

Identification of a lipase gene with a role in tomato fruit short-chain fatty acid-derived flavor volatiles by genome-wide association

Xiang Li^{1,2}, Denise Tieman¹, Zimeng Liu^{1,2}, Kunsong Chen^{2*} and Harry J. Klee^{1,2*} 

¹Horticultural Sciences, Genetics Institute, University of Florida, Gainesville, FL 32611, USA, and

²Laboratory of Fruit Quality Biology/Zhejiang Provincial Key Laboratory of Horticultural Plant Integrative Biology, Zhejiang University, Zijingang Campus, Hangzhou 310058, PR China

Received 23 April 2020; revised 14 July 2020; accepted 21 July 2020.

*For correspondence (e-mail hjklee@ufl.edu; akun@zju.edu.cn).

SUMMARY

Fatty acid-derived volatile organic compounds (FA-VOCs) make significant contributions to tomato (*Solanum lycopersicum*) fruit flavor and human preferences. Short-chain FA-VOCs (C5 and C6) are among the most abundant and important volatile compounds in tomato fruits. The precursors of these volatiles, linoleic acid (18:2) and linolenic acid (18:3), are derived from cleavage of glycerolipids. However, the initial step in synthesis of these FA-VOCs has not been established. A metabolite-based genome-wide association study combined with genetic mapping and functional analysis identified a gene encoding a novel class III lipase family member, *SI-LIP8*, that is associated with accumulation of short-chain FA-VOCs in tomato fruit. *In vitro* assays indicated that SI-LIP8 can cleave 18:2 and 18:3 acyl groups from glycerolipids. A CRISPR/Cas9 gene edited *SI-LIP8* mutant had much lower content of multiple fruit short-chain FA-VOCs, validating an important role for this enzyme in the pathway. *SI-LIP8* RNA abundance was correlated with FA-VOC content, consistent with transcriptional regulation of the first step in the pathway. Taken together, our work indicates that glycerolipid turnover by SI-LIP8 is an important early step in the synthesis of multiple short-chain FA-VOCs.

Keywords: *Solanum lycopersicum*, glycerolipids, fatty acids, lipase, volatile, FA-VOCs, Z-3-hexen-1-ol, hexyl alcohol.

INTRODUCTION

A diverse set of flavor-associated chemicals, including sugars, acids, and various volatile organic compounds (VOCs), contribute to human liking of tomato (*Solanum lycopersicum*) fruits (Klee, 2010; Tieman *et al.*, 2012; Klee and Tieman, 2013, 2018). There is considerable interest in the regulation of flavor-associated VOCs, in large part because modern, intensively bred crops such as tomato have lower concentrations of these chemicals and are much less flavorful than older heirloom varieties (Tieman *et al.*, 2017; Zhao *et al.*, 2019). Poor flavor is a major source of consumer dissatisfaction and there is a keen interest in genetic approaches to improve fruit volatile contents (Folta and Klee, 2016; Klee and Tieman, 2018). An important part of flavor improvement involves understanding the biosynthetic pathways for synthesis of the most important VOCs as well as understanding how the pathways are regulated. While many of the genes encoding biosynthetic enzymes have been identified (Speirs *et al.*, 1998; Howe *et al.*, 2000;

Chen *et al.*, 2004; Tieman *et al.*, 2006a; Goulet *et al.*, 2012, 2015), there are still major gaps in our knowledge, especially regarding regulation of pathway output.

Through a combination of biochemistry and consumer panel evaluations, the most important flavor-associated VOCs contributing to consumer liking of tomatoes have been identified (Tieman *et al.*, 2012, 2017). Genetic loci impacting the synthesis of these VOCs were identified and a genetic roadmap for how to improve flavor in modern cultivars was defined (Tieman *et al.*, 2017). This roadmap involves identifying superior alleles that are capable of increasing synthesis of favorable VOCs and reducing synthesis of undesirable VOCs. Importantly, the quantitative trait loci (QTLs) identified through genome-wide association studies (GWAS), provide targets for causative genes that significantly alter regulation of the metabolic pathways. Thus, identification of causative genes associated with those QTLs is economically valuable and provides fundamental knowledge about the regulation of metabolic flux.

VOC synthesis requires a series of catalytic processes that convert precursors such as lipids into volatile end products. Plant lipids are important membrane components that are essential both structurally and as precursors to a wide range of important metabolites (Andersson and Dörmann, 2009; Boudiere *et al.*, 2014; Wang *et al.*, 2017, 2019; Higashi *et al.*, 2018). Consequently, lipid turnover into fatty acids (FAs) must be highly regulated. The polyunsaturated FA linolenic acid (18:3) is a substrate for synthesis of oxylipins, including jasmonic acid (JA), a multifunctional growth regulator with critical roles in defense against a wide array of biotic stresses (Holopainen and Gershenzon, 2010; Erb *et al.*, 2012; Campos *et al.*, 2016; Guo *et al.*, 2018; Howe *et al.*, 2018). Linolenic (18:3) and linoleic (18:2) acids are also precursors for flavor VOCs (Chen *et al.*, 2004; Shen *et al.*, 2014). FA-derived VOCs (FA-VOCs) constitute the most abundant class of volatiles in tomato fruits and are ubiquitous in the plant kingdom (Schwab *et al.*, 2008). In the tomato fruit, multiple C5, C6, C7, C8, and C10 FA-VOCs are significantly correlated with flavor and human preferences (Tieman *et al.*, 2012, 2017). Many of these same volatiles have vital functions in inter- and intra-kingdom communication, signaling, and disease and insect resistance (Bouwmeester *et al.*, 2019; Hammerbacher *et al.*, 2019). Multiple downstream enzymes in the FA-VOC biosynthetic pathway, including lipoxygenase, hydroperoxide lyase, and alcohol dehydrogenase 2, have been identified (Speirs *et al.*, 1998; Howe *et al.*, 2000; Porta and Rocha-Sosa, 2002; Chen *et al.*, 2004; Shen *et al.*, 2014; Klee and Tieman, 2018). However, little is known about the initial step in the pathway, presumably catalyzed by acyl group hydrolases.

Lipases are acyl group hydrolases having critical roles in multiple biological processes, including seed germination, stress resistance, and senescence responses. For example, PLIP1 is a phosphatidylglycerol lipase that participates in the export of acyl groups from plastids for seed oil biosynthesis in developing *Arabidopsis* embryos (Wang *et al.*, 2017). DEFECTIVE IN ANTER DEHISCENCE 1 (DAD1) is an *Arabidopsis* lipase that catalyzes the initial step of JA biosynthesis (Ishiguro *et al.*, 2001). A triacylglycerol (TAG) lipase in *Arabidopsis* has an important role in plant growth and senescence while mobilizing FAs from complex TAGs (Padham *et al.*, 2007). In tomato, LeLID1, a TAG lipase, contributes to seed germination by mobilizing stored lipids (Matsui *et al.*, 2004). Recently, another TAG lipase, SILIP1, was shown to have a critical role in synthesis of longer-chain volatiles (C10), including Z-4-decenal and E,E-2,4-decadienal, which are significantly correlated with tomato flavor (Garbowicz *et al.*, 2018). The lipase(s) responsible for synthesis of shorter-chain C5 and C6 FA-VOCs have not been identified. In the case of tomato fruits, short-chain C5 and C6 FA-VOCs, including 1-penten-3-ol, 1-pentanol, Z-3-hexen-1-ol, and hexyl alcohol, are important volatiles that contribute to human liking and flavor intensity (Tieman *et al.*, 2012, 2017). These same

volatiles are also important for defense, insect attraction, and pollination in many plant species (Croft *et al.*, 1993; Bate and Rothstein, 1998; Dicke and Baldwin, 2010; Scala *et al.*, 2013; Rowen and Kaplan, 2016; Beck *et al.*, 2018). Given the abundance and importance of FA-VOCs in the plant kingdom, identification of the first step in the biosynthetic pathway of these volatiles is an important step forward.

Tomato is one of the world's most widely grown and highest valued crops, as well as an important source of micronutrients, vitamins, and antioxidants in the human diet. It also serves as a model system for genetic control of fruit development and chemistry. Metabolic QTL (mQTL) mapping and GWAS are powerful tools for the identification of genetic determinants of plant metabolism (Ye *et al.*, 2017; Fang and Luo, 2019). Metabolic GWAS (mGWAS) have also been used with various crops to gain deeper insights into the genetic bases of metabolic diversity of both primary and secondary metabolites (Sauvage *et al.*, 2014; Matsuda *et al.*, 2015; Zhou *et al.*, 2016; Angelovici *et al.*, 2017; Peng *et al.*, 2017; Zhu *et al.*, 2018). A GWAS performed by Tieman *et al.* (2017) on a large population of tomato accessions identified a single nucleotide polymorphism (SNP) on chromosome 9 that is significantly associated with short-chain FA-VOC content, including Z-3-hexen-1-ol and hexyl alcohol. In addition, using lipidomic profiling of *Solanum pennellii* introgression lines (ILs), Garbowicz *et al.* (2018) used mQTL analysis to identify a region on chromosome 9 affecting glycerolipid metabolism. This region was previously observed to contain a QTL affecting Z-3-hexen-1-ol and hexyl alcohol content (Tieman *et al.*, 2006b). Taken together, the results support the presence of a genetic locus on chromosome 9 affecting the metabolism of FAs to short-chain FA-VOCs.

Based on the GWAS results (Tieman *et al.*, 2017) and the previous IL population studies, we identified a candidate gene, encoding a class III lipase (*SI-LIP8*, *Solyc09g091050*), within the overlapping region of two *S. pennellii* introgressions containing the chromosome 9 mQTL, IL9-3-1 and IL9-3-2. We generated *SI-LIP8* mutants using CRISPR/Cas9 to validate an important role in the synthesis of three C6 volatiles (Z-3-hexen-1-ol, E-2-hexen-1-ol, and hexyl alcohol) and two C5 volatiles (1-penten-3-ol and 1-pentanol). We also performed enzymatic assays *in vitro* to confirm the function of *SI-LIP8* in cleaving glycerolipids to release the precursors of the FA-VOCs, linoleic and linolenic acids. Taken together, our study identified a lipase gene that is involved in synthesis of important flavor-associated VOCs.

RESULTS

GWAS analysis identified a candidate lipase gene on tomato chromosome 9

To gain insight into the underlying causes of flavor loss in modern commercial tomatoes, Tieman *et al.* (2017)

performed a GWAS on 398 *S. lycopersicum*, *S. lycopersicum* var. *cerasiforme*, and *Solanum pimpinellifolium* accessions. A wide range of variations observed in volatile chemical content provided important clues to define genetic loci affecting VOC content. One locus on chromosome 9 associated with SNP ch09_70173370 ($P < 6.15 \times 10^{-9}$) was significantly associated with Z-3-hexen-1-ol and hexyl alcohol content (Figure 1(a)). The average contents of these two VOCs were significantly higher in varieties with the Heinz 1706 reference (R) allele compared to those varieties with the alternate (A) allele (Figures 1(b) and S1). Candidate genes within 0.5 Mb (less than the average linkage disequilibrium [LD] of tomato) to either side of the SNP are listed in Table S2.

