All-Flesh Tomato Regulated by Reduced Expression Dosage of AFF Through a Promoter SV Mutation

Lei Liu¹, Kang Zhang¹, JinRui Bai, Jinghua Lu, Xiaoxiao Lu, Junling Hu, Chunyang Pan, Shumin He, Jiale Yuan, Yiyue Zhang, Min Zhang, Yanmei Guo, Xiaoxuan Wang, Zejun Huang, Yongchen Du, Feng Cheng*, Junming Li*

Key Laboratory of Biology and Genetic Improvement of Horticultural Crops of Ministry of Agriculture, Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, China

¹ These authors contributed equally to this article.

*Correspondence: Junming Li (lijunming@caas.cn) and Feng Cheng (chengfeng@caas.cn)

Highlight: The sequence deletion that occurred in the cis-regulatory region of AFF—the core node of locule tissue liquefaction determined here—reduced its expression dosage and produced all-flesh tomato fruit.
ABSTRACT

The formation of locule gel is an important process in tomato and a typical characteristic of berry fruit. In this study, we collected a tomato natural mutant that produces all-flesh fruits (AFF) in which the locule tissue remains in a solid state during fruit development. We built genetic populations to fine-map the causal gene of the AFF trait and identified the gene AFF (SlMBP3) as the locus conferring the locule gel formation. We determined the causal mutation as a 416-bp deletion that occurred in the promoter region of AFF and reduced its expression dosage. The 416-bp sequence is highly conserved among Solanaceae species, as well as within the tomato germplasm. Furthermore, with the BC₆ NIL materials, we revealed that the reduced expression dosage of AFF did not impact the normal development of seeds but produced unique non-liquefied locule tissue, which was distinct from that of normal tomatoes in terms of metabolic components. We further revealed the importance of AFF gene in locule tissue liquefaction through combined analysis using mRNA-seq and metabolomics. Our findings provide clues to investigate fruit type differentiation in Solanaceae crops and also contribute to the application of the AFF gene in tomato breeding programs.

Key Words: locule gel; all-flesh fruit; structural variant; cis-regulatory mutation; dosage effect; processing tomato

Abbreviations: AFF, all-flesh fruit; NIL, near-isogenic line; WT, wild-type; PG, polygalacturonase; PME, pectin methylesterases; SV, structural variant; TGRC, Tomato Genetics Resource Center; KASP, Kompetitive allele specific PCR; π, nucleotide diversity; DAF, days after flowering; CDS, coding sequence; PCA, principal component analysis; HCA, hierarchical cluster analysis; PCC, Pearson correlation coefficient; KEGG, Kyoto Encyclopedia of Genes and Genomes; MG, mature green; BR, breaker ripe; RR, red ripe; BSA-seq, bulked segregant analysis sequencing; GO, Gene Ontology; XET, xyloglucan endo-transglycosylase.
INTRODUCTION

Locule gel is a typical characteristic of berry fruit. As a model plant for the study of fruit development and ripening, tomato (*Solanum lycopersicum*) fruit has clear tissue distribution and structure (Czeredni et al., 2015; Huber and Lee, 1986; Joubès et al., 1999). A large body of information on tomato fruit development has been documented, but this information is mainly focused on fruit type, fruit weight, and fruit ripening. Recent studies have comprehensively described the molecular activities of developing locules through large-scale transcriptional analyses as well as physiological and biochemical surveys (Lamia et al., 2015; Lemaire-Chamley et al., 2005; Lin et al., 2014; Seymour et al., 2008; Shinozaki et al., 2018; Zhu et al., 2018). However, the regulation of locule gel formation and development remains unclear.

Tomato locule tissue, which is the second-most abundant tissue in tomato fruit, represents 23% (w/w) of the fruit fresh weight (Mounet et al., 2009). The formation of locule tissue has been proven to be a complex process involving a series of physiological and biochemical changes that play a critical role in fruit growth and maturation (Lamia et al., 2015; Lemaire-Chamley et al., 2005; Mounet et al., 2009). Generally, tomato locule tissue derives from the placenta and grows up around the ovules (Davies and Cocking, 1965), encloses the developing seeds, undergoes extensive expansion and liquefaction, and transforms into a jelly-like homogenous tissue that is composed of thin-walled giant cells (Atherton and Rudich, 1986; Cheng and Huber, 1996; Joubès et al., 1999). However, the regulation of its differentiation and formation during the development of tomato fruit remains unknown. The natural mutant ‘all-flesh fruit’ (named as AFF) tomato does not produce locule gel (jelly-like tissue), which completely alters the structure of locule tissue in tomato fruit (Macua et al., 2015; Silvestri, 2006). This might provide an ideal material to uncover the regulation mechanism of locule development. This mutant also offers several advantages for the tomato processing industry such as its high solid content, improved firmness, long shelf-life, and color compared to wild-type (WT) tomato (Macua et al., 2015; Silvestri, 2006). Therefore, the exploration of AFF may be quite important, not only for the elucidation of berry fruit formation, but also for breeding programs.

Commonly, phytohormones and cell wall-modifying enzymes have been considered to play important roles in locule gel formation. Evidence clearly indicates that IAA, GA, and ABA are present at high levels in seeds and are transported to the surrounding tissues to
participate in inducing and regulating the development of locule tissue (Kumar and Khurana, 2014; Lemaire-Chamley et al., 2005; Mounet et al., 2009; Sofia et al., 2007). However, it has been verified that ethylene and IAA do not control the determination and liquefaction of locule gel in tomato fruit (Brecht, 1987; Gillaspy et al., 1993; Qin et al., 2012). Rather, the formation of locule gel might be related to the ripening and softening of fruit because the locule gel development progresses along with the dissolution of pectin, deglycosylation, and hemicellulose—the main components of the cell wall matrix—catalyzed by polygalacturonase (PG) and pectin methylesterase (PME) (Bapat et al., 2010; Cheng and Huber, 1997; Nunan et al., 1998). However, PG and PME mainly change the texture of fruit and do not determine the process of locule gel formation (Tieman et al., 1992; Uluisik et al., 2016). Hence, the initial formation of locule gel may involve a mechanism that is different from the classic phytohormones or PME–D-galacturonanase scenario.

The well-known floral ‘ABCDE’ model was established to elucidate the regulation of floral organ development and differentiation. The D-class genes contribute to the formation of the seeds, the ovule, and the funiculus and regulate the expansion and maturation of the carpel and fruit (Dreni and Kater, 2014; Itkin et al., 2010; Vrebalov et al., 2009). For example, as the first set of D-class MADS-box genes reported in petunia, FLORAL BINDING PROTEIN 7 (FBP7) and FBP11 are expressed specifically in ovule differentiation and also participate in seed and coat development (Angenent et al., 1995; Colombo et al., 1995). Another orthologous gene, SEEDSTICK (STK; previously AGL11), isolated from Arabidopsis, is also involved in the development of ovules and affects seed germination (Ezquer et al., 2016; Favaro et al., 2003; Pinyopich et al., 2003). Suppression of STK orthologous genes trigger seedless fruits in tomato and grape (Ocarez and Mejía, 2016), whereas overexpression of the tomato TAGL1 gene results in the dramatic modifications of flower and fruit organization (Huang et al., 2017). In addition, genes SHATTERPROOF1 (SHP1) and SHP2 act redundantly with STK in promoting ovule identity (Liljegren et al., 2000; Pinyopich et al., 2003). Similarly, in tomato, TAGL1, an SHP orthologous gene, controls fruit expansion and fleshiness (Vrebalov et al., 2009). More recently, the D-class gene AGAMOUS MADS-box protein 3 (SlMBP3)—a paralog of TAGL1—was found to modulate both placenta liquefaction and seed formation in tomato (Zhang et al., 2019). The RNAi and over-expression lines of SlMBP3 genes yielded fleshy fruit without locular gel and accelerated placenta liquefaction, respectively. Furthermore, the SlMBP3-RNAi lines produced malformed seeds that could not germinate, while the over-expression lines
generated larger seeds. In addition, as a common feature, all these D-class genes participate in seed development, such as the Arabidopsis gene STK, the mutant of which exhibits reduced seed germination efficiency (Ezquer et al., 2016); the tomato gene TAGL11, which maintains the production of seeds (Huang et al., 2017); and SlMBP3, whose RNAi plants cannot produce germinable seeds (Zhang et al., 2019). In contrast, the natural mutant AFF, as aforementioned, produces normal seeds with a high germination rate. Therefore, this AFF mutant without gel formation provides novel clues for addressing the genetic mechanism underlying ovule development, especially locule gel formation, by ruling out the negative impact on the seed development.

