https://doi.org/10.1093/plphys/kiac509 PLANT PHYSIOLOGY 2023: 191: 110–124

Plant Physiology®

Research Article

A family of methyl esterases converts methyl salicylate to salicylic acid in ripening tomato fruit

Elizabeth M. Frick [1], Manoj Sapkota [2,3], Lara Pereira [2,3,†], Yanbing Wang [2,3], Anna Hermanns [4], James J. Giovannoni [5], Esther van der Knaap [2,3], Denise M. Tieman [1] and Harry J. Klee [1,*]

1 Horticultural Sciences, University of Florida, Gainesville, Florida 32611, USA
2 Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Athens, Georgia 30602, USA
3 Department of Horticulture, University of Georgia, Athens, Georgia 30602, USA
4 Section of Plant Breeding and Genetics, Cornell University, Ithaca, New York 14853, USA
5 United States Department of Agriculture-Agricultural Research Service and Boyce Thompson Institute, Cornell University campus, Ithaca, New York 14853, USA

*Author for correspondence: hjklee@ufl.edu
†Present address: Ecology and Evolutionary Biology, School of Biosciences, University of Sheffield, Sheffield, S10 2TN, UK.

E.M.F., M.S., J.J.G., E.vdK., D.T., and H.J.K. supervised the research. E.M.F., M.S., and A.H. performed the research. E.M.F., M.S., A.H., L.P.G., and Y.W. analyzed the data. E.M.F. wrote the manuscript with contributions from M.S., L.P.G., D.T., E.vdK., and H.J.K.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plphys/pages/general-instructions) is Harry Klee (hjklee@ufl.edu).

Abstract

Methyl salicylate imparts a potent flavor and aroma described as medicinal and wintergreen that is undesirable in tomato (*Solanum lycopersicum*) fruit. Plants control the quantities of methyl salicylate through a variety of biosynthetic pathways, including the methylation of salicylic acid to form methyl salicylate and subsequent glycosylation to prevent methyl salicylate emission. Here, we identified a subclade of tomato methyl esterases, SALICYLIC ACID METHYL ESTERASE1-4, responsible for demethylation of methyl salicylate to form salicylic acid in fruits. This family was identified by proximity to a highly significant methyl salicylate genome-wide association study locus on chromosome 2. Genetic mapping studies in a biparental population confirmed a major methyl salicylate locus on chromosome 2. Fruits from *SlMES1* knockout lines emitted significantly ($P < 0{,}05$, t test) higher amounts of methyl salicylate than wild-type fruits. Double and triple mutants of *SlMES2*, *SlMES3*, and *SlMES4* emitted even more methyl salicylate than *SlMES1* single knockouts—but not at statistically distinguishable levels—compared to the single mutant. Heterologously expressed SlMES1 and SlMES3 acted on methyl salicylate in vitro, with SlMES1 having a higher affinity for methyl salicylate than SlMES3. The *SlMES* locus has undergone major rearrangement, as demonstrated by genome structure analysis in the parents of the biparental population. Analysis of accessions that produce high or low levels of methyl salicylate showed that *SlMES1* and *SlMES3* genes expressed the highest in the low methyl salicylate lines. None of the *MES* genes were appreciably expressed in the high methyl salicylate-producing lines. We concluded that the *SlMES* gene family encodes tomato methyl esterases that convert methyl salicylate to salicylic acid in ripe tomato fruit. Their ability to decrease methyl salicylate levels by conversion to salicylic acid is an attractive breeding target to lower the level of a negative contributor to flavor.

Received May 2, 2022. Accepted October 13, 2022. Advance access publication October 31, 2022

Introduction

Tomatoes (*Solanum lycopersicum*) are the second most valuable crop in the USA, valued at over 2 billion USD per year (Agricultural Marketing Resource Center, 2018), but are consistently criticized by consumers for a lack of flavor (Bruhn et al., 1991). Consumers' perception of tomato flavor is influenced by a combination of sugars (fructose and glucose), acids (citric and malic), and approximately 30 volatile aroma compounds (Tieman et al., 2012). The levels of flavor volatiles can substantially affect fruit liking independent of sugars or acids (Zhang et al., 2016). Additionally, these flavor volatiles are present at extremely low levels and can vary by more than 1,000-fold between cultivars (Tieman et al., 2012), representing an attractive target for breeders to enhance flavor without negatively impacting yield. Despite the importance of volatiles to flavor, the biosynthetic pathways for many volatiles are incompletely described. This lack of knowledge complicates breeding efforts to alter targeted volatile levels.

Flavor is the result of the interaction between taste receptors in the mouth and olfactory receptors in the nose. Differences in sequence, expression, and copy number of human taste and smell receptors can have a large effect on flavor preferences. For example, single nucleotide polymorphisms (SNPs) in a human sweet receptor lead to marked differences in sensitivity to sucrose between individuals (Fushan et al., 2009). Olfactory receptors, of which humans have about 440, contribute to flavor through retronasal olfaction. For example, sequence changes in just two receptors significantly affect perceptions of one tomato flavor-associated volatile, guaiacol (Trimmer et al., 2019). Individuality in flavor perceptions makes it challenging for breeders to prioritize flavor because there is no perfect tomato flavor for all consumers. Moreover, breeders would need to rely on taste panels to test possible consumer acceptance and these are costly and time-consuming. One solution is to link a flavor phenotype with molecular genetic markers and to incorporate the beneficial alleles in breeding programs aimed at creating superior flavor varieties. This approach is less time-consuming than performing taste panels or biochemical quantification and removes the bias of individual preferences.

Because flavor is not a priority in most breeding programs, most modern cultivars contain far less volatile flavor compounds that positively impact consumer liking compared to heirloom varieties (Tieman et al., 2017). Exacerbating the problem, one of the few flavor volatiles that has increased in modern tomato is methyl salicylate (Tieman et al., 2017). Methyl salicylate is a flavor chemical described as tasting medicinal and wintergreen (MacLeod et al., 1982), and it is negatively correlated with liking (Krumbein and Auerswald, 1998; Zanor et al., 2009; Tieman et al., 2017). However, methyl salicylate is also a defense compound in plants, acting as the mobile derivative of the primary defense hormone salicylic acid (Park et al., 2007). Tomato fruits with high salicylic acid and methyl salicylate levels are less susceptible to common postharvest infections such as *Botrytis cinerea* (Wang et al., 2011; Zhang et al., 2017; Min et al., 2018) and against abiotic injury from chilling (Zhang et al., 2011). The dual role of methyl salicylate as both a negative flavor chemical and as a beneficial defense molecule presents a challenge in the effort to balance these competing goals of a flavorful tomato and highly resistant variety.