Based on the gene annotation and publicly available RNA-Seq data, a previously uncharacterized gene, *Solyc09g091050* (*SI-LIP8*), predicted to encode a class III lipase, was identified as a candidate for the causative gene (Figure 1(a)). In a prior multi-omics study of the 392 accessions used in the GWAS, RNA-Seq analysis was performed on ripening fruit samples (Zhu *et al.*, 2018). Analysis of *SI-LIP8* transcript in lines containing either the R or the A allele was consistent with the pattern of Z-3-hexen-1-ol and hexyl alcohol contents, with the average *SI-LIP8* transcript level being significantly higher in varieties with the R allele than those with the A allele (Figure 1(b)). Quantitative PCR analysis was also performed to assess *SI-LIP8* gene expression in three randomly selected R and A varieties. The result

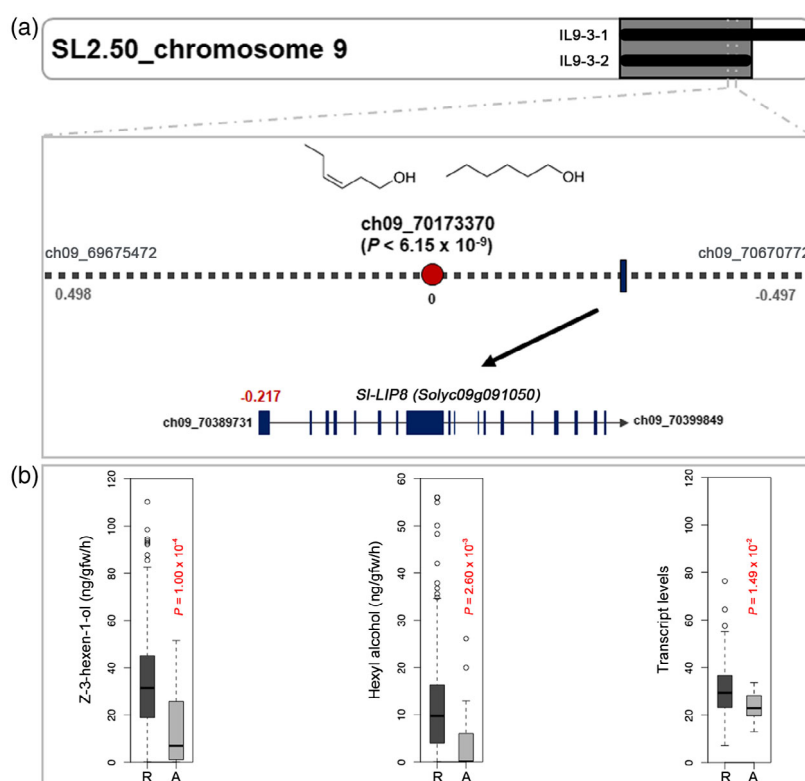
showed that the *SI-LIP8* transcript is significantly more abundant in the three R varieties with higher Z-3-hexen-1-ol and hexyl alcohol content (Figure S2(a)). Taken together, *SI-LIP8* was predicted to be the causal gene underlying the SNP ch09_70173370 identified via GWAS.

Garbowicz *et al.* (2018) conducted a lipidomic profiling of fruit pericarp and leaf tissue of a set of *S. pennellii* ILs that identified an mQTL in the same region of chromosome 9. Within this region, the contents of multiple TAGs, monogalactosyldiacylglycerols (MGDGs), and phospholipids were altered by (a) gene(s) in the overlapping region of IL9-3-1 and IL9-3-2 (Figure 1(a)). They noted a predicted lipase gene within the 19.3-cM region defined by the IL overlap (*Solyc09g091050*, *SI-LIP8*). The transcript abundance of *SI-LIP8* in *S. lycopersicum* (M82) was approximately 10-fold higher than that in *S. pennellii* and approximately 3.5- and 6.2-fold higher than that of IL9-3-1 and IL9-3-2, respectively. The results of genetic mapping in the *S. pennellii* ILs are consistent with the results of GWAS analysis. Based on all these results, we inferred that *SI-LIP8*, located within the region common to IL9-3-1 and IL9-3-2, is the causal gene with an important role in the first step of lipid-derived VOC synthesis.

Molecular analysis of *SI-LIP8*

Although the polyunsaturated long-chain FAs linoleic acid (18:2) and linolenic acid (18:3) are believed to be precursors of the C5 and C6 short-chain FA-VOCs, little is known

Figure 1. Genome-wide association analysis of Z-3-hexen-1-ol and hexyl alcohol. (a) Map of *S. lycopersicum* chromosome 9 and identification of the causative gene (*SI-LIP8*). The gray bar represents the overlapping region of IL9-3-1 and IL9-3-2. The IL population was derived from a cross between *S. lycopersicum* var. M82 and *S. pennellii* LA0716. The red dot and blue box represent the lead SNP (ch09_70173370) and *SI-LIP8* (*Solyc09g091050*), respectively. The horizontal dashed line indicates the region ~ 0.5 Mb to either side of the lead SNP. The gene model of *SI-LIP8* with the chromosome location is shown. (b) Box plot indicating distribution of volatile (Z-3-hexen-1-ol and hexyl alcohol) contents in ripening fruits from plants containing the reference (R, $n = 332$) or alternate (A [A>G variants at position ch09_70173370], $n = 19$), as well as transcript levels of *SI-LIP8* in R ($n = 292$) and A ($n = 13$) varieties (Table S1). The horizontal lines within the boxes represent median values. The box height represents the range of the 25th to the 75th percentile of the data. The whiskers and outer dots indicate interquartile ranges and outliers, respectively (* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$).



about the enzymes responsible for generating the free FAs (FFAs) or how that release is regulated. SI-LIP8 is annotated as a member of the class III lipase family (SGN, <https://solgenomics.net/>). This family constitutes a large number of lipolytic enzymes that catalyze the formation of glycerol and FFAs from TAGs. Typically, lipases belong to a group of serine esterases having a signature motif (an Asp-His-Ser triad), but some possess only a Ser-Asp dyad (Kelly and Feussner, 2016; Wang *et al.*, 2017, 2018). Amino acid sequence alignment of SI-LIP8 with characterized lipases using the NCBI conserved domain database identified a conserved dyad (Ser-682 and Asp-742) and a conserved GxSxG domain with the serine active site residue (Figure S3). A phylogenetic analysis that included SI-LIP8, the top 11 tomato proteins similar to SI-LIP8, and well-characterized lipases from other plant species was performed (Figure 2a; sequence in Data S1). The result showed that SI-LIP8 clusters with Arabidopsis Heat Inducible Lipase 1 (AtHIL1, *At4g13550*), although the two

sequences have only 38.96% similarity (Figure S4). *AtHIL1* encodes a chloroplast lipase with relatively broad substrate specificity for leaf glycolipids and phospholipids. *In vitro* assays indicated that *AtHIL1* hydrolyzes MGDGs to release 18:3-FFA (α -linolenic acid) (Higashi *et al.*, 2018).

SI-LIP8 encodes a polypeptide with 863 amino acids containing a lipase III domain and a C2 domain that are also present in *AtHIL1* (Figure 2(b)). *AtHIL1*, as well as another TAG lipase (*At2g31690*), has chloroplast transit peptides (cTP), and the truncated proteins have higher catalytic activity than the full-length proteins (Padham *et al.*, 2007; Higashi *et al.*, 2018). Analysis of the SI-LIP8 protein sequence, performed using the TargetP and ChloroP programs, predicted that it also possesses a 69-amino acid (~8 kDa) cTP (Figure 2(b)) resulting in a predicted mature protein of approximately 92 kDa. Gene expression data from the Expression Viewer in SGN indicate that *SI-LIP8* is expressed in multiple tissues (buds, flowers, leaves, and roots) as well as different stages of fruit development

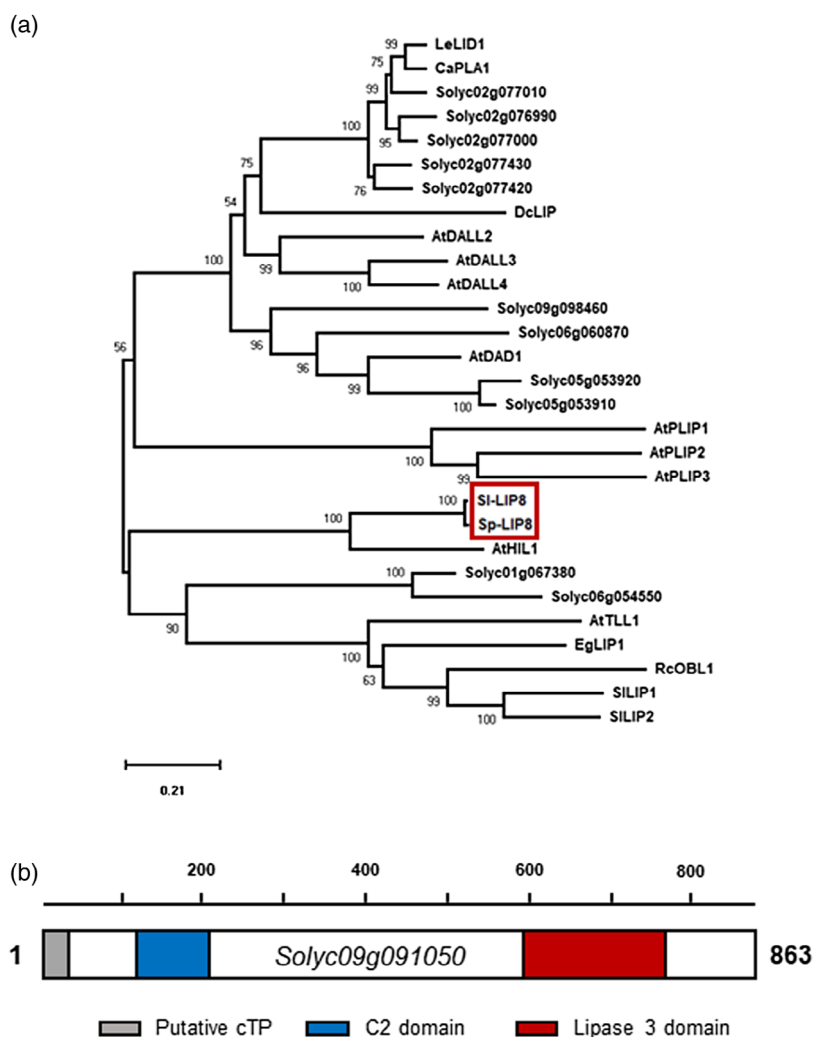


Figure 2. Phylogeny and structure of SI-LIP8. (a) Phylogenetic analysis among SI-LIP8, 11 *S. lycopersicum* proteins similar to Sp-LIP8, and other characterized plant lipases. The alignment of deduced full-length amino acid sequences (Data S1) was used to construct an unrooted neighbor-joining tree (Mega X). The percentage of similarity in which the associated taxa are clustered together is annotated close to the branches. The bootstrap values in which the associated taxa are clustered together is annotated close to the branches. GenBank accession or NCBI reference numbers are as follows: Sp-LIP8 (Sopen09g034030), SI-LIP8 (Solyc09g091050), SILIP1 (Solyc12g055730), SILIP2 (Solyc03g123750), *AtHIL1* (*At4g13550*), *AtPLIP1* (*At3g61680*), *AtPLIP2* (*At1g02660*), *AtPLIP3* (*At3g62590*), *AtDAD1* (*At2g44810*), *AtDLL2* (*At1g51440*), *AtDLL3* (*At2g30550*), *AtDLL4* (*At1g06800*), *AtTLL1* (*At1g45201*), *LeLID* (XP_00432963), *CaPLA1* (EF595843), *DcLIP* (AAD01804), *EgLIP1* (JX556215), and *RcOBL1* (JQ945176). (b) Schematic illustration of the *SI-LIP8* structure. Gray, blue, and red boxes represent the putative cTP sequence, C2 domain, and lipase_3 domain.