In this study, the AFF gene was fine-mapped by combining a genetic analysis and map-based cloning approach. We found that a novel structural variant (SV)—a 416-bp sequence deletion—occurred in the conserved cis-regulatory region of aff. This deletion suppressed the expression of AFF and produced the AFF phenotype. Furthermore, combined transcriptome and metabolome analyses were performed with the aff near-isogenic lines (NILs) to illustrate the regulatory pathways and the effects on the fruit quality of aff. The metabolic components showed a distinct difference between the mutant and the WT. These findings provide useful information for tomato breeding programs and novel insights into the evolution of berry fruits.

MATERIALS AND METHODS

Plant Materials and Crossing

Tomato (S. lycopersicum) plants were cultivated at the Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences (IVF-CAAS), Beijing, China, during the natural growing season and under greenhouse conditions. The seeds of S. lycopersicum cv. 06-790 and 09-1225 (all-flesh cultivar) were obtained from our own stocks (China National Vegetable Germplasm Bank at IVF-CAAS). The seeds of S. lycopersicum cv. LA4069, H1706 (used for the tomato genome sequencing project) (Tomato Genome Consortium., 2012) and Micro-tom tomatoes were obtained from the Tomato Genetics Resource Center (TGRC) at the University of California, Davis (USA).

The aff line 06-790 was crossed with the WT LA4069 to generate F₁ progeny, and F₂ progeny were derived from self-pollination of the F₁ progeny. The F₁ progeny were crossed with 06-790 to generate BC₁P1 progeny and crossed with LA4069 to generate BC₁P2
progeny. All six populations of two crosses were grown for genetic analysis in the greenhouse in the spring of 2015.

The BC$_2$S$_1$ progeny were developed from the aff line 06-790 as donor parents, with continued backcrossing to H1706. The aff individuals of BA-130 and BA-150 and the normal individuals of BA-124 and BA-128 were selected for qRT-PCR.

The NILs of aff were derived from BC$_6$S$_2$ plants generated by 06-790 continual backcrossing to H1706. Among the NILs, BA-1 and BA-2 are aff lines and BA-4 and BA-6 are normal lines, and these were used for seed germination. BA-1 and H1706 were also used for morphology research as well as transcriptome and metabolome profiling.

**Paraffin Sectioning**

The locule tissue was cut into 1 mm×2 mm×2 mm cuboids and fixed in FAA (5% acetic acid, 5% formaldehyde, 50% ethanol, 5% glycerin mixture) for 24 h at room temperature. After dehydration, embedding, slicing, and pretreatment, sections were dyed using safranine and fast green double dye (Wang and He, 2004). The paraffin sections were visualized and photographed with an OLYMPUS IX71 microscope.

**Genome Sequencing, SNP and SV Calling, and KASP**

For rapid identification of the mutation conferring all-flesh fruit in 06-790, we used MutMap, a method based on the whole-genome resequencing of bulked DNA of F$_2$ segregants (Takagi et al., 2013). We designed two mixed DNA pools that combined the 30 F$_2$ progeny that had either the AFF phenotype or the normal phenotype. The DNA pools were subjected to whole-genome resequencing using an Illumina GAIIx DNA sequencer at Beijing Berry Genomics Co., Ltd. The sequencing depth was approximately 20-fold coverage for the two parental lines and approximately 30-fold coverage for the two mixed DNA pools. The paired-end reads of 06-790, LA4069, and the mixed DNA pools were mapped to the tomato reference genome (SL4.0 build; Tomato Genome Consortium, 2012) using Burrows-Wheeler Aligner (Version 0.7.10- r789) with default parameters (Li and Durbin, 2009). The BAM files were further deduplicated using the MarkDuplicate function of Picard (http://broadinstitute.github.io/picard/). The HaplotypeCaller function of GATK was used to call the variants with the default parameters (Mckenna et al., 2010). Variants supported by less than three reads were filtered out. ANNOVAR was used to annotate the retained variants (Wang et al., 2010). The ∆SNP-index was calculated based on the 200-kb sliding window.
with a 20-kb increment. The potential SVs of the aff line were called using BreakDancer (Version 1.1.2, http://gmt.genome.wustl.edu/breakdancer/current/) based on the BAM file. A total of 24 SNPs with strong associations were selected to develop KASP markers, and the SNP polymorphic analysis and genotyping of populations were conducted using the KASP genotyping system (LGC Genomics). Among these SNPs, 13 SNPs that exhibited polymorphism were used for fine mapping (Supplementary Table S1).

**Promoter Sequence Conservation of AFF Orthologous Genes**

Syntenic orthologous genes of AFF among Solanaceae crop species, *S. lycopersicum*, *S. pennellii*, *S. tuberosum*, *S. melongena*, and *Capsicum annuum*, were determined by the SynOrths tool (Cheng et al., 2012); they were *Sopen06g023350*, *Sotub06g020180*, *Sme2.5_02049.1_g0007.1*, and *Capang01g002169*. Then, 5-kb upstream sequences (promoter region) of each of the five orthologous genes were extracted from the genomes of the five species. These sequences were further aligned by MUSCLE (Edgar, 2004). Aligned sequences were further submitted to calculate the conservation level of each aligned nucleotide and then averaged by a 50-bp sliding window with a step of 10 bp using an in-house Perl script (available upon request).

We further investigated the sequence conservation of the AFF gene in the tomato germplasm using the published variome datasets of 360 tomato samples (Lin et al., 2014). We calculated the nucleotide diversity (\(\pi\)) values for the 3-kb upstream region, gene body, and 3-kb downstream region for all 34,075 tomato genes in the genome of *S. lycopersicum* with the variome datasets using the VCFtools (Danecek et al., 2011). The distributions of the \(\pi\) values in the three regions were further plotted as bean-plots with the R package “beanplot” (Kampstra, 2008).

**RNA Extraction and qRT-PCR**

With the use of specific primers and probes, different tomato lines were detected by real-time PCR. They included aff lines BA-130, BA-150, 06-790, and 09-1225, WT lines BA-124, BA-128, LA4069, and H1706, and F1 progeny from the crossing of 06-790 and H1706. They were grown in the greenhouse in the autumn of 2017, and RNA was collected from locule tissues at seven days after flowering (DAF), 10 DAF, 15 DAF, and 25 DAF, with three samples taken per line. *SIFRG27* (*Solyc06g007510*), *SIFRG03* (*Solyc02g063070*), and *ACTIN* (*Solyc11g005330*) in tomato were selected as the reference genes, with the primer sequences provided in Supplementary Table S2 (Cheng et al., 2017). The primer sequence of AFF was
F (5ʻ–3ʼ): GCATCTGGTTGGTGAAGG; R (5ʻ–3ʼ): ATCTGATTCTGCTGATGCC. The primers were designed by Roche LCPDS2 software and synthesized by Beijing TsingKe Biological Technology Co., Ltd. The cDNA was obtained from total RNA using a PrimeScript RT reagent kit, a reverse transcription kit produced by Takara Bio Inc. The qRT-PCR was completed on an ABI Prism®7900 qRT-PCR operating system produced by Applied Biosystems, according to the instructions of the SYBR Prime Script RT-PCR kit. The qRT-PCR and 2^{ΔΔCt} method were used to analyze the expression of the selected genes.

Gene Knockout and Overexpression Assays

To knockout the AFF gene, two sgRNAs targeting the second and the third exon of AFF were designed and constructed into the CRIPSPR/Cas9 expression vector BGK012-DSG to obtain the recombinant plasmids MSG8124/8125. The plasmids were introduced into the cotyledon explants of S. lycopersicum cv. Micro-tom (WT) through the Agrobacterium tumefaciens-mediated transformation method as described previously (Sun et al., 2006). The transgenic plants were confirmed by genotyping PCR using Sanger sequencing. For the AFF overexpression assay, the full-length coding sequence (CDS) of AFF was cloned into vector pEXT06/g to construct the recombinant plasmid 35S::AFF-CDS::GFP. The plasmid was further introduced into Micro-tom to obtain transgenic plants with over-expression of the Solyc06g064840 gene.

Relative Activity of Luciferase

To confirm the function of the 416-bp deletion in promoting the expression of the associated gene, a dual luciferase reporter gene assay was used to check the expression difference. Based on the PCR-based accurate synthesis method, full-length splicing primers were designed and the protective base synthesis gene promoters (Del and WT) designed at both ends of the primers were inserted into sites between PvuII and KpnI in plasmid pGreenII 0800-luc. The recombinant plasmid pGreenII 0800-luc-promoter (Del) was transferred into the epi400 clone strain, and the recombinant plasmid pGreenII 0800-luc-promoter (WT) was transferred to the Top10 clone strain. The sequence of the recombinant plasmid was verified by the sequence of the positive clones.