Much of the biochemical pathway involved in the synthesis and catabolism of methyl salicylate has been described, although not completely in any single species (Figure 1). Methyl salicylate is derived from salicylic acid, which can be synthesized from phenylalanine via cinnamic acid or from isochorismate (reviewed in Métraux, 2002). The balance between salicylic acid, methyl salicylate, and methyl salicylate derivatives is regulated through the action of multiple enzymes. Methylation of salicylic acid to produce methyl salicylate is performed by salicylic acid methyl transferases (SAMTs) in tomato fruit (Tieman et al., 2010), tea (*Camelina sinensis*) leaves (Deng et al., 2017), soybean (*Glycine max*) (Lin et al., 2013), rice (*Oryza sativa*) (Koo et al., 2007), Arabidopsis (*Arabidopsis thaliana*) (Chen et al., 2003), snapdragon (*Antirrhinum majus*) (Negre et al., 2002), *Clarkia breweri* (Ross et al., 1999), and lily (*Lilium* "Yelloween") (Wang et al., 2015). The reverse reaction, removal of a methyl group from methyl salicylate to produce salicylic acid, is performed by methyl esterases that have been described in multiple species, including *Nicotiana benthamiana* (Forouhar et al., 2005), potato (*Solanum tuberosum*) (Manosalva et al., 2010), and peach (*Prunus persica*) (Cao et al., 2019). Thus far no enzyme that converts methyl salicylate to salicylic acid has been described in tomato. In addition to the conversion between methyl salicylate and salicylic acid, plants use other mechanisms to regulate the pool of salicylate derivatives. The action of NONSMOKY GLYCOSYLTRANSFERASE (NSGT) prevents the emission of methyl salicylate and guaiacol via glycosylation, forming a nonvolatile glucoside (Tikunov et al., 2013). The *NSGT1* locus in tomato exhibits both sequence polymorphism and structural variation among varieties. The presence of functional copies of *NSGT1* leads to more glycosylation and lower emission of the negative flavor volatile guaiacol (Alonge et al., 2020).

In this work, we describe a subclade of tomato esterases that convert methyl salicylate to salicylic acid in ripe tomato fruit. This esterase locus is an attractive breeding target to decrease the amount of the negative flavor compound methyl salicylate.

Results

Identification of MeSA2.1 controlling methyl salicylate levels in tomato fruit

To identify genes affecting methyl salicylate emission, a population of 166 tomato accessions encompassing the wild relative *Solanum pimpinellifolium* (SP), semi-domesticated cherry tomato *S. lycopersicum* var.*cerasiforme* (SLC), and the domesticated large-fruited *S. lycopersicum* var. *lycopersicum*

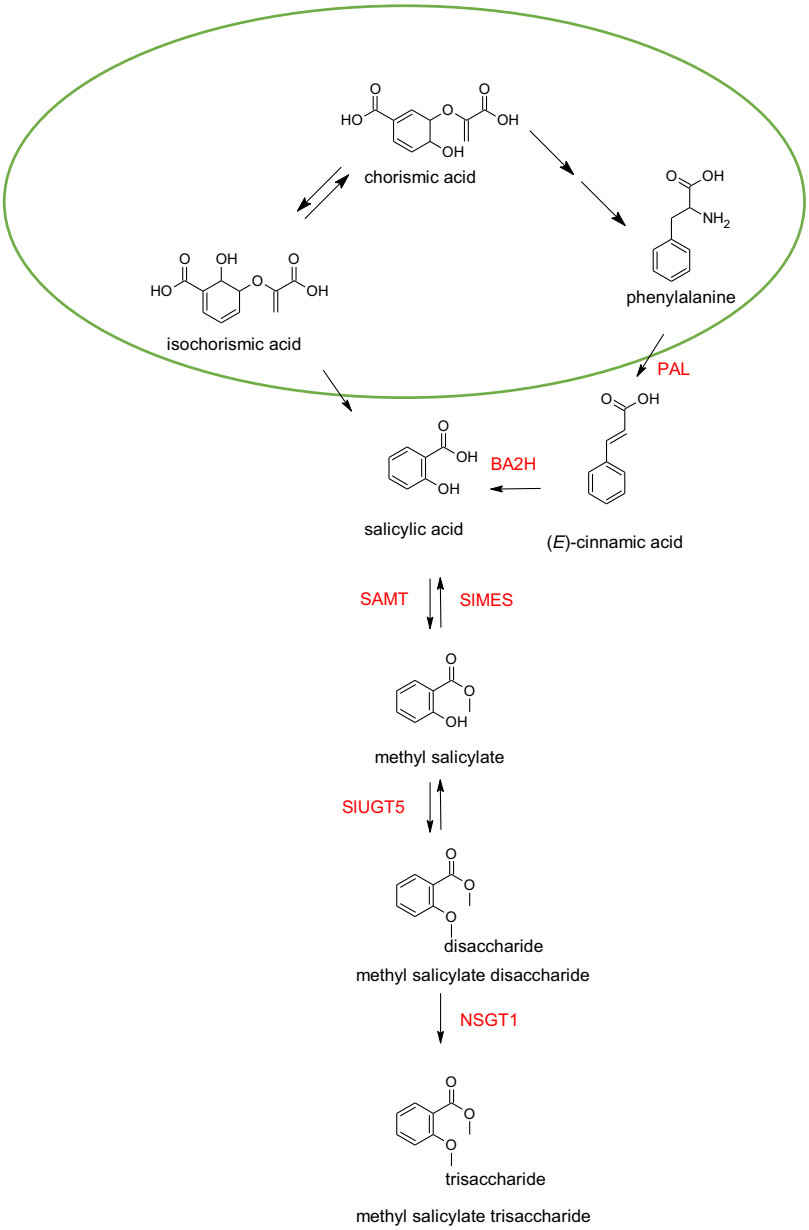

Figure 1 Overview of the methyl salicylate and salicylic acid biosynthesis pathway. Pathway components located in the chloroplast are found within the green circle. Enzymes are in red. PAL, phenyalanine ammonia lyase; BA2H, benzoic acid 2-hydroxylase; SlMES, methyl esterase; SlUGT5, UDP glycosyltransferase; NSGT1, non-smoky glycosyltransferase 1.

(SLL) was used (Razifard et al., 2020). The genome-wide association study (GWAS) for methyl salicylate levels identified loci on chromosomes 2 and 9 that were highly associated with methyl salicylate levels (Figure 2). To confirm the importance of the chromosome 2 locus, we developed a biparental population from two SLC accessions, BGV006931 and BGV014508, which produced near the highest (17.3 ng gfw^{-1} h^{-1}) and the lowest (0.02 ng gfw^{-1} h^{-1}) amounts of methyl salicylate within the GWAS panel. A list of varieties used in this work and a brief description are in Supplemental Table S1. Methyl salicylate levels in the different F_2 plants (n = 145) showed a skewed distribution that was indicative of a recessive trait controlled by more than one locus (Figure 3A). We generated molecular markers for the two GWAS loci on chromosomes 2 and 9 and found two quantitative trait loci (QTLs) each with logarithm of the odds (LOD) scores of 9.6 that were designated *MeSA2.1* and *MeSA9.1* (Figure 3B). These two QTLs together accounted for 20.2% and 19.2% of the phenotypic variance that combined led to a 42.4% effect on the volatile levels in the segregating population (Figure 3, B and C). The locus *MeSA9.1* overlapped with *NSGT1*, which has known activity on both methyl salicylate and guaiacol (Tikunov et al., 2013). Importantly, the recessive and nonfunctional allele of *NSGT1* combined with the recessive allele of *MeSA2.1* contributed to high methyl salicylate levels at a significance level of