(Tomato Genome, 2012). The transcript is most abundant in fruits and its expression increases during ripening (Figure S2(b)). This pattern of fruit expression is consistent with that of the short-chain FA-VOCs (Tiemann *et al.*, 2006b).

Deletion mutants generated via CRISPR/Cas9

Based on candidate single-guide RNAs (sgRNAs) provided by CRISPR-P (Lei *et al.*, 2014), we selected two adjacent sgRNAs separated by 38 bp in exon 1 of *SI-LIP8*. The sgRNA 1 and 2 double cassette was inserted into the pCambia2300 vector (Figure 3(a)) and introduced into the Fla. 8059 cultivar by *Agrobacterium tumefaciens*-mediated transformation. Two independent lines containing the *NPTII* selectable marker gene were grown and self-pollinated to generate T₁ progeny. DNA sequence analysis of the target region indicated that both lines contained an identical (5-bp) deletion (Figures 3(b) and S5a). Digestion of the wild-type (WT) *SI-LIP8* PCR fragment (760 bp) with *DdeI* results in four products of 127, 314, 142, and 177 bp. Cas9 editing of the target site eliminated one of the *DdeI* sites that is adjacent to the protospacer-adjacent motif, leading to only three fragments of 127, 451, and 177 bp (Figure S5(b)). The deletion resulted in a frameshift from the first amino acid of the mature SI-LIP8 protein and a premature stop codon at the sixth amino acid, resulting in a truncated protein (Figure 3(c)).

To determine whether *SI-LIP8* is responsible for the variation in FA-VOCs levels associated with the QTL, the levels of 18 volatiles from fully ripe fruits of the homozygous mutants (T₁, SI-LIP8CR) and WT Fla. 8059 were measured.

The contents of nine FA-VOCs were significantly reduced in SI-LIP8CR mutants. Significantly reduced VOCs included three C5 (1-pentanol, Z-2-penten-1-ol, and 1-penten-3-ol) and three C6 (Z-3-hexen-1-ol, E-2-hexen-1-ol, and hexyl alcohol) VOCs (Figure S6(a); Table S3). In addition, contents of two C8 (hexyl acetate and E-2-octenal) and one C10 (Z-4-decenal) VOC were significantly reduced (Table S3).

Homozygous *SI-LIP8CR* T₃ progeny from subsequent generations were grown in the field and greenhouse. Although some statistically significant differences were observed in hexyl acetate and E-2-octenal, consistent reductions in the levels of two C5 (1-pentanol and 1-penten-3-ol) and three C6 (Z-3-hexen-1-ol, E-2-hexen-1-ol, and hexyl alcohol) VOCs were observed in the *SI-LIP8CR* mutants across every season and location (Figure S6(b)). No consistent differences were observed in longer-chain FA-VOCs (Table S3). The consistent result was also observed in backcrossed F₂ generation (BC1-F₂) fruits (Figure 4). Overall, these results indicate that *SI-LIP8* has an important role in synthesis of multiple short-chain FA-VOCs in tomato fruits, consistent with *SI-LIP8* being responsible for the GWAS phenotype.

SI-LIP8 cleaves 18:2 and 18:3 acyl groups from glycerolipids *in vitro*

In an effort to determine SI-LIP8 activity, cDNAs corresponding to the full-length (SI-LIP8_F) and mature (SI-LIP8_M, without the putative cTP sequence) proteins were overexpressed as recombinant His-tagged proteins in *Escherichia coli*. Proteins of the predicted molecular

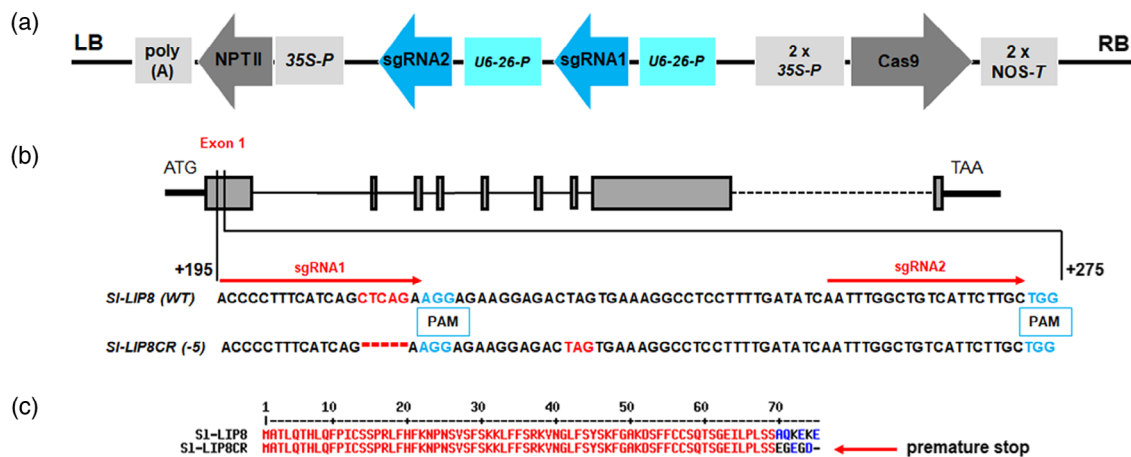


Figure 3. Construction and characterization of *SI-LIP8CR*-edited lines. (a) Schematic of the binary vector pCambia2300_CR3-EF with two single-guide RNAs (sgRNAs) and Cas9 endonuclease used for *Agrobacterium*-mediated transformation of *S. lycopersicum* var. Fla. 8059. The *U6-26* promoter from *Arabidopsis thaliana* drives expression of sgRNA1 and 2, and the enhanced 35S promoter from Cauliflower Mosaic Virus drives expression of the *NPT II* and *Cas9* genes together with sgRNAs to induce mutations in the target region of the *SI-LIP8* gene. Nos-T, Nos terminator; LB, left border; RB, right border. (b) Schematic illustration of the two sgRNA target sites (horizontal red arrow) in *SI-LIP8*. *SI-LIP8CR* carries a single 5-bp deletion (red dash line), and blue font indicates protospacer-adjacent motif (PAM) sequences. (c) Alignment of the first 76 amino acid sequences of wild-type (SI-LIP8) and mutant (SI-LIP8CR) SI-LIP8. A premature stop codon is indicated with a horizontal red arrow. The red font indicates the putative chloroplast transit peptide (cTP).

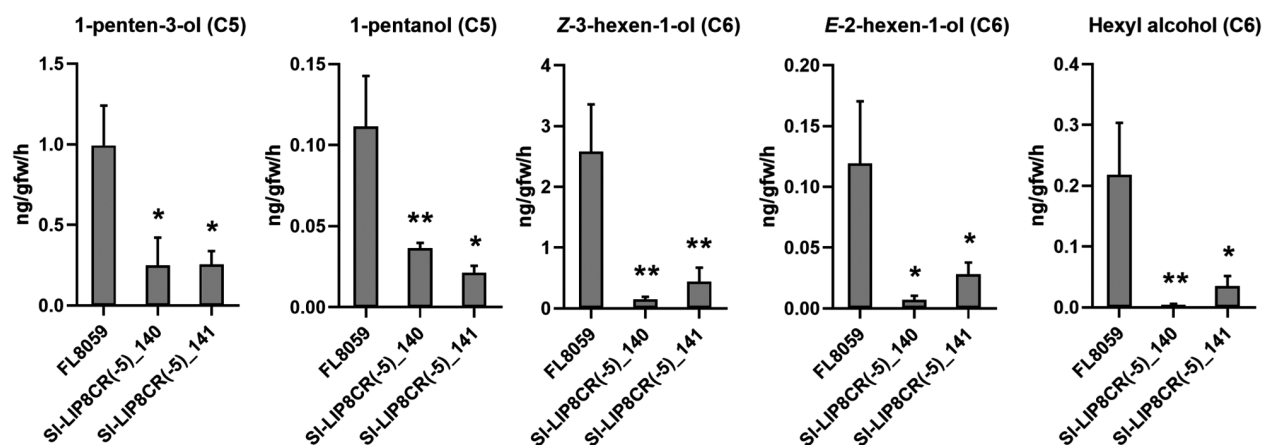


Figure 4. The effects of SI-LIP8 mutation on FA-VOCs in BC1-F₂ and T₃ fruit. Reduced emissions of five FA-VOCs in the ripe fruits of CRISPR/Cas9-engineered transgenic lines ($\pm$ SE, $n = 6$). C# in the parentheses indicates the number of carbons in each FA-VOC. FL8059, wild-type Fla. 8059; BC, backcross; BC1-F₂, SI-LIP8CR(-5)₁₄₀. Asterisks indicate significant differences (* $P < 0.05$; ** $P < 0.01$).

weights of SI-LIP8_F and SI-LIP8_M, 100 kDa and 92 kDa, respectively, were observed (Figures S7(a) and S10(b)).

Enzymatic assays were first conducted using two commercial TAGs (54:9-TAG [glyceryl trilinoleate] and 54:6-TAG [glyceryl trilinolenate]) as substrates. Non-esterified FA (NEFA) assay results showed FFA contents significantly increased when the purified recombinant SI-LIP8_F protein was incubated with 54:6-TAG or 54:9-TAG as substrates (Figure 7S(b)). In addition, an FA methyl esterification (FAME) assay was performed and FA methyl esters were subsequently analyzed by gas chromatography-mass spectrometry (GC-MS). As expected, the volatile esters were characterized as methyl linoleate (18:2) (Figure 5(a)) and methyl linolenate (18:3) (Figure 5(b)), and incubation of recombinant SI-LIP8_F with a mixture of the aforementioned two TAGs resulted in a significant increase in the levels of these two volatile esters. FA methyl esters were not detected in the no substrate control (Figure 5(c,d)).

