931–2945. https://doi.org/10.1007/s00122-021-03869-0

Toppino L, Barchi L, Lo Scalzo R et al (2016) Mapping quantitative trait loci affecting biochemical and morphological fruit properties in eggplant (*Solanum melongena* L.). Front Plant Sci. https://doi.org/10.3389/fpls.2016.00256

Van Eck J, Keen P, Tjahjadi M (2019) Agrobacterium tumefaciens-Mediated Transformation of Tomato. In: Kumar S, Barone P, Smith M (eds) Transgenic Plants: Methods and Protocols, edn. Springer, New York, New York, NY, pp 225–234

Wahl V, Brand LH, Guo YL, Schmid M (2010) The FANTASTIC FOUR proteins influence shoot meristem size in *Arabidopsis thaliana*. BMC Plant Biol 10:285. https://doi.org/10.1186/1471-2229-10-285

Wang J, Zhang Z (2021) Gapit version 3: boosting power and accuracy for genomic association and prediction. Genomics Proteomics Bioinformatics 19:629–640. https://doi.org/10.1016/j.gpb.2021.08.005

Wang K, He J, Zhao Y et al (2018) EAR1 negatively regulates ABA signaling by enhancing 2C protein phosphatase activity. Plant Cell 30:815–834. https://doi.org/10.1105/tpc.17.00875

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. https://doi.org/10.1007/s00122-015-2544-7

Xu C, Liberatore KL, MacAlister CA et al (2015) A cascade of arabinosyltransferases controls shoot meristem size in tomato. Nat Genet 47:784–792. https://doi.org/10.1038/ng.3309

Zhang M, Yuan B, Leng P (2009) The role of ABA in triggering ethylene biosynthesis and ripening of tomato fruit. J Exp Bot 60:1579–1588. https://doi.org/10.1093/jxb/erp026

Zygier S, Chaim AB, Efrati A, Kaluzky G, Borovsky Y, Paran I (2005) QTLs mapping for fruit size and shape in chromosomes 2 and 4 in pepper and a comparison of the pepper QTL map with that of tomato. Theor Appl Genet 111:437–445. https://doi.org/10.1007/s00122-005-2015-7

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