Monoclones were selected for PCR verification after plasmid transformation. Nicotiana benthamiana leaves (one month old) were transiently infected by positive strains using an Agrobacterium-mediated method. Each group had three replicates. The activity of the dual luciferase reporter gene was detected after three days. The transcriptional regulation was
determined by the activity ratio of firefly luciferase and Ranilla luciferase, that is, the relative activity of luciferase.

Metabolic Network (MMN) Analysis Using Transcriptome and Metabolome Profiling

Metabolome profiling was carried out using a widely targeted metabolome method by Wuhan Metware Biotechnology Co., Ltd. (Wuhan, China) (http://www.metware.cn/). Briefly, the tomato tissues were lyophilized and ground into fine powder using a mixer mill (MM 400, Retsch) with a zirconia bead for 1.5 min at 30 Hz. Then, 100 mg tissue powder was weighed and extracted overnight with 1.0 mL 70% aqueous methanol at 4°C, followed by centrifugation for 10 min at 10,000 g. All supernatants were collected and filtered with a membrane (SCAA-104, 0.22 mm pore size; ANPEL, Shanghai, China, http://www.anpel.com.cn/) before LC-MS analysis. The quantification of metabolites was carried out using a scheduled multiple reaction monitoring method (Wei et al., 2013; Zhu et al., 2018). In the data analysis process, unsupervised PCA (principal component analysis) was performed by the function prcomp within R (Version 3.5.0, www.r-project.org). The data were unit-variance scaled before performing unsupervised PCA. The HCA (hierarchical cluster analysis) results of samples and metabolites were presented as heatmaps with dendrograms, while Pearson correlation coefficients (PCC) between samples were calculated by the cor function in R. Both HCA and PCC were carried out using the R package heatmap (Version 1.0.12). Identified metabolites were annotated using the Kyoto Encyclopedia of Genes and Genomes (KEGG) compound database (http://www.kegg.jp/kegg/compound/); annotated metabolites were then mapped to the KEGG pathway database (http://www.kegg.jp/kegg/pathway.html). Pathways with significantly regulated metabolites were then fed into MSEA (Xia and Wishart, 2010).

For the RNA-seq experiments, a total of 3 µg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using the NEBNext Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer’s recommendations, and index codes were added to attribute sequences to each sample. The constructed libraries were then sequenced on an Illumina Hiseq platform, and 125 bp/150 bp paired-end reads were generated. Transcriptome profiling was performed as described previously (Ying et al., 2020). Briefly, clean reads were obtained using a Hiseq-XTen sequencing platform, mapped to the tomato reference genome (Version 4.0) using Hisat 2 (Daehwan et al., 2015), and then normalized to TPM (tags per million reads) reads by StringTie (Pertea et al., 2015).
Samples under different combinations were analyzed by MeV (Version 4.9) with the $k$-means method (Gasch and Eisen, 2002). The normalized expression values of genes and metabolites were calculated by dividing their expression level at different time points and in different tissues. HCA and PCA were performed to facilitate graphical interpretation of relatedness among different time points/tissue samples. The transformed and normalized gene and metabolite expression values with z-scores were used for HCA and PCA. We used Pearson’s correlation algorithm method (Bishara and Hittner, 2012) to construct a transcription factor-related gene and metabolite regulatory network. Mutual information was used to calculate the expression similarity between the expression levels of transcription factors and genes, and metabolite pairs were calculated with R software. All the associations among transcription factors, genes, and metabolites were analyzed by Cytoscape software (Kohl et al., 2011).

**Phylogenetic Analysis**

Protein sequences of genes that were homologous to AFF were collected from the NCBI database through the BLAST service. These protein sequences were then aligned by Clustal Omega (Fábio et al., 2019), and the multiple alignments were further used to construct the phylogenetic tree using IQ-TREE with the maximum likelihood method (Nguyen et al., 2015). The protein structure analysis was performed with the assistance of the Pfam server (Finn et al., 2008).

**RESULTS**

**The All-Flesh Fruit Trait is Controlled by a Single Recessive Locus**

To investigate the genetic characteristics of aff, we built an F$_2$ population using the aff genotype 06-790 as parent P1 and WT LA4069 as parent P2. The fruit traits of the F$_2$ population showed that the ratio of WT to aff samples was 150:41, consistent with a 3:1 segregation (chi-squared test: $\chi^2 = 1.176$, non-significant) (**Table 1**), suggesting a single recessive genetic model of the AFF trait. To further confirm the result, we built a BC$_1$P1 population. The progenies showed that the ratio of WT to aff samples was 46:42, conforming to a 1:1 segregation ratio (chi-squared test: $\chi^2 = 0.102$, non-significant) (**Table 1**). We also built a BC$_1$P2 population, and all progenies of this population had WT fruits. In conclusion, these data together confirmed that the AFF trait is controlled by a single recessive mutation.
The Locule Cell in *aff* Maintains a Complete Structure during Fruit Ripening

To determine the time point for the initiation of locule gel development, we observed the difference in locule tissue between *aff* and WT by crosscutting fruits at an interval of every five days. We found that no jelly-like tissue formed in the locule cavity area during the whole development process of the *aff* genotype (Fig. 1A). In contrast, obvious jelly-like tissue was observed in WT 25 DAF and reached complete liquefaction after the mature green (MG) stage (Fig. 1B). These findings indicated that 25 DAF is an important time point for the formation and development of locule gel in tomato fruit. We further examined the structure of the locule tissues of the 25 DAF samples of the WT and *aff* fruits using the paraffin sectioning method followed by microscopic examination. The individual locule cells of the WT fruit continued to collapse and showed a tendency to fracture intercellularly within the plane of the cell wall at the MG stage (Fig. 1B), which was consistent with previous reports (Cheng and Huber, 1996; Lemaire-Chamley et al., 2005). However, these distinct changes did not occur in cells of the *aff* locule; the locule cells of *aff* maintained a complete structure instead. This made the morphology of the locule tissue in *aff* tomato more like that of the placenta tissue (Fig. 1A; Supplementary Fig. S1).

A Large Sequence Deletion is Identified in the Promoter of the *aff* Gene

Bulked segregant analysis sequencing (BSA-seq) was applied to locate the *AFF* gene. Using the genome sequence of *S. lycopersicum* (SL4.0 ITAG4.0) as the reference, we called out 298,942 SNPs that were polymorphic between P1 and P2 while homozygous in each of the two parental genomes. These SNPs were further used in SNP index analysis (Takagi et al., 2013) with the *aff* and the WT pools of the F2 population. We detected a significant signal (Δindex=1.448, above the 99% confidence level) located between 37.25 Mb and 37.75 Mb on chromosome six (Fig. 2A), and 21 SNPs were located in this region. The average SNP indexes in this region for *aff* and WT pools were 0.98 and 0.32, respectively.

Linkage analysis of two populations (F2 with 215 individuals and BC2S1 with 249 individuals) was used to fine-map the *AFF* gene. Molecular markers (Supplementary Table S1) were selected from these polymorphic SNPs and SVs between P1 and P2, and were genotyped by PCR and KASP. The linkage signals of *AFF* overlapped with those of BSA-Seq (Fig. 2B). With the genotype of these markers in the BC2S1 population, the *AFF* gene was finally mapped between markers SNP-14 and SNP-15 (SL4.0ch06: 37,945,500–38,129,705), which was about 184.2 kb and harbored 27 genes (Fig. 2C–D; Supplementary
Table S3). A 416-bp deletion was found (SL4.0ch06: 38,062,128–38,062,543), located 1,775 bp upstream of the gene Solyc06g064840 (Fig. 2E), which contains no sequence polymorphisms in the coding sequences between aff and WT. A marker named SV-12 was designed with the 416-bp deletion and it showed complete co-segregation with all-flesh individuals (Fig. 2F). This deletion-associated gene Solyc06g064840 is a member of the AGAMOUS family belonging to the MADS-box D-class genes and is also named SlMBP3 (Zhang et al., 2019). The gene is specifically expressed in the developing locule (include the seeds) of tomato (Supplementary Fig. S2–3) (Fernandez-Pozo et al., 2017; Koenig et al., 2013).

Gene Editing of AFF Confirmed its Function in Tomato Locule Gel Formation

To prove that Solyc06g064840 is the causal gene of the aff genotype, we generated knockout mutations using the CRISPR/Cas9 system with two sgRNAs (Fig. 3A) that targeted the second and third exon of the AFF gene. These transgenic plants were confirmed by PCR amplification and DNA sequencing (Fig. 3B). We evaluated the first-generation (T0) diploid lines that were homozygous on edited mutant alleles; the aff-cr mutants produced the expected all-flesh fruits without the locule gel and normal seeds as in the WT (Fig. 3C). We also generated transgenic plants with over-expression of the Solyc06g064840 gene driven by the 35S promoter. As shown in Fig. 3C, the AFF-overexpression transgenic T1 homozygous lines developed more locule gel compared to the normal locule gel in the WT Micro-tomato fruit. These results indicate that the AFF gene does possess the key function in locule gel formation in tomato fruit.