Another quantitative FFA assay was performed using SI-LIP8_M, and the result confirmed SI-LIP8 has lipase activity in cleaving the acyl group from the aforementioned two TAGs. In addition, two commercial phosphatidylcholines (PCs) (36:4-PC [1,2-dilinoleoyl-sn-glycero-3-phosphatidylcholine] and 36:6-PC [1,2-dilinenoyl-sn-glycero-3-phosphatidylcholine]) were also used for lipase activity measurement (Figure 6). Kinetic assay showed that the specificity constants (K_{cat}/K_m) of SI-LIP8_M for 36:4-PC and 36:6-PC were $0.086 \mu\text{M}^{-1}\text{s}^{-1}$ and $1.506 \mu\text{M}^{-1}\text{s}^{-1}$, respectively (Table 1). The results indicated that SI-LIP8_M has greater activity on PCs than TAGs and greater activity on 18:3 than 18:2 acyl groups (Figure 7(a)). These results are similar to those reported for AtHIL1, an SI-LIP8 homolog, which has a preference for 18:3 FA (Higashi *et al.*, 2018). In addition, the SI-LIP8_M recombinant protein exhibited much higher lipase activity than SI-LIP8_F using 36:4-PC as

a substrate (Figure 7(b)). These results indicate that SI-LIP8 encodes an acyl group hydrolase that is able to cleave various glycerolipids with a probable preference to release 18:3-FFA (Figure 7(c)) and removal of the cTP from SI-LIP8 is necessary for superior lipase activity.

Transcriptional control of SI-LIP8 underlies short-chain FA-VOC variation

All of our *in vivo* and *in vitro* assays validated the hypothesis that SI-LIP8 functions in 18:2 and 18:3 acyl group hydrolysis, providing substrates for the later steps of short-chain FA-VOC synthesis catalyzed by lipoxygenases and hydroperoxide lyase. Based on the genome sequencing data of 398 tomato accessions (Tieman *et al.*, 2017), the ch09_70173370 SNP is functionally linked to the variation in FA-VOCs observed within the population. Since this SNP is located 0.217 Mb upstream of SI-LIP8 (Figure 1(a)), additional nonsynonymous SNPs present within the coding region of SI-LIP8 were investigated. In particular, two SNPs were observed to be associated with the FA-VOC contents: the ch09_70393277 G>A variant, resulting in an Ala-213 to Thr-213 change, and the ch09_70374515 G>T variant, resulting in an Asp-428 to Tyr-428 change (sequence data in Table S4). When the volatile contents of reference- versus alternate allele-containing accessions were compared, there were significant differences observed for the contents of multiple short-chain FA-VOCs, with the alternate allele-containing accessions having lower contents of these VOCs (Figures S8 and S9; Table S5(a,b)).

To determine whether the amino acid changes are responsible for the variation in FA-VOC contents, we performed kinetic assays comparing lipase activity between the reference (SI-LIP8_M^{3277_G/4515_G}) and alternate (SI-LIP8_M^{3277_A} or SI-LIP8_M^{4515_T}) forms of the SI-LIP8 enzymes (Figure S10(a)). Recombinant enzymes were

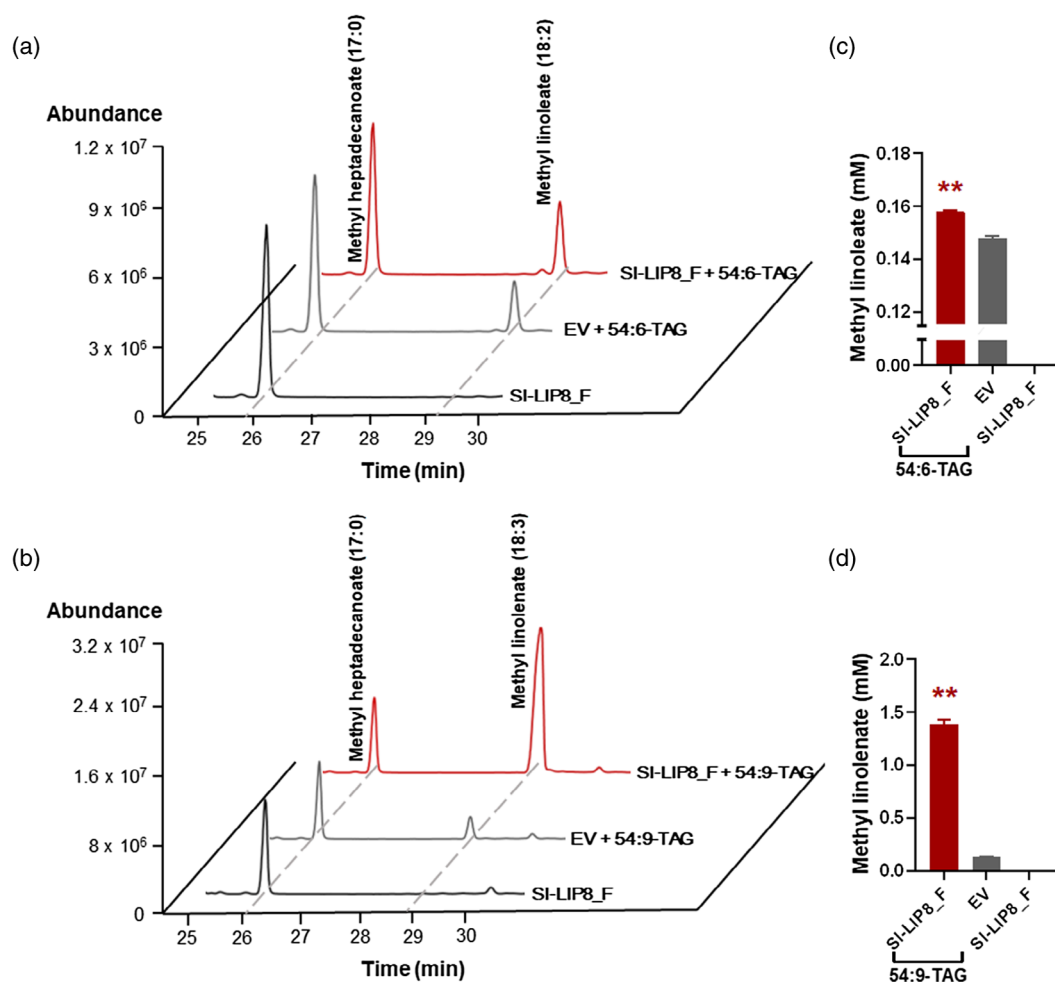


Figure 5. Fatty acid methyl esterification (FAME) assay. (a,b) Identification of enzymatic products after full-length recombinant protein (SI-LIP8_F, ~2 μ g) was incubated with 54:6-TAG or 54:9-TAG (~0.3 mg). After esterification, the peaks of putative methyl linoleate (18:2) and methyl linolenate (18:3) were analyzed by GC-MS. (c,d) Determination of volatiles that were detected using GC-MS ($\pm$ SE, $n = 3$). Heptadecanoic acid (17:0) was used as the internal standard. Red line and bar, SI-LIP8_F with 54:6-TAG or 54:9-TAG. Gray line and bar, EV with 54:6-TAG or 54:9-TAG (control). Black line and bar, SI-LIP8_F without substrates (control). EV, empty vector (pET28a). Asterisks indicate significant differences (** $P < 0.01$).

expressed in *E. coli* and purified (Figure S10(b)). The affinity constant (K_m) and specificity constant (K_{cat}/K_m) were determined for each enzyme using 36:4-PC as a substrate. There were no significant differences between the reference and SI-LIP8_M^{4515_T} enzymes; SI-LIP8_M^{3277_A} was somewhat less efficient (Figures S8 and S9; Table S6). Thus, the amino acid differences are unlikely to explain the observed differences in VOC contents.

Table 1 Enzymatic activity of SI-LIP8_M using 36:4-PC and 36:6-PC as substrates

Substrate	K_m (μ M)	V_{max} (μ M s ⁻¹)	K_{cat}/K_m (μ M ⁻¹ s ⁻¹)
36:4-PC	2.22 $\pm$ 0.74	0.06 $\pm$ 0.01	0.09
36:6-PC	3.43 $\pm$ 1.54	1.51 $\pm$ 0.36	1.51

We then examined the contents of SI-LIP8 transcripts in accessions containing the aforementioned two alleles. The average transcript levels of SI-LIP8 in accessions with the alternate 3277_A or 4515_T allele were significantly lower than those in accessions with the reference 3277_G or 4515_G allele (Zhu *et al.*, 2018) (Figure S11). These results indicate that transcript abundance is likely to be the underlying cause of the differences in FA-VOC emissions, although we cannot exclude a contribution from the SI-LIP8_M^{3277_A} amino acid change. The overall results are consistent with the differences in LIP8 transcript abundances observed between *S. lycopersicum* cv. M82 and *S. pennellii* LA0716. Garbowicz *et al.* (2018) reported that in ripening fruits, the *S. pennellii* LIP8 transcript abundance is approximately 10% of that in *S. lycopersicum*, consistent with transcriptional regulation of FA-VOC contents. When we compared the sequences of the genes from these two

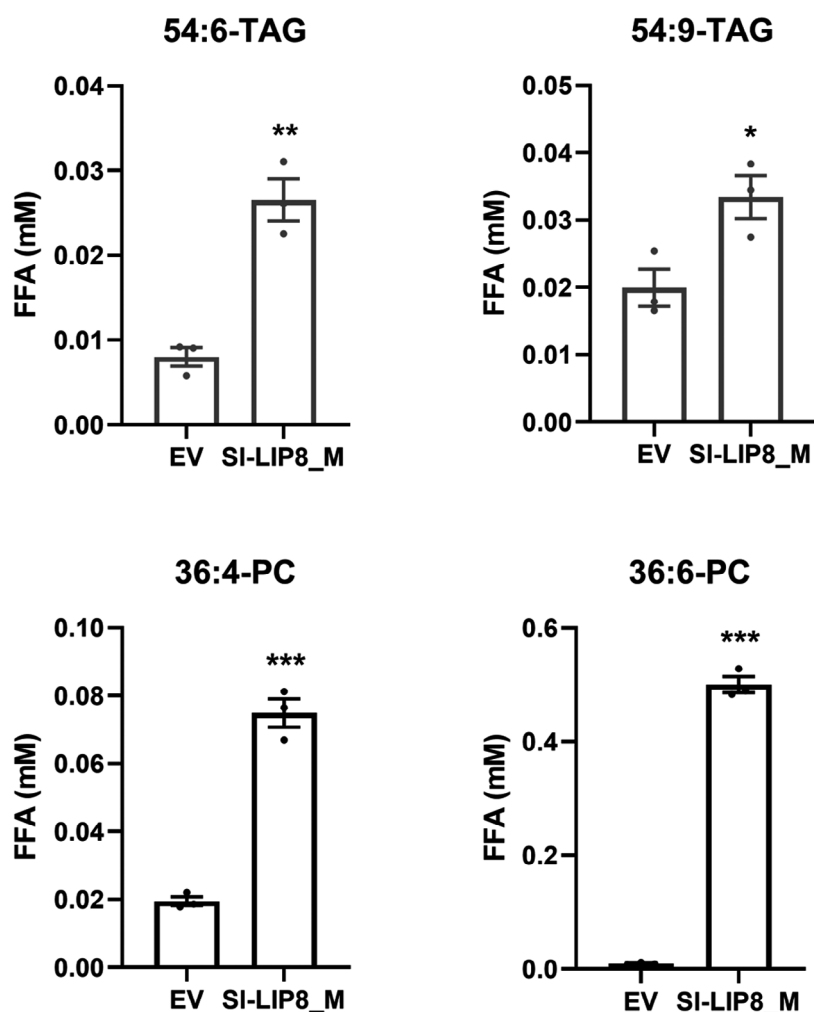


Figure 6. *In vitro* lipase activity of SI-LIP8. Quantitative analysis of FFAs generated from mature recombinant protein (SI-LIP8_M, ~0.5 μ g) mixed with 54:6-TAG, 54:9-TAG, 36:4-PC, or 36:6-PC (~0.1 mg). Each data point represents biological replicates ($\pm$ SE, $n = 3$). EV, empty vector (pET28a); FFA, free fatty acid. Asterisks indicate significant differences (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$).