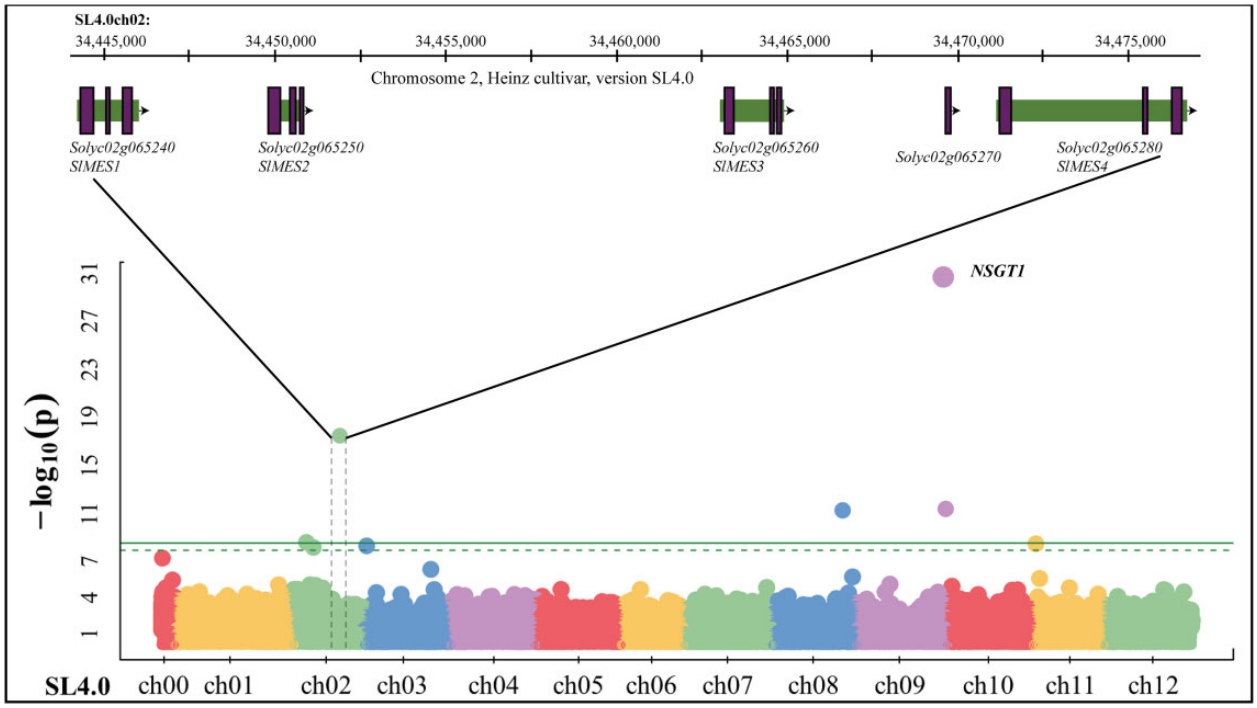

Figure 2 Identification of *SlMES* candidate genes by a GWAS. The *X*-axis represents different chromosomes in different colors and the *Y*-axis represents the −log10(*P*-value) of the variants used for analysis. Two major loci on chromosomes 2 and 9 are associated with methyl salicylate production. The locus on chromosome 2 (green) overlaps with a cluster of methyl esterases, whereas the locus on chromosome 9 (purple) corresponds to the *NSGT1* locus (Tikunov et al., 2013). The green horizontal solid and dashed lines represent significance thresholds for the association a significant level of 5% and 1%, respectively. The zoomed in region from most significant peak on chromsome 2 represents the gene models in *MES* locus. The green and purple bars represents the introns and exons of the genes, respectively, with each gene's name and Solyc number below the models. The scale above the gene model represents the scale of physical position of the genes in the tomato genome.

$P < 7.4\ e^{-13}$ for the interaction. The significant SNPs from the GWAS at the *MeSA2.1* locus and the most significant marker in the biparental population were less than 1 kb away from *Solyc02g065240* (Figure 2). The genome structure of *MeSA2.1* in the two parents was dramatically different (Supplemental Figure S1). The low methyl salicylate-producing parent BGV014508 is more similar to the tomato reference genome, while the high-producing parent BGV006931 has shorter distances between genes at the locus (Figure 3D; Supplemental Figure S1). There are four full-length genes annotated as methyl esterases at *MeSA2.1* in SL4.0, which we have named *SlMES1-4* (NCBI Gene IDs: *SlMES1*-101261578 *SlMES2*-101261293, *SlMES3*-112940921 *SlMES4*-101260990) (Figures 2 and 3, D).

MeSA2.1 is a major effect locus controlling methyl salicylate accumulation

Because methyl salicylate is a polygenic trait with several important loci characterized, we wanted to determine the effect of variation at *SlMES* without contribution from other genes, particularly the chromosome 9 *NSGT* locus. We compared two SLC line accessions that were nearly identical for many of the known genes in the methyl salicylate pathway but differ at the *MeSA2.1* locus (Table 1 and Figure 4). These known genes are a catechol O-methyltransferase (*Solyc10g005060*) which methylates the precursor catechol (Mageroy et al., 2012), salicylic acid methyltransferases (*Solyc09g091530* and *Solyc09g091540*) which converts salicylic acid into methylsalicylate (Tieman et al., 2010), and the previously described *NSGT1* locus (Tikunov et al., 2013; Alonge et al., 2020) (Table 1). BGV008218 is most similar to the Heinz 1706 reference genome and has low methyl salicylate levels. BGV006779 has multiple insertions and SNPs compared to the reference sequence at *MeSA2.1* and contains high levels of methyl salicylate (Table 1). The significant difference in methyl salicylate accumulation shows that *MeSA2.1* likely affects methyl salicylate independently from other known methyl salicylate genes (Figure 4).

Molecular characterization of SlMES1-4

A phylogenetic analysis of the encoded *MES* complementary DNAs (cDNAs) showed that *MES1* (*Solyc02g065240*) was most likely ancestral to *MES2*, *MES3*, and *MES4* (Figure 5A). In the Heinz1706 reference tomato, *SlMES1* and *SlMES4* were predicted to encode proteins of approximately 200 amino acids, whereas *SlMES2* and *SlMES3* were shorter (Figure 5, B and C). *SlMES2* carried a premature stop codon and *SlMES3* was truncated due to an insertion leading to shorter version of the protein. The full-length *MES* genes encoded proteins that comprised 7–8 motifs, which was consistent with methyl esterases from potato, *N. benthamiana*, and Arabidopsis (Figure 5, B and C). We amplified cDNA