The Deleted Promoter Sequence Shows Strong Conservation

To understand the detailed function of the 416-bp deletion, the 2-kb sequence (416-bp deletion included) was analyzed by the promoter prediction tool TSSP in the PlantProm DB database (Shahmuradov et al., 2003) and the PlantCARE database (Lescot et al., 2002). We found that the 416-bp deletion in the promoter region of Solyc06g064840 and functional elements including the TATA box and CAAT box (Supplementary Table S4) were located in this deleted sequence. Therefore, it is interesting to investigate the conservation status of the 416-bp sequence across Solanaceae crops. We selected five genomes from four Solanaceae crops, including two tomato genomes, S. lycopersicum and S. pennellii, as well as the genomes of potato (S. tuberosum), capsicum (C. annuum), and eggplant (S. melongena). We then determined the syntenic orthologous genes of AFF in the five selected genomes.
(Methods), which were Solyc06g064840, Sopen06g023350, Sotub06g020180, Capang01g002169, and Sme2.5_02049.1_g00007.1, respectively. The promoter sequences of these five syntenic genes were extracted from corresponding genomes and aligned using MUSCLE (Edgar, 2004). Based on the results of multiple sequence alignment, we estimated the conservation level of these promoter sequences. As shown in Fig. 4A, using the revised \( \pi \) as the measure and a threshold of 0.3, we determined that five main local regions showed a relatively higher conservation level (low mismatch ratio in multiple sequence alignment) in the promoter sequences of these Solanaceae crops. These five regions should have important roles in regulating the expression of associated genes. Moreover, the 416-bp sequence deletion was located at one of the two most conserved regions (the grey bar in Fig. 4A). This suggests that the deletion may have a large effect in altering the expression of the gene \( AFF \) in \( aff \) tomato.

We further investigated the conservation of the promoter region within the tomato germplasm by analyzing its sequence diversity using the previously published variome dataset of 360 tomato accessions (Lin et al., 2014). Generally, we estimated the sequence diversity (selective sweep) by calculating \( \pi \) for the gene body, 3-kb upstream, and 3-kb downstream regions for each of the 34,075 tomato genes and checked the selection strength, i.e., the diversity level of the \( AFF \) gene under the background of all tomato genes. For \( AFF \), the \( \pi \) value of the gene body was 0.21, slightly larger than the mode value of all genes, while its downstream region had a \( \pi \) value of 0.46, which indicated higher diversity, whereas its upstream region had a \( \pi \) value of 0.026, less than 93.80% of all other genes (Fig. 4B). This suggests that the promoter region of \( AFF \) gene might undergo stronger selection pressure against mutations than many other genes in the tomato germplasm. Together, these findings support the importance of sequence conservation in the promoter region of \( AFF \), which further suggests that the 416-bp deletion may have a significant impact on the function of \( AFF \).

The Deletion in the Promoter Region Down-regulates the Expression Level of the Gene \( AFF \)

We investigated the expression of \( AFF \) by quantitative real-time PCR assay and the dual luciferase reporter system. First, we performed qRT-PCR analysis to measure the expression variation of \( AFF \) in different developmental stages in locule tissues from the four BC\(_2\)S\(_1\) lines, together with their parental materials P1, H1706, and F\(_1\), as well as another \( aff \) line 09-
1225 and the WT line P2 (LA4069). As shown in Fig. 5A, in all of these samples, the gene was highly expressed at seven DAF and 10 DAF, and significantly decreased on 15 DAF, which was synchronous with the differentiation of locule tissues, and was consistent with previous reports (Fernandez-Pozo et al., 2017; Koenig et al., 2013). More importantly, the expression of AFF was significantly lower in aff samples than in WT samples. The aff BC$_2$S$_1$ lines BA-130 and BA-150 had a significantly lower expression of AFF than the WT BC$_2$S$_1$ lines BA-124 and BA-128 (Fig. 5A). We further evaluated the transcriptional activity of the promoter sequences in WT and aff samples, i.e., the promoter sequences with or without the 416-bp deletion, through experiments on the relative activity of luciferase (the ratio of luc to Rluc). The relative activity of luciferase in the aff promoter was significantly lower than that of the WT promoter (Fig. 6A).

Furthermore, we examined the locule tissues of aff and WT tomato fruits using the NILs. The NILs were generated by back-crossing the aff tomato material 06-790 to H1706 for six generations (Methods), assisted by molecular selection of marker SV-12. These aff lines all produced all-flesh fruit, which is distinct to the WT line H1706 (Fig. 5B), in which the locule tissues of all-flesh tomatoes maintain a solid state during fruit development. Additionally, we checked the seed characteristics, thousand-seed weight, and seed germination activity using the aff NILs; the seed structure or appearance did not differ between the aff fruit lines and the WT fruit lines (Supplementary Fig. S4). As shown in Fig. 6B–D, the aff lines BA-1 and BA-2 had a similar thousand-seed weight, germination index, and germination rate to those of WT fruit lines BA-4 and BA-6. These results suggest that the deletion stops the formation of gel in tomato but does not impact the function of SlMBP3/AFF involved in the normal development of seeds. The deletion mutation in the aff was different from that of the aff-cr5 plants that cannot produce seeds and the SlMBP3 RNAi plants whose seeds cannot germinate (Zhang et al., 2019).

The aff Mutation Largely Alters Gene Expression and Metabolic Components

A low dosage of AFF had an impact on systematic gene expression variations in the locule tissue of aff; more genes showed down-regulated expression. We compared whole genome gene expression patterns between the tomato material HZ106 (WT) and its NIL BA-1 (aff line), whose AFF gene was replaced by the mutated one with the 416-bp deletion in its promoter region. mRNA-seq analyses were performed on two tissues, the locule and placenta, for both the WT and aff lines, at three time points (10, 15, and 25 DAF). We found...
that genes belonging to GO (Gene Ontology) terms including lipid metabolism, plant-type cell wall, phyto-hormones, metabolism and catabolism, flavonoid biosynthesis, glucosyltransferase activity, and nutrient reservoir activity were enriched in these differentially expressed gene sets (Supplementary Table S5) between WT and aff, as well as KEGG pathways including sugar metabolism and phyto-hormone biosynthesis (Supplementary Table S6). Among the top 50 enriched GO terms, 1,110 genes were down-regulated, while only 359 genes were up-regulated (Supplementary Table S5). In the KEGG pathway ‘MAPK signaling’, 55 differentially expressed genes were involved, with 42 genes down-regulated and 13 genes up-regulated. These results clearly show a large number of genes whose expression was altered—mostly down-regulated—along with the reduced expression of AFF in the aff line. Furthermore, we performed detailed comparisons between pair-wise transcriptome datasets. When comparing gene expressions between locule and placenta tissues in WT (group 1), we found that genes involved in GO terms including lipid transport, apoplast, flavonoid metabolism, transferase, and hydrolase activity, or KEGG pathways including metabolism, protein kinase, and phyto-hormone were enriched (Supplementary Fig. S5). However, the GO terms lipid transport and flavonoid metabolism were not enriched in differentially expressed genes between the locule and placenta tissues in aff (group 2). Additionally, there was over-representations of GO terms such as phloem or xylem, as well as symporter activity or transmembrane transmission-related, in genes that were differentially expressed in group 2 (Supplementary Fig. S6). To focus on the differences between the locule tissues from WT and aff (group 3), we further compared their differentially expressed genes and found that similar GO terms or KEGG pathways to those observed in group 1 were enriched (Fig. 7A–B). Moreover, the GO terms DNA replication, plasma membrane, photosystem II, plant-type cell wall, glucosyltransferase activity, as well as nutrient reservoir activity were specifically enriched in group 3 (Supplementary Table S7; Fig. 7A). Additionally, we also compared the gene expression differences between placenta tissues of WT and aff lines (group 4). However, the aforementioned enriched GO terms or KEGG pathways were not observed (Supplementary Fig. S7). These results together suggest that the reduced expression of AFF is associated with the down-regulated expression of genes involved in DNA replication, phyto-hormone metabolism, photosynthesis, sugar metabolism, and MAPK signaling, which might then prevent the locule liquefaction process that normally occurs in WT tomato.
Genes involved in the development of locule tissues showed strong expression differentiation between locule tissues of WT and *aff* tomato. Genes involved in hydrolases, phyto-hormone metabolism, and DNA replication have been reported to regulate the locule tissue liquefaction of tomato fruit (Christian et al., 2014; Huber and Lee, 1986; Mounet et al., 2009; Takizawa et al., 2014; Uluisik et al., 2016). A total of 188 genes showed strong and stable differential expression between WT and *aff*, of which 122 were down-regulated and 66 were up-regulated in *aff* (Supplementary Table S8). Many of these differentially expressed genes were locule tissue liquefaction process-related genes. First, six gibberellin-related genes were down-regulated in *aff*. Among them, four are gibberellin-regulated proteins, while one is involved in the gibberellin biosynthesis process, and the last one is involved in the gibberellic acid-mediated signaling pathway (Supplementary Table S8). We also found six auxin-related genes that were differentially expressed, with the auxin repressed protein up-regulated, while the other five genes, auxin transporters, auxin responsive protein IAA9, auxin-related genes from the GH3 gene family, and those involved in the auxin signaling pathway, were down-regulated in *aff*. Second, there were three copies of cytochrome P450 genes whose expressions were down-regulated in *aff* (Supplementary Table S7), indicating a low level of energy-related activity in *aff*. Third, pectinesterase, which is strictly regulated and functions in the softening of tomato fruit, was largely down-regulated in *aff*. This may have an important impact on the solidness of *aff* tomato fruits. Fourth, we identified two copies of xyloglucan endo-transglycosylase (XET)-related genes that were down-regulated in *aff*. XET is involved in the induction of fruit ripening and softening, and its down-regulation should hinder the softening of *aff* tomato fruits. Fifth, we found three copies of glycosyl hydrolase genes that were down-regulated in *aff*, indicating the suppressed metabolism of glycolysis in *aff* compared with WT. More importantly, we found that TAG1 (Solyc02g071730) and TAGL1 (Solyc07g055920), both of which are paralogs of *AFF* and show high levels of sequence homology (Supplementary Fig. S8), were up-regulated in the locule tissue of *aff* compared with WT. TAGL11 (Solyc11g028020), the paralog with the highest homology to *AFF* in tomato (Huang et al., 2017) (Supplementary Fig. S8), showed no expression difference between *aff* and WT.