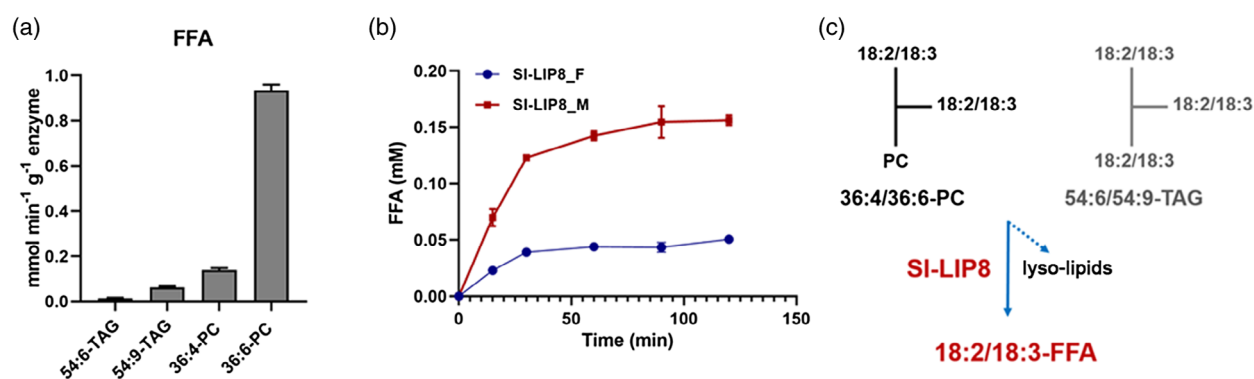


Figure 7. Comparative analysis of lipase activity. (a) SI-LIP8_M activity on TAGs and PCs ($\pm$ SE, $n = 3$). In each assay, ~0.1 mg of each glycerolipid and ~0.5 μ g SI-LIP8_M were used. (b) Comparative analysis of lipase activity between recombinant SI-LIP8_F and SI-LIP8_M proteins (~0.5 μ g) using 36:4-PC (~0.1 mg) as the substrate ($\pm$ SE, $n = 3$). (c) Schematic illustration of the SI-LIP8 lipase reaction. PC, phosphatidylcholine; TAG, triacylglycerol.

accessions, we observed a ~660-bp insertion in the *S. lycopersicum* transcriptional promoter region (Figure 8). This result is similar to our prior observation of a large

insertion into the promoter of the *S. lycopersicum* *CXE1* relative to its *S. pennellii* ortholog, which also resulted in much higher expression of *SICXE1* (Goulet *et al.*, 2012).

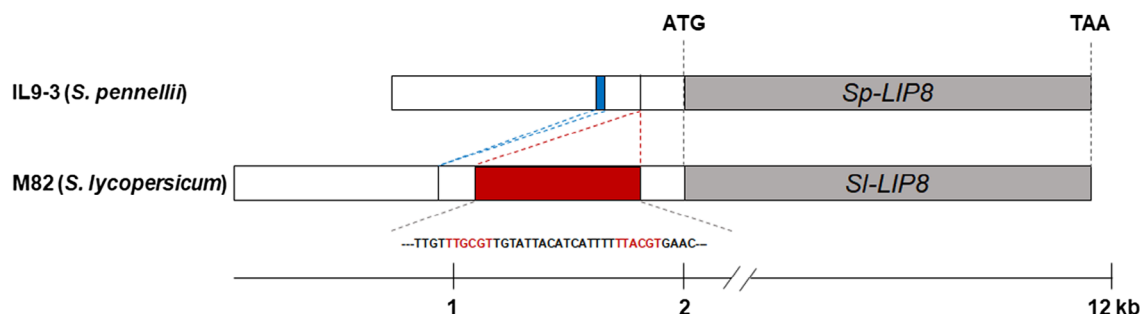


Figure 8. Comparative analysis of *Sp-LIP8* and *SI-LIP8* transcriptional promoter sequences. A ~660-bp insertion is present in the promoter region in M82 (red box) containing a tandem repeat of a NAC-binding *cis*-element (red font). M82 lacks a 33-bp segment (blue box) present in IL9-3.

DISCUSSION

Deterioration in fruit flavor quality in recent decades has become a major cause of consumer complaints. In the case of tomatoes, modern commercial varieties are generally substantially less flavorful than heirloom varieties, in large part due to cumulative loss of superior alleles affecting biosynthesis of important flavor-associated chemicals over many breeding cycles. Understanding the molecular basis of variation within the species in flavor-associated VOCs is an essential part of restoring and maintaining the maximum potential flavor in commercial cultivars. We previously analyzed GWAS data from 398 diverse *S. lycopersicum* and *S. pimpinellifolium* accessions to identify loci that quantitatively influence flavor-associated VOCs (Tiemann *et al.*, 2017). With knowledge of how each VOC positively or negatively contributes to consumer preferences, it is possible to identify the most desirable (superior) alleles at each locus. We can further determine which superior alleles are present in any cultivar and design breeding strategies to introduce and/or maintain the superior alleles. Since each QTL is defined by one or more SNPs that are not necessarily in the causative genes, it is most desirable to identify the linked underlying causative gene for each QTL.

Multiple short-chain FA-VOCs have a significant impact on consumer preferences (Tiemann *et al.*, 2017). Lipids are presumed to be the initial precursors for these FA-VOCs. An essential first step in the pathway is release of FFAs catalyzed by various acyl hydrolases. C5 and C6 VOCs are particularly abundant in ripening tomato fruits (Buttery *et al.*, 1989). These VOCs are derived from 18:2 (linoleic acid) and 18:3 (linolenic acid) FAs by subsequent action of lipoxygenases and hydroperoxide lyase (Chen *et al.*, 2004; Shen *et al.*, 2014). Thus, lipases that release linoleic and linolenic acids are of particular interest for modifying flavor volatile content.

One QTL significantly associated with C5 and C6 FA-VOC content was identified near the bottom of chromosome 9 (Tiemann *et al.*, 2017). The SNP defining this QTL is close to

a gene annotated as a lipase, *SI-LIP8* (*Solyc09g091050*). A lipid mQTL analysis of an IL population derived from a cross between *S. lycopersicum* cv. M82 and *S. pennellii* also identified a locus in the same area of chromosome 9 linked to altered C18 glycerolipid content (Garbowicz *et al.*, 2018). That region includes the *SI-LIP8* gene, and Garbowicz *et al.* (2018) suggested that it could be the underlying causative gene. Very recently, Kuhalskaya *et al.* (2020) have, through network analysis, shown that the level of the *SL-LIP8* transcript is inversely correlated with fruit digalactosyldiacylglycerol content. To investigate the GWAS and mQTL results, we conducted experiments to validate the function of the lipase gene *SI-LIP8* in regulating the contents of short-chain FA-VOCs.

In Arabidopsis, although approximately 300 genes have been annotated as putative lipases (Li-Beisson *et al.*, 2013; Kelly and Feussner, 2016), only a fraction of them have been biochemically validated. These lipases function in plant defense, development, membrane maintenance, and signaling. Less is known about the functions of lipases in lipid metabolism for other plant species, including tomato, particularly as related to VOC synthesis. A gene encoding a tomato lipase (*LIP1*) on chromosome 12 has an important role in TAG and diacylglycerol metabolism, but it has low protein sequence identity (27.32%) to *SI-LIP8*. In addition, *LIP1* is mainly involved in synthesis of longer-chain FA-VOCs such as *Z*-4-decenal (C10) and *E,E*-2,4-decadienal (C10) (Garbowicz *et al.*, 2018). Notably, we did not observe a correlation between short-chain FA-VOCs and the *LIP1* gene in our GWAS analysis (Tiemann *et al.*, 2017), indicating that *LIP1* is not likely to be a source of VOC variation within *S. lycopersicum* accessions. Phylogenetic analysis indicated that *SI-LIP8* is a member of the class III lipase family that mainly catalyzes hydrolysis of the ester bonds of TAGs into FFAs and glycerol. The protein clusters with *AtHIL1*, which can cleave multiple lipids with a preference for MGDG to release 18:3-FFA (Higashi *et al.*, 2018). Our enzymatic assays demonstrated that *SI-LIP8* can hydrolyze 18:2 and 18:3 acyl groups from TAGs and PCs with a preference

for 18:3-FFA (Figure 7(a); Table 1). Based on the mQTL lipid profiling data of Garbowicz *et al.* (2018), SI-LIP8 is predicted to act upon MGDG and phospholipids. Our enzymatic assays indicate that SI-LIP8 can hydrolyze TAGs and PCs.

The genome-edited *SI-LIP8* mutants exhibited considerable decreases in the levels of several C5 (1-pentanol and 1-penten-3-ol) and C6 (Z-3-hexen-1-ol, E-2-hexen-1-ol, and hexyl alcohol) FA-VOCs (Figures 4 and S6(a,b)), indicating an important role for *SI-LIP8* in flavor-associated FA-VOC synthesis. In particular, production of C6 VOCs was almost completely abolished in fruit. Nonetheless, these FA-VOCs are not entirely absent in the *SI-LIP8* mutants, indicating at least some degree of functional redundancy in the pathway.

RNA-Seq analysis of accessions carrying the reference or alternate allele at ch09_70173370 indicated that variation in gene expression is likely responsible for the variation in FA-VOC content associated with this QTL (Figure 1(b)). The transcript level of *LIP8* in *S. pennellii* ripening fruit is approximately 10% of that of the *S. lycopersicum* ortholog, again consistent with transcriptional regulation of FA-VOC contents (Garbowicz *et al.*, 2018). To investigate the underlying cause of this variation in expression, we found a ~660-bp insertion in the transcriptional promoter region of *SI-LIP8* compared to the *Sp-LIP8* promoter. This insertion contains tandemly repeated NAC transcription factor binding motifs (TTG/ACGT) (Lindemose *et al.*, 2014) (Figure 8). In climacteric fruits, including tomato, it has been demonstrated that the NAC transcription factor NONRIPENING (NOR) positively regulates ripening and autocatalytic ethylene synthesis by binding to the promoters of many ripening-associated genes and activating gene expression (Giovannoni *et al.*, 2017; Gao *et al.*, 2018; Lu *et al.*, 2018). The promoter insertion may be responsible for the variation in gene expression of *LIP8*. This result is similar to our prior observation of a large insertion into the promoter of *S. lycopersicum* *CXE1* relative to its *S. pennellii* ortholog that also resulted in much higher expression of *SICXE1* (Goulet *et al.*, 2012). Thus, within and across species, *LIP8* activity appears to be transcriptionally regulated.