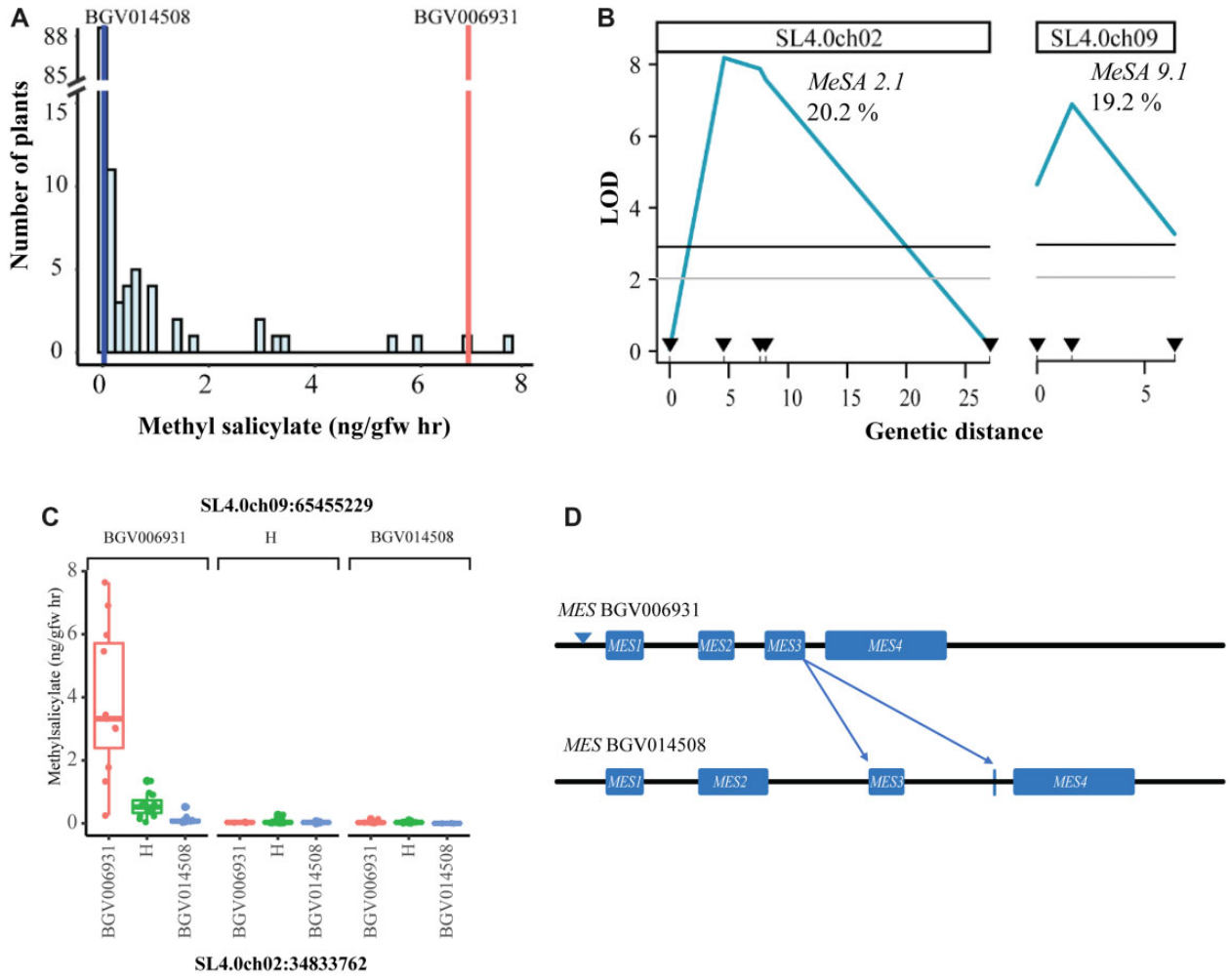

Figure 3 Mapping of *MeSA2.1* (QTL associated with methyl salicylate on chromosome 2) and *MeSA9.1* (QTL associated with methyl salicylate on chromosome 9) in a biparental population. A, Frequency histogram of the number of plants (*Y*-axis) and methyl salicylate levels in nanogram per gram fruit weight per hour (*X*-axis). *X*-axis has a scale break from 15 to 85 to accommodate the high level of skewness for low levels of methyl salicylate and present the low counts of accessions having high level of methyl salicylate. B, Composite interval mapping of methyl salicylate with the markers indicated by the black triangles. logarithm of the odds (LOD) score (blue) and confidence interval at 95% (gray) and 99% (black) are represented. The percentage of trait variation accounted by each significant QTL is also shown. LOD is calculated as −log10(*P*-value) and is the measure of strength of evidence for the presence of a QTL at a particular location. C, Strong synergistic interaction between *MES* and *NSGT1* loci on methyl salicylate levels. The boxplots represents the interaction between *MES* and *NSGT1* loci using the most significant marker (position provided in the parenthesis below the locus name) from each locus identified from QTL mapping. The *Y*-axis represents the methyl salicylate levels in red ripe fruits (in nanogram per gram fruit weight per hour) of F_2 plants. The primary *X*-axis represents different alleles at *MES* (BGV006931, BGV014508, and Heterozygous alleles) represented by the most significant KASP marker on chromosome 2 from QTL mapping and secondary *X*-axis represents different alleles of *NSGT1* (BGV006931, BGV014508, and Heterozygous alleles) represented by the most significant marker on chromosome 9. The center line is the median, with individual measurements represented as dots, and the upper and lower quartiles shown in boxes. D, A representation of the parental haplotypes demonstrate that the four *SlMES* genes are at different spacing from one another and an insertion in *SlMES3* leading to a truncated version of the protein.

from two accessions that lack the insertion in *SlMES3*, BGV006775, and PAS014479, and both transcribed a longer version of the gene which is predicted to encode two additional functional protein motifs (Figure 5B).

We next investigated the expression of the four *MES* genes by reverse transcription–quantitative PCR (RT–qPCR) using RNA extracted from ripe tomato fruits. The accessions chosen either produced very high or very low levels of methyl salicylate and included the high methyl salicylate-producing parent from the mapping population. Expression of all four *MES* genes was low to nondetectable in the high methyl salicylate-accumulating accessions (Figure 6). Low expression is consistent with the genetic analyses predicting loss-of-function and recessive alleles. In low methyl salicylate-producing accessions, *SlMES1* and *SlMES3* were well expressed in ripe tomato fruit. The expression of *SlMES2* was undetectable in nearly all accessions whereas the expression of *SlMES4* was low or undetectable regardless of methyl salicylate production. Functional MES proteins contain five motifs, 1 through 5, that are predicted to form

Table 1 Genotype of BGV008218 and BGV006779 at known methyl salicylate loci

Gene	Annotation	Reference	BGV008218 SNPs	BGV006779 SNPs
Solyc10g005060	Catechol O-methyltransferase	Mageroy et al. (2012)	Insertion at SL25ch10p55319	Insertion at SL25ch10p55319
Solyc09g091530	SAMT	Tieman et al. (2010)	None	None
Solyc09g091540	SAMT	Tieman et al. (2010)	SNP at SL25ch09p70794373; insertion at SL25ch09p70794965	Insertion at SL25ch09p70794965
NSGT1 locus	Nonsmoky glycosyltransferase	Tikunov et al. (2013); Alonge et al. (2020)	Truncated NSGT1	Truncated NSGT1
SlMES1	Methyl esterase	This work	2 SNPs	7 SNPs
SlMES2	Methyl esterase	This work	2 insertions, 1 SNP	9 insertions, 3 SNPs
SlMES3	Methyl esterase	This work	1 SNP	16 insertions, 29 SNPs
SlMES4	Methyl esterase	This work	5 insertions, 2 SNPs	24 insertions, 20 SNPs