Our metabolic data support the results observed in the mRNA-seq analysis. We measured the metabolites and their quantities in tomato fruits of WT and *aff* (Methods). PCA of metabolites from the WT and *aff* (Fig. 7C) showed that the placenta tissues from WT and *aff* had similar metabolic components. However, the metabolites of locule tissue were
different from those of the placenta tissue in WT or *aff*. More importantly, the pattern of locule metabolites in *aff* was located between that of the placenta and locule tissues of WT, which indicated that the down-regulated expression of *AFF* changed the metabolic components of the locule tissue in *aff*. Furthermore, we investigated the metabolites whose contents were changed in *aff* compared to those in WT and found higher levels of flavonoids and lipids (Supplementary Table S9) in *aff*, whereas there were more alkaloids and phenolic acids (Supplementary Table S9) in WT than in *aff*. The differences in the metabolic components were caused by the down-regulated expression of *AFF* and the associated large-scale gene expression variations, which further resulted in the distinct fruit quality of the *aff* tomato compared to the WT tomato.

**DISCUSSION**

Locule gel liquefaction is not only a significant process in development and ripening but also a typical characteristic of tomato fruit. In this study, the causal gene *AFF* of the *AFF* trait and the 416-bp deletion mutation in the cis-regulatory region of *AFF* were identified. We further found that the expression dosage of *AFF* was crucial for locule tissue liquefaction, which coincided with the gene functional characterization of *AFF/SIMBP3* in a recent study (Zhang et al., 2019). *AFF* belongs to the AGAMOUS subfamily and contains a typical MADs-box domain; its paralogous genes in tomato are *TAG1*, *TAGL1*, and *TAGL11*, which have a high sequence homology to each other (Supplementary Fig. S8). These genes and their orthologs have been found to play important roles in ovule differentiation and formation, participate in seed and coat development, or regulate the expansion and ripening processes of the carpel and fleshy fruit in many species (Angenent et al., 1995; Colombo et al., 1995; Favaro et al., 2003; Itkin et al., 2010; Ocaez and Mejía, 2016; Pan et al., 2010; Vrebalov et al., 2009; Zhang et al., 2019).

The *AFF* gene plays a major role in locule liquefaction. Its function likely could not be compensated for by its paralogs *TAG1*, *TAGL1*, or *TAGL11*. The cis-regulatory sequence deletion mutation of the *AFF* gene is associated with the differential expression of many important genes. Among them, we observed that the expressions of *TAG1* and *TAGL1* were significantly up-regulated, accompanied by the down-regulated expression of *AFF*. Another paralog, *TAGL11* (Supplementary Fig. S8), which functions on early fleshy fruit development in tomato (Huang et al., 2017), showed stable expression between *aff* and WT. More importantly, considering the fact that both the up-regulated expression of *TAG1* and
TAGL1 and the stable expression of TAGL11 did not recover the development of normal liquefied locule tissue in the aff tomato, the function of mediating locule tissue liquefaction is likely to mainly belong to the gene AFF (Supplementary Fig. S8), though TAGL11 was indicated to have some function redundancy with AFF in seed development of tomato (Zhang et al., 2019; Huang et al., 2017). Furthermore, based on metabolomics analysis, we found that the pattern of locule metabolites in aff was located between that of the placenta and locule tissues of WT tomato. This finding indicates that tomato locule tissue is derived from the placenta and is formed from the development of the carpel (Davies and Cocking, 1965; Lemaire-Chamley et al., 2005; Pedro et al., 1991). The carpel development process is regulated by D-class genes in the ‘ABCDE’ flower development model, which is consistent with the fact that the locule gel develops along with the degradation of the cell wall matrix (Brecht, 1987; Joubès et al., 1999; Lemaire-Chamley et al., 2005).

The reduced expression dosage of the AFF gene caused by a 416-bp cis-regulatory deletion is the key factor that promoted the formation of the AFF trait. The dosage of gene expression has been proven to play an important role in the variation of plant traits, especially for the floral organ identity that determines genes. Up- or down-regulation of the expression of one ABCDE-class gene may easily shift the boundaries between different types of floral organs (Ito et al., 2007; Wang et al., 2016; Wuest et al., 2012). For example, a dosage imbalance between B- and C-class proteins can change stamen morphology (Liu et al., 2018), while the expression variation of TAGL1 and TAGL11 can also affect the development of tomato seeds and the fleshy characteristics of tomato fruits (Gimenez et al., 2016; Huang et al., 2017; Ocarez and Mejía, 2016; Vrebalov et al., 2009). In addition to floral-determining genes, another example showed that gene editing of different loci in the promoter region of tomato genes resulted in fruits with different sizes (Rodriguez-Leal et al., 2017). Meanwhile, the SV has been found to be a major genetic resource that can be used to employ gene expression dosage variations (Alonge et al., 2020). Unlike SNPs, the gene-associated SVs located in cis-regulatory regions always cause expression dosage changes of corresponding genes and further produce genetic and phenotypic changes. SV was recently reported to be involved in the formation of many traits in plants, and plays an important role in plant evolution and crop domestication (Alonge et al., 2020; Lye and Purugganan, 2019; Rodriguez-Leal et al., 2017). In our study, a 416-bp sequence deletion—a type of SV—that lies in the cis-regulatory region of AFF down-regulated the expression of AFF and led to its dosage effect as the AFF trait (Fig. 5 and 6). As exemplified in this study, SVs present as
useful quantitative variants, which might be used in next-generation breeding strategies through genetic engineering in the future (Alonge et al., 2020; Rodriguez-Leal et al., 2017; Swinnen et al., 2016).

The variation of AFF may also contribute to the evolution of fleshy fruit in Solanaceae, which provides insights into fruit type evolution in plants. Evidence obtained through archaeology and molecular biology has shown that fleshy fruit plants evolved from dry fruit plants, but the molecular mechanisms responsible for the shift from dry plants to fleshy fruit plants remains unknown (Kumar and Khurana, 2014; Maheepala et al., 2019; Seymour et al., 2008). Therefore, revealing the genetic basis and mechanism underlying the alteration process between fruit types is critical for understanding the evolution of biodiversity. However, the lack of intermediate or transition fruit types has limited research progress in this area (Annette et al., 2011; Wang et al., 2015). Comparative genetic analysis has shown that there are widespread genomic synteny and collinearity of genes among Solanaceae species, especially in Solanaceae vegetable crops (potato, tomato, capsicum, and eggplant), whose fruits show similar characteristics but are varied in many aspects. These differences in fruit development could be caused by gene expression or sequence variations of similar genes (Kim et al., 2014). For example, there is more locule gel in the wild tomato S. lycopersicum var. cerasiforme and S. pimpinellifolium than in cultivated tomato (Lemaire-Chamley et al., 2005). The example suggests a positive relationship between the quantity of liquefied locule tissues and the expression level of the AFF gene through the process of tomato domestication and breeding.