Unlike many of the favorable flavor-associated VOC QTLs that we have identified, all of the modern commercial tomato varieties for which we have whole-genome sequences contain the favorable allele (Tieman *et al.*, 2017) conferring higher contents of C5 and C6 VOCs (Table 2). The alternate allele is present in some varieties but is most abundant in *S. pimpinellifolium* accessions from Ecuador. This pattern indicates that the alternate allele was lost during domestication and improvement. Consistent with that interpretation, this segment of chromosome 9 was identified as being part of an improvement sweep associated with selection for the *fw9.3* fruit weight locus (Lin *et al.*, 2014). This reference allele was likely fixed as a consequence of its linkage to a gene conferring larger fruit size.

Table 2 Allele (ch09_70173370) distribution in tomato accessions. R, reference (A); A, alternate (G); H, heterozygous (A/G)

Accessions	R	A	H
Modern	47	0	0
Transitional	34	7	1
Heirloom	209	2	2
<i>S. lycopersicum</i> var. <i>cerasiforme</i>	136	8	0
<i>S. pimpinellifolium</i>	8	16	6

Despite the strong selection in commercial tomato varieties, it is worth noting that we did not observe any obvious phenotypic differences in plant or fruit phenotypes in the engineered loss-of-function lines. While *SI-LIP8* is expressed in leaves (Figure S2), we found no difference in wound-associated C6 volatiles in leaves of the knockout lines, indicating that SI-LIP8 is not involved in wounding responses.

In summary, GWAS analysis identified a SNP on chromosome 9 that is significantly associated with short-chain FA-VOCs that contribute to tomato flavor intensity and liking. Here, we identified a gene, *SI-LIP8*, with a clear role in providing substrates for subsequent steps in the FA-VOC synthetic pathway. We confirmed that it functions as an acyl group hydrolase. Knockout of *SI-LIP8* significantly reduced the contents of multiple short-chain FA-VOCs in ripe tomato fruit. *In vitro* assays showed SI-LIP8 cleaves multiple lipid substrates into FFAs, including linoleic and linolenic acids. The activity of this first step in FA-VOC synthesis is regulated at the level of transcript abundance within and between species. The *SI-LIP8* gene is an important component of the synthetic pathway of flavor-associated VOCs that affect human preferences for tomatoes.

EXPERIMENTAL PROCEDURES

Plant materials and growth conditions

Plants were grown in a heated greenhouse on the University of Florida campus or in a field in Live Oak, Florida, using recommended commercial practices. Fruits for mRNA quantification were sampled, frozen in liquid nitrogen, and then stored at -80° . All fruits were harvested at a full-red ripe stage.

Tomato expression and metabolite data

The transcript levels were obtained from previous RNA-Seq analysis (Tomato Genome, 2012; Zhu *et al.*, 2018) and the database available from http://tea.solgenomics.net/expression_viewer/input. The metabolite contents of ILs were analyzed using publicly available data from the Tomato Functional Genomics Database (<http://ted.bti.cornell.edu/>).

Tissue collection, RNA extraction, and quantitative PCR

Total RNA was extracted from pericarp tissue of fully ripe fruit harvested at the University of Florida with an Isolate II RNA Plant Kit (BioLine, London, UK). Possible contaminating DNA was removed by DNase (Promega, Madison, WI, USA) treatment. The

quantitative reverse transcriptase-PCR (qRT-PCR) was conducted on a GeneAmp 5700 Sequence Detection System (Applied Biosystems, Waltham, MA, USA) with a SensiFAST SYBR Hi-ROX One-Step Kit (Bioline). The oligonucleotide primers of *SI-LIP8* used for qRT-PCR were designed using Primer 3 (<http://bioinfo.ut.ee/prime3-0.4.0/>) and are listed in Table S7.

Sequence and phylogenetic analysis

The NCBI database (<https://www.ncbi.nlm.nih.gov/>) was used for sequence confirmation and conserved domain analysis (CD-search) of *SI-LIP8*. The top 11 *S. lycopersicum* amino acid sequences similar to *SI-LIP8* were obtained by comparing the protein sequence against the tomato proteome using the BLASTp program in SGN BLAST (<https://solgenomics.net/tools/blast/>), and sequences were aligned using MUSCLE with default settings employing MEGA (Version X). The unrooted neighbor-joining phylogenetic tree (2000 bootstrap replicates) was built using MEGA (Version X) software.

CRISPR/Cas9 mutagenesis, plant transformation, and mutant allele selection

Two sgRNAs in close proximity (within 150 bp) in the coding sequence of the target gene were designed using the CRISPR-P (<http://cbi.hzau.edu.cn/cgi-bin/CRISPR>) tool as described by Lei *et al.* (2014). The *In vitro* efficacy test of two sgRNAs (sgRNA1 and sgRNA2) was performed using the Guide-itTM sgRNA *In vitro* Transcription and Screening System (Takara, Mountain View, CA, USA) according to the user manual. The primers used for the sgRNA efficacy test are listed in Table S7. The two sgRNAs were inserted into a CRISPR/Cas9 binary vector (pCambia2300_CR3-EF, provided by Brian Staskawicz, University of California, Berkeley, CA, USA), in which the target sequence was driven by the *Arabidopsis U6-26* promoter and *Cas9* by 2 × 35S. The recombinant vector with the two sgRNAs alongside the *Cas9 endonuclease* gene were designed to produce defined deletions in the coding sequence of *SI-LIP8*. The two sgRNA sequences are listed in Table S7.

The final binary vector was transformed into the tomato cultivar Fla. 8059 by *Agrobacterium*-mediated transformation (McCormick *et al.*, 1986) using kanamycin as a selective agent. The *NPT II*-positive T₀ plants were genotyped for deletions by sequencing with a forward primer specific for the 5'-end of sgRNA1 and a reverse primer specific for the 3'-end of sgRNA2. The plants heterozygous for the engineered deletion were self-pollinated to identify homozygous plants in the T₁ generation (*SI-LIP8CR*), and a 5-bp deletion was confirmed by sequencing. The T₁ homozygotes with deletions in the *SI-LIP8* gene were self-pollinated and T₃ homozygous plants (*SI-LIP8CR*_141, field; *SI-LIP8CR*_363 and 364, greenhouse) lacking the transgene were eventually isolated. Additionally, the T₁ homozygous plants with deletions were also backcrossed to WT, in order to achieve progeny with a genetic identity close to that of the parent. The F₁ generation from backcrossed plants (BC1-F₁) was genotyped for the desired deletion and absence of the *NPT II* transgene. Finally, plants heterozygous for the engineered deletion and lacking the *NPT II* transgene were self-pollinated to isolate homozygous plants from the BC1-F₂ generation (*SI-LIP8CR*_140, field). Additionally, genomic DNA was extracted from T₃ and BC1-F₂ tomato leaves and its flanks containing the target sites were amplified using the specific primers *SI-LIP8-F* and *SI-LIP8-R*. The PCR products were digested with *DdeI* (NEB) at 37°C for 4 h and then run on a 2% agarose gel. The primers used for amplification and genotyping are listed in Table S7.

Volatile collection and analysis

At least four ripe fruits from each line grown in randomized plots in the greenhouse or field were used for analysis, and fruits were obtained from three to four weekly harvests. Volatiles were collected from approximately 100 g chopped ripe tomato fruits during a 1-h period at room temperature as previously described (Tiemann *et al.*, 2006a, 2006b), trapped on a divinylbenzene resin column (SuperQ), and eluted with methylene chloride using nonyl acetate as an internal standard. The samples were analyzed on an Agilent 6890N gas chromatograph. Retention times were compared with known standards and identities of volatile peaks were confirmed by GC-MS (Agilent 5975, www.agilent.com).

Recombinant protein purification

The full-length and mature coding sequences of *SI-LIP8* were cloned using Fla. 8059 cDNA as the template and subcloned into the pET28a vector through *EcoRI* and *XhoI* restriction sites. Primers used are listed in Table S7. The recombinant N- and C-terminal His-tagged protein constructs were transformed into *E. coli* strain RosettaTM 2 (DE3) SinglesTM component cells (Novagen). Recombinant protein was expressed as described by Li *et al.* (2017). Briefly, 10 ml of the overnight culture was combined with 500 ml LB liquid medium (50 µg ml⁻¹ kanamycin and 34 µg ml⁻¹ chloramphenicol). Isopropyl β-D-1-thiogalactopyranoside was added to a final concentration of 0.3 mM, the bacteria were incubated at 16°C for 20 h and harvested by centrifugation (4000 rpm, 4°C, 15 min). Then the cells were resuspended in 1 × PBS buffer, followed by sonication and centrifugation, and the supernatant was purified using a HisTALONTM Gravity Column (Clontech, Mountain View, CA, USA) according to the user manual. Buffer change and clean-up of proteins were conducted using a PD-10 Desalting Column (GE Healthcare, Chicago, IL, USA). Purified proteins were concentrated using an Amicon Ultra-4 Centrifuge Filter (Millipore UFC803008), after which the proteins were quantified using the Bio-Rad Bradford (Bio-Rad, Hercules, CA, USA) assay, aliquoted, and stored at -20°C in 30% glycerol for future use. SDS-PAGE was performed and the proteins were visualized using Coomassie brilliant blue.

Site-directed mutagenesis

Lipase mutagenesis was performed using overlap extension PCR as described by Luo *et al.* (2012). Briefly, target mutations at position ch09_70393277 (G>A) and ch09_70394515 (G>T) were introduced into primers, respectively, and then the target fragments were amplified using the two previous PCR products as templates. The final products were sequenced to confirm the mutations in the protein-coding region. The *SI-LIP8*_M^{3277_A} and *SI-LIP8*_M^{4515_T} were subcloned into the pET-28a vector, and purification was conducted as described above. The primers are listed in Table S7.

NEFA assay and product identification

Compounds were purchased for *in vitro* lipase activity assay, including 54:6-TAG (T9517, purity ≥ 98%), 54:9-TAG (T6513, purity ≥ 97%), and 36:4-PC (T0537, purity ≥ 99%) from Sigma Aldrich and 36:6-PC (37-1803-0, purity > 98%) from Larodan. Each lipid was diluted in the reaction buffer containing 100 mM Tris-HCl, pH 7.5, and 5% (w/v) Gum Arabic (Sigma Aldrich, St. Louis, MO, USA) without detergents (Eastmond, 2004; Higashi *et al.*, 2018), followed by sonication for 30 sec (3 × 10 sec) on ice. The substrate solutions were stored in -20°C until use. The lipase *in vitro* reaction assay was conducted essentially as described by Wang *et al.* (2017).