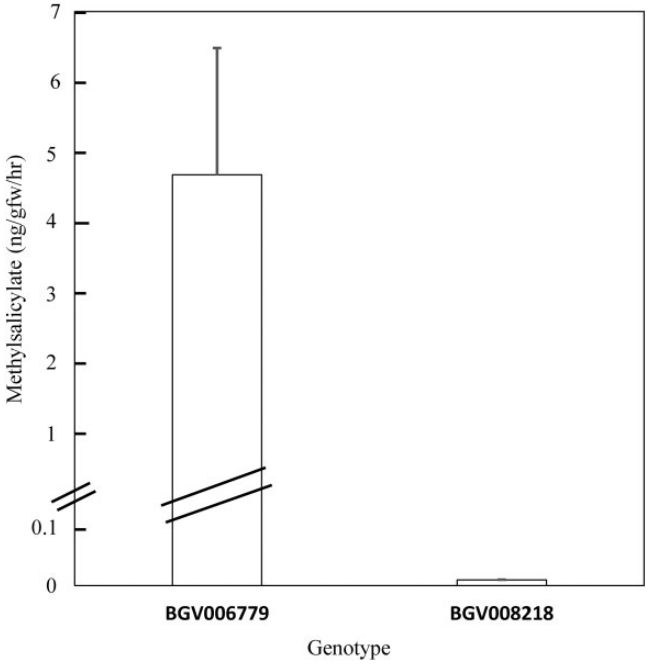

Figure 4 *SlMES* affects methyl salicylate levels independently of characterized methyl salicylate synthetic genes. Methyl salicylate levels in BGV008218 and BGV006779, which contain the same *NSGT1* haplotype. Methyl salicylate was collected from the headspace of chopped tomato fruit for 1 h, then eluted and run on an Agilent LC–MS. Values are reported as the nanograms of methyl salicylate per gram of fresh tomato weight per hour of collection and represent the average of at least three technical and biological replicates collected throughout the growing season; bars are standard error. Each replicate contained five to six plants.

the active site pocket and catalytic residues. *SlMES2* and *SlMES3* were predicted to encode a shorter protein of approximately 150 amino acids. The shorter proteins lack Motifs 2 and 5 which contain catalytic and pocket residues (Figure 5B). Both SlMES2 and SlMES3 were predicted to retain Motifs 1, 3, and 4, which also contained some of the pocket and catalytic residues. Therefore, the activity of these two MES proteins was expected to be lower or abolished. Combined, these results suggested that the most important gene in methyl salicylate volatile production in tomato fruits is *SlMES1* because the gene is expressed in ripe fruits and predicted to encode a full-length protein.

SlMES1 acts on methyl salicylate in planta

To test the roles of the *MES* genes on the methyl salicylate to salicylic acid conversion in tomato fruit, we used CRISPR editing to target subsets of the genes in two cultivars, Brandywine Sudduth and Moneymaker. All guide RNAs targeted the first exon of genes. In Brandywine Sudduth, we obtained frameshift deletions in *SlMES1* in one transgenic line and both *SlMES1* and *SlMES2* in another line (Table 2). In both cases, the frameshift causes a premature stop codon in the first exon (Supplemental Table S2). In Moneymaker, we obtained frameshift deletions in *SlMES1/SlMES3* and *SlMES1/SlMES3/SlMES4*, which also result in premature stop codons (Supplemental Table S2). Multiple alleles of *SlMES1* single mutants were evaluated in field and greenhouse-grown plants. In *slmes1* knockout lines, methyl salicylate levels were significantly higher than the Brandywine control (Figure 7). The increased methyl salicylate phenotype was consistent across experiments, although environment caused substantial variation in total methyl salicylate. A greenhouse experiment showed that *slmes1-1* and *slmes1-8* plants contained 17.4× and 6.8× more methyl salicylate than wild-type, respectively, a significant increase over the control (Figure 7B). There was substantial pest and disease pressure during the greenhouse experiment, which may explain the higher absolute levels of methyl salicylate in all genotypes compared to the field experiments.

Effects from loss of additional *SlMES* genes

The combination of *slmes1* single mutants with *slmes3* and *slmes4* double and triple mutants in the Moneymaker background allowed us to evaluate whether *SlMES1* contributed the most to methyl salicylate content. In addition to *slmes1* single knockouts, we tested *slmes1* knockouts in combination with the loss of other *SlMES* genes. The *slmes1 slmes2* double mutant was in the Brandywine Sudduth background and the *slmes1 slmes3 slmes4* triple mutants were in the Moneymaker background. The amount of methyl salicylate made by the *slmes1 slmes2* double mutant was statistically indistinguishable from the *slmes1-8* mutant, indicating that *SlMES2* is not an important contributor to the