To summarize, the aff tomato, whose locule tissue changes from a jelly-like substance to a solid-state cavity, was found to be caused by an SV of a 416-bp sequence deletion in the cis-regulatory region of the AFF gene. The SV mutation reduced the expression dosage of AFF, which further impacted the normal liquefication process of locule tissue through the altered expression of many important genes and the subsequent changes in the metabolic components of tomato. Our findings are valuable for revealing the mechanism that underlies changes inside tomato fruit and shed new light on the evolution of berry fruits. In the future, with systematic studies on the dosage effects of AFF expression and extensive research on the formation and development processes of fruit locule tissues, the evolutionary mechanism of the berry fruits will be revealed in depth.
Funding Information

This work was supported by The National Key Research and Development Program of China (2016YFD0100204-05), the Fundamental Research Funds for Central Non-profit Scientific Institution (IVF-BRF2018006), the Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, China, and the Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (Grant No. CAAS-ASTIPIVFCAAS).

Author Contributions

J.L and L.L. designed and organized the study. L.L., J.B., J.L., X.L., J.H., C.P., S.H., J.Y., and M.Z. conducted the research. F.C., K.Z., and L.L. analyzed the data. All authors discussed and interpreted the results. L.L., F.C., and J.L. wrote the paper. L.L. agrees to serve as the author responsible for contact and ensures communication.

Acknowledgements

We thank Zhenxian Zhang and Wencai Yang (both China Agricultural University), and Jianchang Gao, Zhonghua Zhang and Xiaowu Wang (all Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences) for providing critical comments.

Competing Interests

The authors declare no competing financial interests.

Data Availability Statement

The mRNA-seq and metabolic data are openly available at http://www.bioinformaticslab.cn/files/tomato_AFF/.
References

Alonge M, Wang XG, Benoit M, et al. 2020. Major impacts of widespread structural variation on gene expression and crop improvement in tomato. Cell 182, 1-47.

Angenent GC, Franken J, Busscher M, et al. 1995. A novel class of MADS box genes is involved in ovule development in petunia. The Plant Cell 10, 1569-1582.

Annette B, Karine A, Catherine D. 2011. The evolution of flower development: Current understanding and future challenges. Annals of Botany 107, 1427-1431.

Atherton JG, Rudich J. 1986. The tomato crop: A scientific basis for improvement. London: Chapman and Hall.

Bapat VA, Trivedi PK, Ghosh A, et al. 2010. Ripening of fleshy fruit: Molecular insight and the role of ethylene. Biotechnology Advances 28, 94-107.

Bishara AJ, Hittner JB. 2012. Testing the significance of a correlation with nonnormal data: Comparison of Pearson, Spearman, transformation, and resampling approaches. Psychological Methods 17, 399-417.

Brecht JK. 1987. Locular gel formation in developing tomato fruit and the initiation of ethylene production. Hortscience 22, 476-479.

Cheng F, Wu J, Fang L, Wang X. 2012. Syntenic gene analysis between Brassica rapa and other Brassicaceae species. Frontiers in Plant Science 3, 198.

Cheng GW, Huber DJ. 1996. Alterations in structural polysaccharides during liquefaction of tomato locule tissue. Plant Physiology 111, 447-457.

Cheng GW, Huber DJ. 1997. Carbohydrate solutilization of tomato locule tissue cell walls: Parallels with locule tissue liquefaction during ripening. Physiologia Plantarum 101, 51-58.

Cheng Y, Bian WY, Pang X, et al. 2017. Genome-wide identification and evaluation of reference genes for quantitative RT-PCR analysis during tomato fruit development. Frontiers in Plant Science 8, 1440.

Christian C, Matthieu B, Julien P, et al. 2014. Endoreduplication and fruit growth in tomato: Evidence in favour of the karyoplasmic ratio theory. Journal of experimental Botany 65, 2731-2746.

Colombo L, Franken J, Koetje E, et al. 1995. The petunia MADS box gene FBP11 determines ovule identity. The Plant Cell 7, 1859-1868.

Czerednik A, Busscher M, Angenent GC, De Maagd RA. 2015. The cell size distribution of tomato fruit can be changed by overexpression of CDKAI. Plant Biotechnology Journal 13, 259-268.
Daehwan K, Ben L, Salzberg SL. 2015. HISAT: A fast spliced aligner with low memory requirements. Nature Methods 12, 357-360.

Danecek P, Auton A, Abecasis G, Albers CA, Anks EB, Depristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST. 2011. The variant call format and VCFtools. Bioinformatics 27, 2156-2158.

Davies JW, Cocking EC. 1965. Changes in carbohydrates, proteins and nucleic acids during cellular development in tomato fruit locule tissue. Planta 67, 242-253.

Dreni L, Kater MM. 2014. MADS reloaded: evolution of the AGAMOUS subfamily genes. New Phytologist 201, 717-732.

Ezquer I, Mizzotti C, Nguemaona E, et al. 2016. The developmental regulator SEEDSTICK controls structural and mechanical properties of the Arabidopsis seed coat. The Plant Cell 28, 2478-2492.

Fábio M, Mi PY, Joon L, Nicola B, Tamer G, Nandana M, Prasad B, Tivey A, Potter SC, Finn RD. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research 47, W636-W641.

Favaro R, Pinyopich A, Battaglia R, et al. 2003. MADS-Box protein complexes control carpel and ovule development in Arabidopsis. The Plant Cell 15, 2603-2611.

Fernandez-Pozo N, Zheng Y, Snyder SI, et al. 2017. The tomato expression atlas. Bioinformatics 33, 2397-2398.

Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL. 2008. The Pfam protein families database. Nucleic Acids Research 32, D138.

Gasch AP, Eisen MB. 2002. Exploring the conditional coregulation of yeast gene expression through fuzzy k-means clustering. Genome Biology 3, research0059.

Gillaspy G, Ben-David H, Gruissem W. 1993. Fruits: A developmental perspective. The Plant Cell 5, 1439-1451.

Gimenez E, Castañeda L, Pineda B, et al. 2016. TOMATO AGAMOUS1 and ARLEQUIN/TOMATO AGAMOUS-LIKE1 MADS-box genes have redundant and divergent functions required for tomato reproductive development. Plant Molecular Biology 91, 513-531.

Huang B, Routaboul JM, Liu M, et al. 2017. Overexpression of the class D MADS-box gene SL-AGL11 impacts fleshy tissue differentiation and structure in tomato fruits. Journal of experimental Botany 68, 4869-4884.
Huber DJ, Lee JH. 1986. Comparative analysis of pectins from pericarp and locular gel in developing tomato fruit. Acs Symposium Series. Acs Symposium Series 310, 141-156.

Itkin M, Seybold H, Breitel D, Rogachev I, Aharoni A. 2010. TOMATO AGAMOUS-LIKE 1 is a component of the fruit ripening regulatory network. The Plant Journal 60, 1081-1095.

Ito T, Ng KH, Lim TS, Meyerowitz YEM. 2007. The homeotic protein AGAMOUS controls late stamen development by regulating a jasmonate biosynthetic gene in Arabidopsis. The Plant Cell 19, 3516-3529.

Joubès J, Phan T, Just D, et al. 1999. Molecular and biochemical characterization of the involvement of Cyclin-Dependent kinase a during the early development of tomato fruit. Plant Physiology 121, 857-869.

Kampstra P. 2008. Beanplot: A Boxplot Alternative for Visual Comparison of Distributions. Journal Of Statistical Software 28, 1-9.

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. Nature Genetics 46, 270-278.

Koenig D, Jimenez-Gomez JM, Kimura S, et al. 2013. Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. Proceedings of the National Academy of Sciences of the United States of America 110, E2655-E2662.

Kohl M, Wiese S, Warscheid B. 2011. Cytoscape: software for visualization and analysis of biological networks. Methods in molecular biology 696, 291-303.

Kumar R, Khurana A. 2014. Functional genomics of tomato: Opportunities and challenges in post-genome NGS era. Journal of Biosciences 39, 917-929.

Lamia A, Cynthia D, Frédéric G, et al. 2015. Fruit growth-related genes in tomato. Journal of experimental Botany 66, 1075-1086.

Lemaire-Chamley M, Petit J, Garcia V, et al. 2005. Changes in transcriptional profiles are associated with early fruit tissue specialization in tomato. Plant Physiology 139, 750-769.

Lescot M, Dehais P, Thijs G, et al. 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research 30, 325-327.

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

Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Yanofsky MF. 2000. SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature 404, 766-770.
Lin T, Zhu G, Zhang J, et al. 2014. Genomic analysis provides insights into the history of tomato breeding. Nature Genetics 46, 1220-1226.

Liu J, Li CQ, Dong Y, Yang X, Wang YZ. 2018. Dosage imbalance of B- and C-class genes causes petaloid-stamen relating to F1 hybrid variation. BMC Plant Biology 18, 341.