For each reaction, the substrate solution (~0.3 mg) and purified SI-LIP8_F recombinant protein (~2 µg) were added, the suspension was dispersed by sonication for 15 sec, and the reaction mixture was incubated at room temperature (~25°) for 4 h. Following the incubation, the reaction was stopped by lipid extraction. The extraction was done as described by Bligh and Dyer (1959). A total of 300 µl samples dissolved in methanol:chloroform (6:1) was used for FFA determination. A NEFA Assay Kit (Solarbio, Beijing, China) was used to investigate the biochemical function of SI-LIP8, according to the manufacturer's recommendations. The FFA contents were calculated by the copper reagent method (Xu *et al.*, 2013). In addition, FFAs present in the samples were analyzed by FAME assay as described by Kail *et al.* (2012). The FAME assay used acid-catalyzed esterification in methanolic solutions under mild conditions (50°C for 10 min) to achieve complete conversion of more reactive FFAs while TAGs were largely unaffected. Heptadecanoic acid (17:0) (Sigma Aldrich) was added as an internal standard according to Li-Beisson *et al.* (2013). The retention time was compared with known standards (methyl linoleate and methyl linolenate, ACROS ORGANICS, Pittsburgh, PA, USA) and identities of volatile peaks were confirmed by GC-MS. Eluates prepared from *E. coli* containing the empty vector (pET-28a) and SI-LIP8_F without substrate were used as controls.

FFA quantitation and kinetic assay

A commercial kit (Free Fatty Acid Quantitation Kit, Sigma Aldrich) was used with palmitic acid as a standard to confirm the function of the lipase and compare the activity of SI-LIP8_F and SI-LIP8_M according to the user manual. For the functional analysis, the aforementioned four glycerolipids were used as substrates. Each substrate solution (~0.1 mg) mixed with purified lipase (SI-LIP8_M, ~0.5 µg) was incubated at 30°C for 15 min. Comparative analysis of the lipase activity was conducted with a mixture of SI-LIP8_F or SI-LIP8_M (~0.5 µg) and 36:4-PC (~0.1 mg), which were incubated at 30°C for the time course. For kinetic analysis of the recombinant SI-LIP8_M proteins, ~0.5 µg enzyme was incubated with substrate (36:4-PC or 36:6-PC) at 30°C for 15 min. The concentrations of the tested substrates range from 0 to 10 µM. All the kinetic assays were repeated in triplicate. The kinetic parameters K_m and K_{cat} were calculated using the Michaelis–Menten model using GraphPad Prism 8. All reactions were stopped by 99.9°C heat treatment, and the mixture was directly used for FFA measurements. Finally, the absorbance was measured at 570 nm using the colorimetric method. The reaction buffer used in the assays consisted of 100 mM Tris-HCl, pH 7.5, and 5% (w/v) Gum Arabic, and each mixture was sonicated for 15 sec before reaction.

Promoter, subcellular location, and transit peptide analysis

Transcription binding motifs were analyzed with PlantPAN 3.0 (<http://plantpan.itps.ncku.edu.tw/>) (Chow *et al.*, 2016; Garbowicz *et al.*, 2018). The promoter analysis of *SI/Sp-LIP8* was conducted on the accessions of *S. lycopersicum* (M82) and *S. pennellii* (LA0716). The promoter region alignment of *SI/Sp-LIP8* was conducted with Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). The subcellular location and transit peptide analysis were performed with TargetP-2.0 Server (<http://www.cbs.dtu.dk/services/TargetP/>) and ChloroP 1.1 Server (<http://www.cbs.dtu.dk/services/ChloroP/>), respectively (Padham *et al.*, 2007).

Statistical analysis

The Student *t*-test was used for two-sample comparisons (JMP Pro 15 software). The significance levels are indicated in each figure.

ACCESSION NUMBERS

Sequences of *Sp-SI-LIP8* (Sopen09g034030/Solyc09g091050), SI-LIP8 homologs, and SILIP1/2 (Solyc12g055730/Solyc03g123750) were obtained from the SGN database. Sequences of AtHIL1 (At4g13550), AtPLIP1/2/3 (At3g61680/At1g02660/At3g62590), AtDAD1 (At2g44810), AtDLL2/3/4 (At1g51440/At2g30550/At1g06800), and AtTLL1 (At1g45201) were obtained from the Arabidopsis TAIR database. GenBank accessions or NCBI reference numbers are accessible for LeLID (XP_00432963), CaPLA1 (EF595843), DcLIP (AAD01804), EgLIP1 (JX556215), and RcOBL1 (JQ945176).

ACKNOWLEDGEMENTS

This work was supported by a grant from the National Science Foundation to HK and DT (IOS 1844237) as well as Rijk Zwaan Breeding N.V. Part of this work was also supported by the National Key Research and Development Program (2016YFD0400101) and the 111 Project (B17039). The China Scholarship Council supported XL with living expenses for his research in the USA. We thank Dawn Bies and Mark Taylor for their technical assistance. The authors appreciate helpful discussions with Zhongyuan Liu.

AUTHOR CONTRIBUTIONS

XL and DT performed experiments, analyzed data, and contributed to the writing of the manuscript. ZL performed part of the experiments. DT, KC, and HK supervised part of the research. KC and HK also contributed to writing the manuscript.

CONFLICT OF INTEREST

No conflict of interest declared.

DATA AVAILABILITY STATEMENT

All of the data not previously published are provided in the Supplemental Tables.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. FA-VOC emissions in reference (R) and alternate (A) varieties at position ch09_70173370.

Figure S2. *SI-LIP8* transcript levels in reference (R) and alternate (A) varieties.

Figure S3. Alignment of SI-LIP8 with class III lipase protein sequences available in the NCBI conserved domain database (CDD).

Figure S4. Alignment of the deduced amino acid sequences of SI-LIP8 and AtHIL1.

Figure S5. Restriction analysis and sequence alignment of Cas9-edited lines.

Figure S6. The effects of SI-LIP8 mutation on FA-VOCs in tomato fruits.

Figure S7. Non-esterified fatty acid (NEFA) measurement after incubating recombinant full-length SI-LIP8_F with TAGs.

Figure S8. FA-VOC emissions in fruits containing reference (R) and alternate (A) (G>A) alleles at position ch09_70393277.

Figure S9. FA-VOC emissions in fruits containing reference (R) and alternate (A) (G>T) alleles at position ch09_70394515.

Figure S10. Expression constructs and purified proteins used for *in vitro* activity measurement.

Figure S11. Transcript levels of *SI-LIP8* in reference (R) and alternate (A) (G>A variant at position ch09_70393277 and G> T variant at position ch09_70394515) tomato varieties.

Table S1. Fatty acid-derived volatile contents and transcript levels of *SI-LIP8* (RNA-Seq) in R and A variants at position ch09_70173370.

Table S2. List of candidate genes located within 0.5 Mb (less than the average linkage disequilibrium [LD] of tomato) on either side of the lead SNP.

Table S3. Contents of FA-VOCs in *SI-LIP8* knock-out (KO) lines. Data are expressed as the relative fold change compared to the control Fla. 8059 (**P* < 0.05, ***P* < 0.01, and ****P* < 0.001).

Table S4. Sequencing data of *SI-LIP8* region.

Table S5. Fatty acid-derived volatile contents and transcript levels of *SI-LIP8* (RNA-Seq) in R and A variants at position ch09_70393277 and ch09_70394515.

Table S6. Enzymatic activity of SI-LIP8_M^{3277_G/4515_G}, SI-LIP8_M^{3277_A}, and SI-LIP8_M^{4515_T} using 36:4-PC as the substrate (*n* = 3).

Table S7. List of oligonucleotides used in this study.

Data S1. Text file of the amino acid sequence alignment of SI-LIP8 and its homologs used in Figure 2(a).