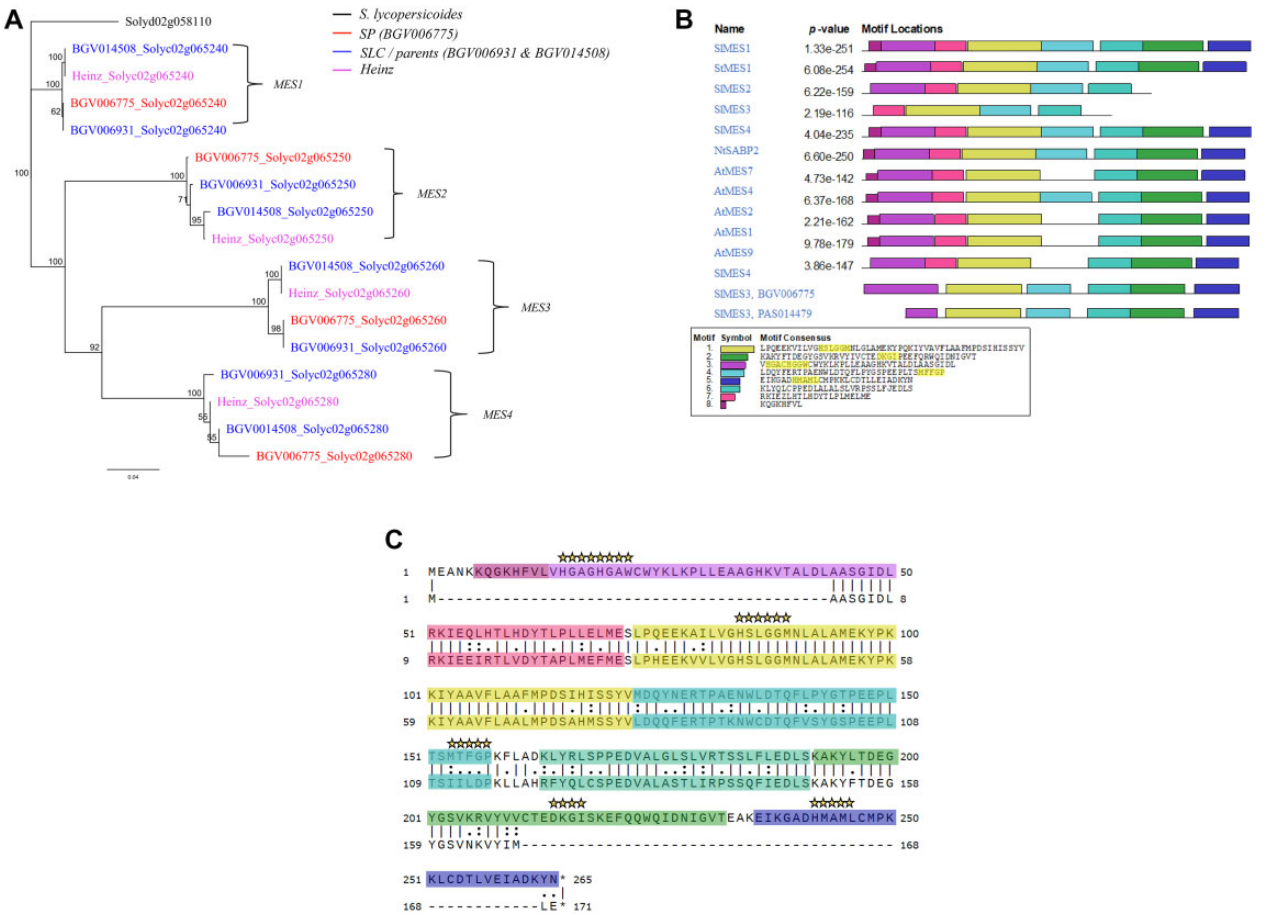

Figure 5 Conservation between MES locus genes. A, Phylogeny of the MES locus genes. MES1 from *S. lycopersicoides* (*Solyd02g058110*) was selected as the outgroup. *Solyd02g058110* is *methylesterase* in *S. lycopersicoides*. Scale bar indicates nucleotide substitutions per site. B, Conserved protein motifs in methyl esterases from the tomato reference genome Heinz1706 and other plant species. Motifs were numbered in descending size order. Predicted active site and pocket residues are highlighted in yellow, and show that SlMES3 in the reference variety is missing Motif 5, which is expected to decrease protein activity. C, Alignment of SlMES3 to SlMES2, using the Heinz reference sequence for both and the Needleman–Wunsch algorithm. A line connecting two residues indicates they are identical, two dots connecting residues indicates they are similar but not identical, and a single dot means the residues are not similar. Coloring corresponds with the Motifs 1–8 in (B). Predicted active site and pocket residues are indicated with stars.

demethylation of methyl salicylate (Figure 7C). This conclusion agreed with the expression data and locus organization, which predicted a premature stop codon in *SlMES2* and low expression in most genotypes (Figures 5, B and 6). The triple *slmes1 slmes3 slmes4* frameshift mutant contained more methyl salicylate than the Moneymaker control. However, the double and triple frameshift mutants were not significantly different from the *slmes1* single mutant (Figure 7C). These results support the notion that *SlMES1* is the most important methyl esterase to convert methyl salicylate to salicylic acid in tomato fruit.

SlMES1 and SlMES3 are methyl esterases acting on methyl salicylate

The cDNAs of *SlMES1* and *SlMES3* were cloned from ripe Heinz 1706 fruit and heterologously expressed in *Escherichia coli*. The assay quantified the loss of the volatile substrate, assuming that the substrate was converted to the nonvolatile salicylic acid, as previously described for other methyl salicylate methyl esterases (Forouhar et al., 2005). Reactions containing equal amounts of SlMES1 and SlMES3 proteins showed that SlMES1 was more active than SlMES3, as SlMES1 was capable of completely demethylating 300-µM methyl salicylate after 30 min, while SlMES3 had about 20% of the starting concentration of methyl salicylate left after 30 min of reaction (Figure 8A). The *SlMES3* allele of Heinz1706 encoded a truncated version of the protein and lacked two critical motifs. This finding was consistent with reduced activity of SlMES3 compared to SlMES1 (Figures 5, B and 8, B). To confirm that the mechanism of action by the tomato methyl esterases is similar to previously characterized methyl esterases in other plant species, we evaluated whether SlMES1 activity is inhibited by the addition of salicylic acid. Addition of salicylic acid competitively inhibited the activity of SlMES1 (Figure 8B), consistent with previous studies showing that salicylic acid binds to the active site

pocket and competes with methyl salicylate for the active site (Forouhar et al., 2005; Manosalva et al., 2010). The specific activity of SlMES1 and SlMES3 with methyl salicylate was also calculated. SlMES1 has a K_M of 37.0 ± 6.2 μM for methyl salicylate, comparable to other methyl esterases converting methyl salicylate into salicylic acid, and SlMES3 has a K_M of 237 ± 6.8 μM, consistent with a slower conversion of methyl salicylate to salicylic acid by SlMES3 (Figure 8, A and C). StMES1 from potato has a K_M of 57.9 μM, tobacco SABP2 is 8.6 μM (Forouhar et al., 2005), and two poplar methyl esterases with K_M of 68 μM and 24 μM have been reported (Zhao et al., 2009). Thus, SlMES1 and SlMES3 have K_M values similar to previously characterized plant methyl esterases.

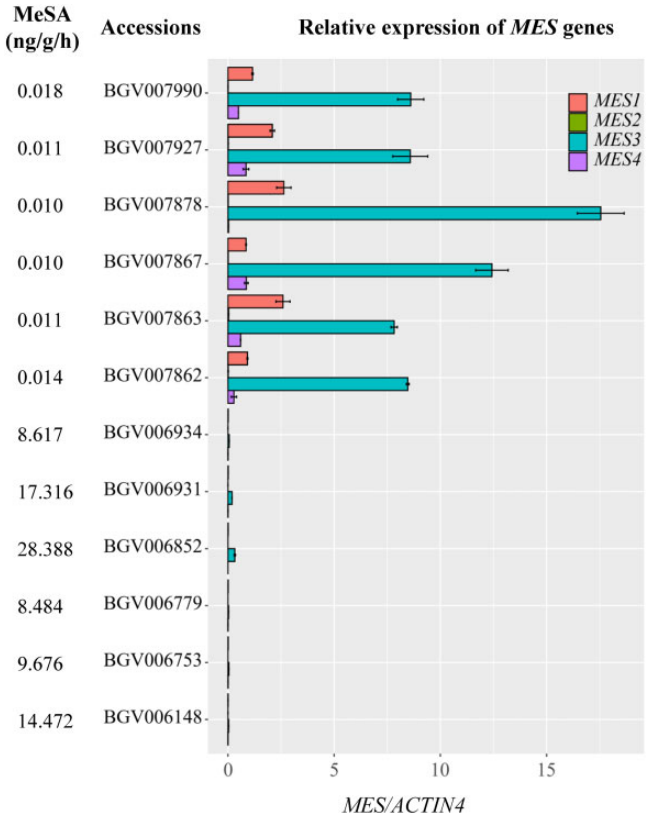

Figure 6 Relative expression of methyl esterase (*MES*) genes in different tomato accessions in ripe tomato and their corresponding methyl salicylate levels are also presented besides the accessions name. The methyl salicylate levels are in nanogram per gram fruit weight per hour). The *MES* genes expression is normalized using the expression of housekeeping gene, *ACT4 (Solyc04g011500)*. Error bars show the standard error of two technical replicates.