Lye ZN, Purugganan MD. 2019. Copy number variation in domestication. Trends in Plant Science 24, 352-365.

Macua JI, Santos Á, Malumbres Á, et al. 2015. Processing tomato in Navarra - "All flesh" cultivars. Acta Horticulturae 1081, 175-180.

Maheepala DC, Emerling CA, Rajewski A, et al. 2019. Evolution and diversification of FRUITFULL genes in Solanaceae. Frontiers in Plant Science 10, 43.

Mckenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20, 1297-1303.

Mounet F, Moing A, Garcia V, et al. 2009. Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development. Plant Physiology 149, 1505-1528.

Nguyen LT, Schmidt HA, Haeseler AV, Minh BQ. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution 32, 268-274.

Nunan KJ, Sims IM, Bacic A, Robinson SP, Fincher GB. 1998. Changes in cell wall composition during ripening of grape berries. Plant Physiology 118, 783-792.

Ocarez N, Mejía N. 2016. Suppression of the D-class MADS-box AGL11 gene triggers seedlessness in fleshy fruits. Plant Cell Reports 35, 239-254.

Pan IL, Mcquinn R, Giovannoni JJ, Irish VF. 2010. Functional diversification of AGAMOUS lineage genes in regulating tomato flower and fruit development. Journal of experimental Botany 61, 1795-1806.

Pedro P, Gallego, Donald, J, Huber, Ignacio, Zarra. 1991. Cell wall hydrolysis in tomato locule gel formation. Hortscience 26, 735.

Pertea M, Pertea GM, Antonescu CM, et al. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature Biotechnology 33, 290-295.

Pinyopich A, Ditta GS, Savidge B, et al. 2003. Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature 424, 85-88.
Qin GZ, Wang YY, Cao BH, Wang WH, Tian SP. 2012. Unraveling the regulatory network of the MADS box transcription factor \textit{RIN} in fruit ripening. The Plant Journal 70, 243-255.

Rodriguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB. 2017. Engineering quantitative trait variation for crop improvement by genome editing. Cell 171, 470-480.

Seymour G, Poole M, Manning K, King GJ. 2008. Genetics and epigenetics of fruit development and ripening. Current Opinion in Plant Biology 11, 58-63.

Shahmuradov IA, Gammerman AJ, Hancock JM, Bramley PM, Solovyev VV. 2003. Plantprom: a database of plant promoter sequences. Nucleic Acids Research 31, 114.

Shinozaki Y, Nicolas P, Fernandez-Pozo N, et al. 2018. High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening. Nature Communications 9, 364.

Silvestri G. 2006. Mutant allele of tomato. US:US7112727.

Sofia M, Esra C, Yury T, et al. 2007. Tissue specialization at the metabolite level is perceived during the development of tomato fruit. Journal of experimental Botany 58, 4131-4146.

Sun HJ, Sayaka U, Shin W, Hiroshi E. 2006. A highly efficient transformation protocol for Micro-Tom, a model cultivar for tomato functional genomics. Plant & Cell Physiology 47, 426-431.

Swinnen G, Goossens A, Pauwels L. 2016. Lessons from Domestication: Targeting Cis-Regulatory Elements for Crop Improvement. Trends in Plant Science 21, 506-515.

Takagi H, Abe A, Yoshida K, et al. 2013. QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. The Plant Journal 74, 174-183.

Takizawa A, Hyodo H, Wada K, et al. 2014. Regulatory Specialization of Xyloglucan (XG) and Glucuronoarabininoxylan (GAX) in Pericarp Cell Walls during Fruit Ripening in Tomato (\textit{Solanum lycopersicum}). Plos One 9, e89871.

Tieman DM, Harriman RW, Ramamohan G, Handa AK. 1992. An antisense pectin methylesterase gene alters pectin chemistry and soluble solids in tomato fruit. The Plant Cell 4, 667-679.

Tomato Genome and Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485, 635-641.

Uluisik S, Chapman NH, Smith R, et al. 2016. Genetic improvement of tomato by targeted control of fruit softening. Nature Biotechnology 34, 950-952.
Vrebalov J, Pan IL, Arroyo AJM, et al. 2009. Fleshy fruit expansion and ripening are regulated by the tomato SHATTERPROOF gene TAGL. Plant Cell 21, 3041-3062.

Wang JF, He YM. 2004. Cell biology experimental course. Beijing: The Science Publishing Company.

Wang K, Li M Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research 38, e164.

Wang L, Li J, Zhao J, He C. 2015. Evolutionary developmental genetics of fruit morphological variation within the Solanaceae. Frontiers in Plant Science 6, 248.

Wang P, Liao H, Zhang W, et al. 2016. Flexibility in the structure of spiral flowers and its underlying mechanisms. Nature Plants 2, 15188.

Wei C, Liang G, Guo ZL, et al. 2013. A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: Application in the study of rice metabolomics. Molecular Plant 6, 1769-1780.

Wuest SE, O Maoileidigh DS, Rae L, et al. 2012. Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. Proceedings of the National Academy of Sciences of the United States of America 109, 13452-13457.

Xia J, Wishart DS. 2010. MSEA: A web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. Nucleic Acids Res 38, W71-W77.

Ying S, Su M, Wu Y, et al. 2020. Trichome regulator SIMIXTA like directly manipulates primary metabolism in tomato fruit, Plant Biotechnology Journal 18, 354-363.

Zhang J, Wang Y, Naeem M, et al. 2019. An AGAMOUS MADS-box gene, SlMBP3, regulates the velocity of placenta liquefaction and seed formation in tomato. Journal of experimental Botany 70, 909-924.

Zhu G, Wang S, Huang Z, et al. 2018. Rewiring of the fruit metabolome in tomato breeding. Cell 172, 249-261.
**Tables**

**Table 1.** The Traits of the Mature Fruit Locule Tissue of the Populations.

| Generation | Population   | Normal | All | Segregation | Theoretical Ratio | $\chi^2$ | Significance |
|------------|--------------|--------|-----|-------------|-------------------|---------|--------------|
| P1 (06-790)| 20           | 0      | 20  | -           | -                 | -       | -            |
| P2 (LA4069)| 20           | 20     | 0   | -           | -                 | -       | -            |
| F1         | 20           | 20     | 0   | -           | -                 | -       | -            |
| F2         | 191          | 150    | 41  | 3:1         | 1.176             | N.s.    |              |
| BC1P1      | 88           | 46     | 42  | 1:1         | 0.102             | N.s.    |              |
| BC1P2      | 40           | 40     | 0   | -           | -                 | -       | -            |
FIGURE LEGENDS

Fig. 1. Morphology and Micrograph of the Locule Tissues of Normal (WT) and All-Flesh Fruit Tomato. (A) The appearance of locule tissue at different developmental stages of WT tomato LA4069 and aff tomato 06-790. DAF: days after flowering; MG: mature green; BR: breaker ripe; RR: red ripe. (B) The cell structure of locule tissues of WT and aff tomato at their mature green stage. Scale bars: (A), 1 cm; (B), 50 μm.

Fig. 2. Map-based Cloning of the AFF Gene. (A) Δ(SNP index) from BSA-Seq. The x-axis is the physical position of tomato chromosomes; the y-axis is the value of the SNP-index. (B) Initial mapping of the AFF gene using 215 F2 plants derived from a cross between 06-790 and LA4069. (C) Genotypes and phenotypes of homozygous recombinant plants derived from 249 BC2S1 plants generated by continued backcrossing of 06-790 to H1706 (B51, B68, B228, B111, and B69 are normal lines; B64 and B166 are aff lines). (D) Annotated gene models in Tomato SL4.0 ITAG4.0 (H1706) in the mapping region. These local genes are indicated by rectangles with arrows. (E) Gene structure of AFF. The gray dashed-box represents the SV of 416-bp deletions in the cis-regulatory region of the AFF gene. (F) The PCR results of different tomato varieties or lines using the marker SV-12 designed by the 416-bp deletion. M: 100-bp DNA ladder.

Fig. 3. Characterization of CRISPR/Cas9-aff (aff-cr) Lines and Over-expression (AFF-over) Lines. (A) Schematic illustrating single-guide RNA targeting the AFF coding sequence (red triangle). (B) aff-cr mutants generated using CRISPR/Cas9. The red lines indicate the target sites of the guide RNAs. The nucleotides underlined in black bold font represent the protospacer-adjacent motif (PAM) sequences. aff-cr alleles identified by cloning and sequencing PCR products of the AFF-targeted region from two T0 plants under the Micro-tom background. (C) Representative fruit transection from CRISPR/Cas9-aff (aff-cr) lines compared with the WT and over-expression (AFF-over) lines at 25 days after flowering. Scale bars: 1 cm.
Fig. 4. The Sequence Conservation of the AFF Promoter in Solanaceae Species and the Tomato Germplasm. (A) Sequence conservation of promoter regions of AFF orthologous genes among five Solanaceae species. (B) Beanplot of π values for the three regions: the gene body, 3-kb upstream, and 3-kb downstream regions of all genes. The 3-kb upstream region of gene AFF shows strong conservation compared to other genes in the tomato germplasm. The yellow stars show the π values of the three regions of the AFF gene.