REFERENCES

- Andersson, M. and Dörmann, P. (2009) Chloroplast membrane lipid biosynthesis and transport. In *The chloroplast—Interactions with the environment* (Sandelius, A.S. and Aronsson, H., eds). New York, NY: Springer, pp. 125–158.
- Angelovici, R., Batushansky, A., Deason, N., Gonzalez-Jorge, S., Gore, M.A., Fait, A. and DellaPenna, D. (2017) Network-Guided GWAS improves identification of genes affecting free amino acids. *Plant Physiol.* **173**, 872–886.
- Bate, N.J. and Rothstein, S.J. (1998) C6-volatiles derived from the lipoxygenase pathway induce a subset of defense-related genes. *Plant J.* **16**, 561–569.
- Beck, J.J., Alborn, H.T., Block, A.K. *et al.* (2018) Interactions among plants, insects, and microbes: elucidation of inter-organismal chemical communications in agricultural ecology. *J. Agric. Food Chem.* **66**, 6663–6674.
- Bligh, E.G. and Dyer, W.J. (1959) A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **37**, 911–917.
- Boudiere, L., Michaud, M., Petroustos, D. *et al.* (2014) Glycerolipids in photosynthesis: composition, synthesis and trafficking. *Biochim. Biophys. Acta*, **1837**, 470–480.
- Bouwmeester, H., Schuurink, R.C., Bleeker, P.M. and Schiestl, F. (2019) The role of volatiles in plant communication. *Plant J.* **100**, 892–907.
- Buttery, R.G., Teranishi, R., Flath, R.A. and Ling, L.C. (1989) Fresh tomato volatiles. In *Flavor Chemistry: Trends and Developments* (Teranishi, R., Buttery, R.G., and Shahidi, F., eds). Washington, DC: American Chemical Society, pp. 213–222.
- Campos, M.L., Yoshida, Y., Major, I.T. *et al.* (2016) Rewiring of jasmonate and phytochrome B signalling uncouples plant growth-defense tradeoffs. *Nat. Commun.* **7**, 12570.
- Chen, G., Hackett, R., Walker, D., Taylor, A., Lin, Z. and Grierson, D. (2004) Identification of a specific isoform of tomato lipoxygenase (TomloxC) involved in the generation of fatty acid-derived flavor compounds. *Plant Physiol.* **136**, 2641–2651.
- Chow, C.N., Zheng, H.Q., Wu, N.Y. *et al.* (2016) PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Res.* **44**, D1154–D1160.
- Croft, K., Juttner, F. and Slusarenko, A.J. (1993) Volatile products of the lipoxygenase pathway evolved from *Phaseolus vulgaris* (L.) Leaves inoculated with *Pseudomonas syringae* pv *phaseolicola*. *Plant Physiol.* **101**, 13–24.
- Dicke, M. and Baldwin, I.T. (2010) The evolutionary context for herbivore-induced plant volatiles: beyond the 'cry for help'. *Trends Plant Sci.* **15**, 167–175.
- Eastmond, P.J. (2004) Cloning and characterization of the acid lipase from castor beans. *J. Biol. Chem.* **279**, 45540–45545.
- Erb, M., Meldau, S. and Howe, G.A. (2012) Role of phytohormones in insect-specific plant reactions. *Trends Plant Sci.* **17**, 250–259.
- Fang, C. and Luo, J. (2019) Metabolic GWAS-based dissection of genetic bases underlying the diversity of plant metabolism. *Plant J.* **97**, 91–100.
- Folta, K.M. and Klee, H.J. (2016) Sensory sacrifices when we mass-produce mass produce. *Hortic. Res.* **3**, 16032.
- Gao, Y., Wei, W., Zhao, X. *et al.* (2018) A NAC transcription factor, NOR-like1, is a new positive regulator of tomato fruit ripening. *Hortic. Res.* **5**, 75.
- Garbowicz, K., Liu, Z., Alseekh, S. *et al.* (2018) Quantitative trait loci analysis identifies a prominent gene involved in the production of fatty acid-derived flavor volatiles in tomato. *Mol. Plant*, **11**, 1147–1165.
- Giovannoni, J., Nguyen, C., Ampofo, B., Zhong, S. and Fei, Z. (2017) The epigenome and transcriptional dynamics of fruit ripening. *Annu. Rev. Plant Biol.* **68**, 61–84.
- Goulet, C., Kamiyoshihara, Y., Lam, N.B., Richard, T., Taylor, M.G., Tieman, D.M. and Klee, H.J. (2015) Divergence in the enzymatic activities of a tomato and *Solanum pennellii* alcohol acyltransferase impacts fruit volatile ester composition. *Mol. Plant*, **8**, 153–162.
- Goulet, C., Mageroy, M.H., Lam, N.B., Floystad, A., Tieman, D.M. and Klee, H.J. (2012) Role of an esterase in flavor volatile variation within the tomato clade. *Proc. Natl. Acad. Sci. USA*, **109**, 19009–19014.
- Guo, Q., Major, I.T. and Howe, G.A. (2018) Resolution of growth-defense conflict: mechanistic insights from jasmonate signaling. *Curr. Opin. Plant Biol.* **44**, 72–81.
- Hammerbacher, A., Coutinho, T.A. and Gershenzon, J. (2019) Roles of plant volatiles in defence against microbial pathogens and microbial exploitation of volatiles. *Plant Cell Environ.* **42**, 2827–2843.
- Higashi, Y., Okazaki, Y., Takano, K., Myouga, F., Shinozaki, K., Knoch, E., Fukushima, A. and Saito, K. (2018) HEAT INDUCIBLE LIPASE1 remodels chloroplastic monogalactosyldiacylglycerol by liberating alpha-linolenic acid in arabidopsis leaves under heat stress. *Plant Cell*, **30**, 1887–1905.
- Holopainen, J.K. and Gershenzon, J. (2010) Multiple stress factors and the emission of plant VOCs. *Trends Plant Sci.* **15**, 176–184.
- Howe, G.A., Lee, G.I., Itoh, A., Li, L. and DeRocher, A.E. (2000) Cytochrome P450-dependent metabolism of oxylipins in tomato. Cloning and expression of allene oxide synthase and fatty acid hydroperoxide lyase. *Plant Physiol.* **123**, 711–724.
- Howe, G.A., Major, I.T. and Koo, A.J. (2018) Modularity in jasmonate signaling for multistress resilience. *Annu. Rev. Plant Biol.* **69**, 387–415.
- Ishiguro, S., Kawai-Oda, A., Ueda, J., Nishida, I. and Okada, K. (2001) The DEFECTIVE IN ANTHR DEHISCENCE1 gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in Arabidopsis. *Plant Cell*, **13**, 2191–2209.
- Kail, B.W., Link, D.D. and Morreale, B.D. (2012) Determination of free fatty acids and triglycerides by gas chromatography using selective esterification reactions. *J. Chromatogr. Sci.* **50**, 934–939.
- Kelly, A.A. and Feussner, I. (2016) Oil is on the agenda: lipid turnover in higher plants. *Biochim. Biophys. Acta*, **1861**, 1253–1268.
- Klee, H.J. (2010) Improving the flavor of fresh fruits: genomics, biochemistry, and biotechnology. *New Phytol.* **187**, 44–56.
- Klee, H.J. and Tieman, D.M. (2013) Genetic challenges of flavor improvement in tomato. *Trends Genet.* **29**, 257–262.
- Klee, H.J. and Tieman, D.M. (2018) The genetics of fruit flavour preferences. *Nat. Rev. Genet.* **19**, 347–356.
- Kuhalskaya, A., Ahchige, M.W., Perez de Souza, L., Vallarino, J., Brotman, Y. and Alseekh, S. (2020) Network analysis provides insights into tomato lipid metabolism. *Metabolites*, **10**, 152. <https://doi.org/10.3390/metab10040152>
- Lei, Y., Lu, L., Liu, H.Y., Li, S., Xing, F. and Chen, L.L. (2014) CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. *Mol. Plant*, **7**, 1494–1496.
- Li-Beisson, Y., Shorrosh, B., Beisson, F. *et al.* (2013) Acyl-lipid metabolism. *Arabidopsis Book*, **11**, e0161.

- Li, X., Xu, Y., Shen, S., Yin, X., Klee, H., Zhang, B., Chen, K. and Hancock, R. (2017) Transcription factor CitERF71 activates the terpene synthase gene CitTPS16 involved in the synthesis of E-geraniol in sweet orange fruit. *J. Exp. Bot.* **68**, 4929–4938.
- Lin, T., Zhu, G., Zhang, J. *et al.* (2014) Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* **46**, 1220–1226.
- Lindemose, S., Jensen, M.K., Van de Velde, J., O'Shea, C., Heyndrickx, K.S., Workman, C.T., Vandepoele, K., Skriver, K. and De Masi, F. (2014) A DNA-binding-site landscape and regulatory network analysis for NAC transcription factors in *Arabidopsis thaliana*. *Nucleic Acids Res.* **42**, 7681–7693.
- Lu, P., Yu, S., Zhu, N. *et al.* (2018) Genome encode analyses reveal the basis of convergent evolution of fleshy fruit ripening. *Nat. Plants*, **4**, 784–791.
- Luo, L., Wang, S. and Wang, Y. (2012) Site-directed mutagenesis based on overlap extension PCR. *Agric. Sci. Technol. Hunan*, **13**, 719–722.
- Matsuda, F., Nakabayashi, R., Yang, Z., Okazaki, Y., Yonemaru, J., Ebana, K., Yano, M. and Saito, K. (2015) Metabolome-genome-wide association study dissects genetic architecture for generating natural variation in rice secondary metabolism. *Plant J.* **81**, 13–23.
- Matsui, K., Fukutomi, S., Ishii, M. and Kajiwara, T. (2004) A tomato lipase homologous to DAD1 (LeLID1) is induced in post-germinative growing stage and encodes a triacylglycerol lipase. *FEBS Lett.* **569**, 195–200.
- McCormick, S., Niedermeier, J., Fry, J., Barnason, A., Horsch, R. and Fraley, R. (1986) Leaf disc transformation of cultivated tomato (*L. esculentum*) using *Agrobacterium tumefaciens*. *Plant Cell Rep.* **5**, 81–84.
- Padham, A.K., Hopkins, M.T., Wang, T.W., McNamara, L.M., Lo, M., Richardson, L.G., Smith, M.D., Taylor, C.A. and Thompson, J.E. (2007) Characterization of a plastid triacylglycerol lipase from *Arabidopsis*. *Plant Physiol.* **143**, 1372–1384.
- Peng, M., Shahzad, R., Gul, A. *et al.* (2017) Differentially evolved glucosyl-transferases determine natural variation of rice flavone accumulation and UV-tolerance. *Nat. Commun.* **8**, 1975.
- Porta, H. and Rocha-Sosa, M. (2002) Plant lipoxigenases. Physiological and molecular features. *Plant Physiol.* **130**, 15–21.
- Rowen, E. and Kaplan, I. (2016) Eco-evolutionary factors drive induced plant volatiles: a meta-analysis. *New Phytol.* **210**, 284–294.
- Sauvage, C., Segura, V., Bauchet, G., Stevens, R., Do, P.T., Nikoloski, Z., Fernie, A.R. and Causse, M. (2014) Genome-Wide association in tomato reveals 44 candidate loci for fruit metabolic traits. *Plant Physiol.* **165**, 1120–1132.
- Scala, A., Allmann, S., Mirabella, R., Haring, M.A. and Schuurink, R.C. (2013) Green leaf volatiles: a plant's multifunctional weapon against herbivores and pathogens. *Int. J. Mol. Sci.* **14**, 17781–17811.
- Schwab, W., Davidovich-Rikanati, R. and Lewinsohn, E. (2008) Biosynthesis of plant-derived flavor compounds. *Plant J.* **54**, 712–732.
- Shen, J., Tieman, D., Jones, J.B., Taylor, M.G., Schmelz, E., Huffaker, A., Bies, D., Chen, K. and Klee, H.J. (2014) A 13-lipoxygenase, TomloxC, is essential for synthesis of C5 flavour volatiles in tomato. *J. Exp. Bot.* **65**, 419–428.
- Speirs, J., Lee, E., Holt, K., Yong-Duk, K., Steele Scott, N., Loveys, B. and Schuch, W. (1998) Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavor aldehydes and alcohols. *Plant Physiol.* **117**, 1047–1058.
- Tieman, D., Bliss, P., McIntyre, L.M. *et al.* (2012) The chemical interactions underlying tomato flavor preferences. *Curr. Biol.* **22**, 1035–1039.
- Tieman, D., Taylor, M., Schauer, N., Fernie, A.R., Hanson, A.D. and Klee, H.J. (2006a) Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde. *Proc. Natl. Acad. Sci. USA*, **103**, 8287–8292.
- Tieman, D., Zhu, G., Resende, M.F. Jr *et al.* (2017) A chemical genetic roadmap to improved tomato flavor. *Science*, **355**, 391–394.
- Tieman, D.M., Zeigler, M., Schmelz, E.A., Taylor, M.G., Bliss, P., Kirst, M. and Klee, H.J. (2006b) Identification of loci affecting flavour volatile emissions in tomato fruits. *J. Exp. Bot.* **57**, 887–896.
- Tomato Genome, C. (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, **485**, 635–641.
- Wang, K., Durrett, T.P. and Benning, C. (2019) Functional diversity of glycerolipid acylhydrolases in plant metabolism and physiology. *Prog. Lipid Res.* **75**, 100987.
- Wang, K., Froehlich, J.E., Zienkiewicz, A., Hersh, H.L. and Benning, C. (2017) A plastid phosphatidylglycerol lipase contributes to the export of acyl groups from plastids for seed oil biosynthesis. *Plant Cell*, **29**, 1678–1696.
- Wang, K., Guo, Q., Froehlich, J.E., Hersh, H.L., Zienkiewicz, A., Howe, G.A. and Benning, C. (2018) Two abscisic acid-responsive plastid lipase genes involved in jasmonic acid biosynthesis in *Arabidopsis thaliana*. *Plant Cell*, **30**, 1006–1022.
- Xu, B., Miao, W., Hu, Q., Gao, C. and Dong, Y. (2013) A modified colorimetric method for determining the activity of wheat germ lipase in low-aqueous media. *Qual. Assur. Safety Crops Foods*, **5**, 113–118.
- Ye, J., Wang, X., Hu, T. *et al.* (2017) An InDel in the promoter of AI-ACTIVATED MALATE TRANSPORTER9 selected during tomato domestication determines fruit malate contents and aluminum tolerance. *Plant Cell*, **29**, 2249–2268.
- Zhao, J., Sauvage, C., Zhao, J. *et al.* (2019) Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. *Nat. Commun.* **10**, 1534.
- Zhou, Y., Ma, Y., Zeng, J. *et al.* (2016) Convergence and divergence of bitterness biosynthesis and regulation in Cucurbitaceae. *Nat. Plants*, **2**, 16183.
- Zhu, G., Wang, S., Huang, Z. *et al.* (2018) Rewiring of the fruit metabolome in tomato breeding. *Cell*, **172**(1–2), 249–261.e12.