Discussion

Here, we describe tomato methyl esterases responsible for conversion of methyl salicylate into salicylic acid. The in vitro activity of SlMES1 on its in planta substrate, methyl salicylate, was confirmed, and salicylic acid was confirmed as the product through SlMES1 activity inhibition. SlMES1 has a K_M similar to methyl salicylate methyl esterases from other species (Forouhar et al., 2005; Zhao et al., 2009; Manosalva et al., 2010).

The *SlMES* locus is a major contributor to methyl salicylate levels in tomato fruit and its characterization is an important step toward a more complete understanding of methyl salicylate homeostasis in fruits. Methyl salicylate and salicylic acid can be derived from both phenylalanine and isochorismate in plants, with the preferred pathway varying by species (reviewed in Huang et al., 2020). In synthesis from phenylalanine, phenylalanine ammonia lyase (PAL) converts phenylalanine into (*E*)-cinnamic acid (Rigano et al., 2016). Several tomato *PALs* are expressed in ripe fruit (Martina et al., 2021). (*E*)-cinnamic acid is converted into salicylic acid by benzoic acid 2-hydroxylase (BA2H), which has not been identified in tomato fruit (Lee et al., 1995). SlSAMT converts salicylic acid to methyl salicylate (Tieman et al., 2010). With the addition of the *SlMES* locus, interconversion between salicylic acid and methyl salicylate provides another important layer of regulation. Methyl salicylate emission is regulated by multiple glycosylation steps into nonvolatile di- and tri-glycoside methyl salicylate and deglycosylation back into a volatile form. The final step into the nonvolatile glycosylated form is a primary control point mediated by NSGT1 (Tikunov et al., 2010, 2013). The first glycosylation can be catalyzed by the glycosyltransferase *SlUGT5*, which acts on many other substrates in addition to methyl salicylate (Louveau et al., 2011). It is likely that other fruit-specific glycosyltransferases with activity on methyl salicylate also exist. The enzymes catalyzing the deglycosylation of methyl salicylate also remain unknown. Studies in petunia and mango showed that β-glucosidases convert diglycoside methyl salicylate into methyl salicylate and facilitate volatile release (Cna'ani et al., 2017; Sudheeran et al., 2020). It is expected that a similar method exists in tomato fruits for release of

Table 2 Mutant lines generated by CRISPR-induced deletions

Mutant Name	Brandywine			Moneymaker		
	slmes1-1	*slmes1-8*	*slmes1-7 slmes2-4*	*slmes1-5 slmes4-3*	*slmes1-5 slmes3-1 slmes4-3*	*slmes1-5 slmes3-1 slmes4-5*
SlMES1	−1 bp	−8 bp	−7 bp	−5 bp	−5 bp	−5 bp
SlMES2	Wild-type	Wild-type	−4 bp	Wild-type	Wild-type	Wild-type
SlMES3	Wild-type	Wild-type	Wild-type	Wild-type	−1 bp	−1 bp
SlMES4	Wild-type	Wild-type	Wild-type	−3 bp	−3 bp	−5 bp

Alleles are designated by the number of base pairs deleted in each line. All alleles are the result of a minimum of one independent transformation event.

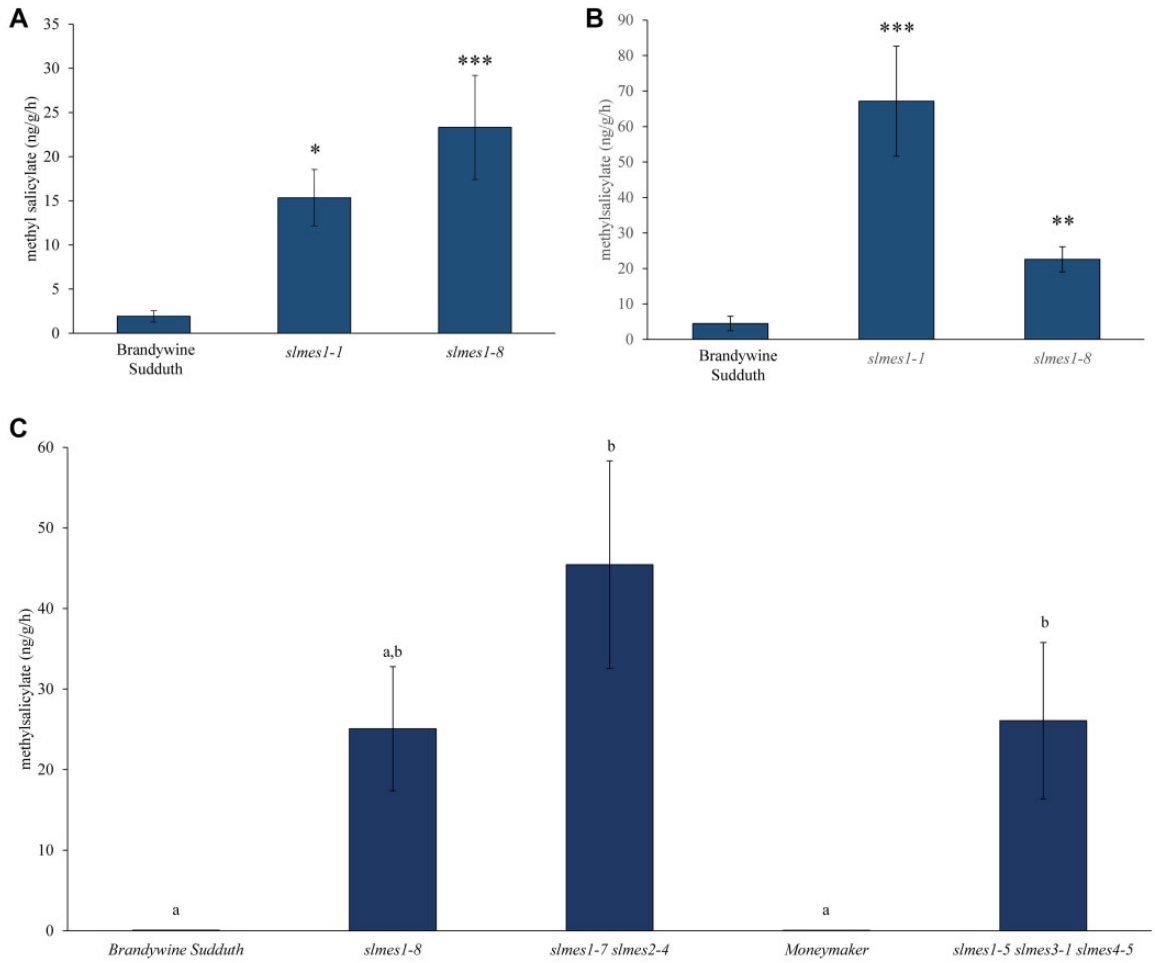

Figure 7 Methyl salicylate levels in single, double, and triple *slmes* mutants. Methyl salicylate was collected from chopped ripe fruit for an hour and measured on an Agilent GC–MS, reported as nanograms of volatile per gram of ripe tomato fruit. Results are from three individual mutation lines with two to three reps per line and three to five plants per rep. A, Levels of methyl salicylate in ripe fruit from Brandywine (wild-type) and the single *slmes1* loss-of-function mutants grown in field conditions. B, Levels of methyl salicylate in ripe fruit from wild-type and single *slmes1* loss-of-function mutants grown in greenhouse conditions. For all graphs, numbers above each bar are the mean volatile content, and error bars represent the standard error. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ $^{****}P < 0.0001$ by Student's *t* test to the wild-type. C, Methyl salicylate levels in the fruit of Moneymaker and Brandywine and in double and triple *slmes* loss of function mutants. The letter above the standard error bars show significance using Tukey's significance test ($P < 0.05$).