Fig. 5. The Expression of Gene AFF and the Phenotypes of Locule Tissues of WT and All-Flesh Tomato Fruits at Different Development Stages. (A) qRT-PCR of AFF transcripts in different locule tissues and stages from seven to 25 days after flowering (DAF). BA-130 and BA-150, aff lines derived from BC₂S₁ plants generated by the continued backcrossing of 06-790 to H1706. The 09-1225 and 06-790 are all-flesh cultivars. To obtain 06-790xH1706, F₁ progeny, the aff line 06-790 was crossed to WT H1706. H1706 and LA4069 are normal cultivars obtained from TGRC. BA-124 and BA-128 are normal lines derived from BC2S1 plants generated by continued backcrossing of 06-790 to H1706. Note: To normalize the expression data, the SlFRG27 (Solyc06g007510), SlFRG03 (Solyc02g063070), and ACTIN (Solyc11g005330) were used as the internal control (Cheng et al., 2017). The bars represent the standard deviation. (B) The longitudinal section of fruit locule tissue at different stages of the WT and aff NIL (PA-1) created by backcrossing of 06-790 to H1706 for six generations followed by two generations of selfing. MG: mature green; BR: breaker ripe; RR: red ripe. Scale bars: 1 cm.

Fig. 6. The Ratio of Firefly and Ranilla Luciferase Signals, as well as the Thousand-Seed Weight and the Seed Germination of aff NILs. (A) Relative luciferase activity (the ratio of luc to Rluc) of the two constructs. pGreenll luc, the blank vector with the 35S promoter; pGreenll luc-Del, the vector with the aff promoter; pGreenll luc-WT, the vector with the AFF promoter. Different letters above the bars indicate statistically significant differences. **: P < 0.01 (Student's t test). (B) The thousand-seed weight of four aff NILs. (C) The germination index of four aff NILs. (D) The seed germination percentage of four aff NILs. BA1-1 and BA2-1 are aff lines derived from BC₆S₂ plants generated by 06-790 continued backcrossing to H1706; BA4-1 and BA6-1 are normal lines derived from BC₆S₂ plants generated by 06-790 continued backcrossing to H1706. The bars represent the standard deviation.
Fig. 7. The Differential Gene Expression and Metabolite Contents between Fruits from the All-Flesh and WT Tomatoes. The significantly enriched GO terms (A) and KEGG pathways (B) of differentially expressed genes between locule tissues from the all-flesh fruit and WT tomatoes. (C) The principal component analysis of metabolites from the locule and placenta of WT and all-flesh fruit tomatoes. WT-L: locule tissue of WT tomatoes; WT-P: placenta tissue of WT tomatoes; AFF-L: locule tissue of aff tomatoes; AFF-P: placenta tissue of aff tomatoes.
Fig. 1. Morphology and Micrograph of the Locule Tissues of Normal (WT) and All-Flesh Fruit Tomato. (A) The appearance of locule tissue at different developmental stages of WT tomato LA4069 and aff tomato 06-790. DAF: days after flowering; MG: mature green; BR: breaker ripe; RR: red ripe. (B) The cell structure of locule tissues of WT and aff tomato at their mature green stage. Scale bars: (A), 1 cm; (B), 50 μm.
**Fig. 2. Map-based Cloning of the AFF Gene.** (A) Δ(SNP index) from BSA-Seq. The x-axis is the physical position of tomato chromosomes; the y-axis is the value of the SNP-index. (B) Initial mapping of the AFF gene using 215 F2 plants derived from a cross between 06-790 and LA4069. (C) Genotypes and phenotypes of homozygous recombinant plants derived from 249 BC2S1 plants generated by continued backcrossing of 06-790 to H1706 (B51, B68, B228, B111, and B69 are normal lines; B64 and B166 are aff lines). (D) Annotated gene models in Tomato SL4.0 ITAG4.0 (H1706) in the mapping region. These local genes are indicated by rectangles with arrows. (E) Gene structure of AFF. The gray dashed-box represents the SV of 416-bp deletions in the cis-regulatory region of the AFF gene. (F) The PCR results of different tomato varieties or lines using the marker SV-12 designed by the 416-bp deletion. M: 100-bp DNA DNA ladder.
Fig. 3. Characterization of CRISPR/Cas9-\textit{aff} (\textit{aff}-cr) Lines and Over-expression (\textit{AFF-over}) Lines. (A) Schematic illustrating single-guide RNA targeting the \textit{AFF} coding sequence (red triangle). (B) \textit{aff}-cr mutants generated using CRISPR/Cas9. The red lines indicate the target sites of the guide RNAs. The nucleotides underlined in black bold font represent the protospacer-adjacent motif (PAM) sequences. \textit{aff}-cr alleles identified by cloning and sequencing PCR products of the \textit{AFF}-targeted region from two T\textsubscript{0} plants under the Micro-tom background. (C) Representative fruit transsection from CRISPR/Cas9-\textit{aff} (\textit{aff}-cr) lines compared with the WT and over-expression (\textit{AFF-over}) lines at 25 days after flowering. Scale bars: 1 cm.
Fig. 4. The Sequence Conservation of the \textit{AFF} Promoter in \textit{Solanaceae} Species and the Tomato Germplasm. (A) Sequence conservation of promoter regions of \textit{AFF} orthologous genes among five \textit{Solanaceae} species. (B) Beanplot of $\pi$ values for the three regions: the gene body, 3-kb upstream, and 3-kb downstream regions of all genes. The 3-kb upstream region of gene \textit{AFF} shows strong conservation compared to other genes in the tomato germplasm. The yellow stars show the $\pi$ values of the three regions of the \textit{AFF} gene.
Fig. 5. The Expression of Gene AFF and the Phenotypes of Locule Tissues of WT and All-Flesh Tomato Fruits at Different Development Stages. (A) qRT-PCR of AFF transcripts in different locule tissues and stages from seven to 25 days after flowering (DAF). BA-130 and BA-150, aff lines derived from BC\textsubscript{2}S\textsubscript{1} plants generated by the continued backcrossing of 06-790 to H1706. The 09-1225 and 06-790 are all-flesh cultivars. To obtain 06-790×H1706, F\textsubscript{1} progeny, the aff line 06-790 was crossed to WT H1706. H1706 and LA4069 are normal cultivars obtained from TGRC. BA-124 and BA-128 are normal lines derived from BC2S1 plants generated by continued backcrossing of 06-790 to H1706. Note: To normalize the expression data, the SIFRG27 (Solyc06g007510), SIFRG03 (Solyc02g063070), and ACTIN (Solyc11g005330) were used as the internal control (Cheng et al., 2017). The bars represent the standard deviation. (B) The longitudinal section of fruit locule tissue at different stages of the WT and aff NIL (PA-1) created by backcrossing of 06-790 to H1706 for six generations followed by two generations of selfing. MG: mature green; BR: breaker ripe; RR: red ripe. Scale bars: 1 cm.
Fig. 6. The Ratio of Firefly and Ranilla Luciferase Signals, as well as the Thousand-Seed Weight and the Seed Germination of *aff* NILs. (A) Relative luciferase activity (the ratio of luc to RLuc) of the two constructs. pGreenII luc, the blank vector with the 35S promoter; pGreenII luc-Del, the vector with the *aff* promoter; pGreenII luc-WT, the vector with the *AFF* promoter. Different letters above the bars indicate statistically significant differences. **: *P* < 0.01 (Student's *t*-test). (B) The thousand-seed weight of four *aff* NILs. (C) The germination index of four *aff* NILs. (D) The seed germination percentage of four *aff* NILs. BA1-1 and BA2-1 are *aff* lines derived from BC6S2 plants generated by 06-790 continued backcrossing to H1706; BA4-1 and BA6-1 are normal lines derived from BC6S2 plants generated by 06-790 continued backcrossing to H1706. The bars represent the standard deviation.
Fig. 7. The Differential Gene Expression and Metabolite Contents between Fruits from the All-Flesh and WT Tomatoes. The significantly enriched GO terms (A) and KEGG pathways (B) of differentially expressed genes between locule tissues from the all-flesh fruit and WT tomatoes. (C) The principal component analysis of metabolites from the locule and placenta of WT and all-flesh fruit tomatoes. WT-L: locule tissue of WT tomatoes; WT-P: placenta tissue of WT tomatoes; AFF-L: locule tissue of aff tomatoes; AFF-P: placenta tissue of aff tomatoes.