methyl salicylate, but as yet no β-glucosidases have been identified with this function. Figure 1 summarizes the current understanding of the synthesis and catabolism pathways described above. The SlMES enzymes, like the glycosyltransferases, reduce the pool of methyl salicylate and provide another level of control. This level of control is relevant, as demonstrated by the natural variation associated with this locus among semi-domesticated tomato accessions. Together, *SlMES* and *NSGT* loci accounted for 44% of the variability in methyl salicylate levels in a biparental cross. We expect that identifying the genes responsible for other steps, such as additional glycosyltransferases and β-glucosidases, will explain more of the observed variation in methyl salicylate levels. Better understanding the regulation of methyl salicylate in tomato fruit provides additional tools for breeders to regulate the balance between defense and flavor compounds in the tomato fruit. Additionally, the amount of natural variation and impact of sequence changes on methyl salicylate levels indicates that the *SlMES* QTL is a good candidate for breeding, as long as the *NSGT* locus in parents is known and accounted for.

The *SlMES* locus is another example of how structural variation can have a large effect on agronomic traits. Gene expression is strongly impacted by structural variation (Alonge et al., 2020). The structural variants in the *SlMES* locus are also associated with changes to the overall gene expression as well as relative expression of these family members (Figures 3, D and 6). The most obvious change to the *SlMES* coding sequences is due to transposable element insertion in the *SlMES* locus, which separated the last two motifs of *SlMES3*, annotated in the reference variety as a separate gene, *Solyc02g065270*. The separated motifs in *Solyc02g065270* contain catalytic sites, which probably cause a decrease in functionality between full-length *SlMES3* and the truncated gene.

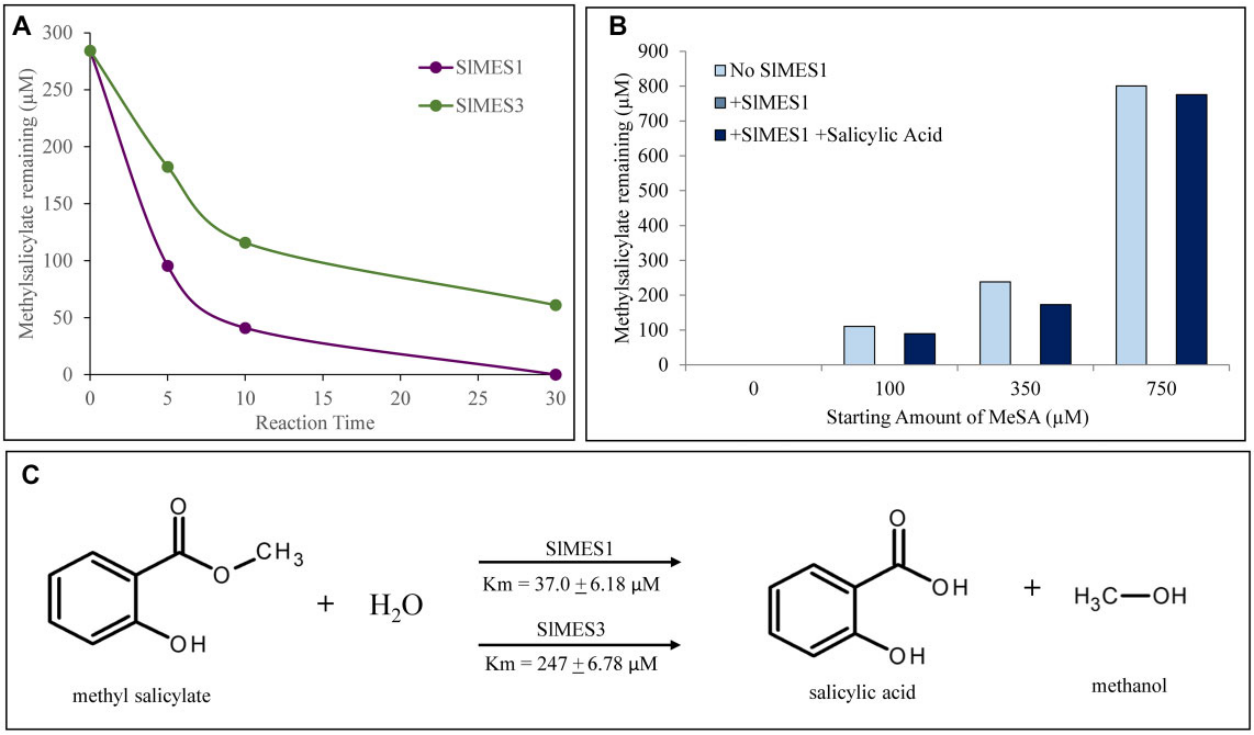

Figure 8 SlMES1 is a methyl esterase acting on methyl salicylate and methyl benzoate. A, Purified His-tagged SlMES1 and SlMES3 were incubated with methyl salicylate for up to 30 min before the reaction was stopped with HCl and volatile starting material remaining quantified on a mass spectrometer. Methyl salicylate was undetectable after 30 min incubation with SlMES1 and decreased by ~80% after incubation with SlMES3 for the same length of time. B, Salicylic acid inhibits the conversion of methyl salicylate to salicylic acid by SlMES1. When salicylic acid is added to the reaction mix, more methyl salicylate remains at the end of the reaction. C, Predicted reaction and K_m's of SlMES1 and SlMES3 based on action on methyl salicylate, and observations from other species that methyl salicylate is converted to salicylic acid and methanol by the action of similar methyl esterases. Results for SlMES1 are from four replicates; results for SlMES3 are from three replicates.

A less-active SlMES3 is supported by both the relative activity on methyl salicylate compared to SlMES1 (Figure 8A), and the smaller effect of *SlMES3* knockout on fruit methyl salicylate content than *SlMES1* knockout (Figure 7C). Interestingly, *SlMES3* is the second recent example of an intragenic transposable element altering the functionality of a flavor gene. Recent work identified a 2-phenylethanol gene, *PPEAT*, that contains a large intronic Copia retrotransposon element insertion. Varieties with the insertion made more 2-phenylethanol than those lacking the insertion (Domínguez et al., 2020). The Copia insertion was found primarily in the wild SP accessions and semi-domesticated varieties, but rarely in modern varieties, which also make overall less flavor chemicals (Domínguez et al., 2020). Transposable elements within the *SlMES* locus have also caused other changes, including increasing the total locus size in SLL compared to SP (Supplemental Table S3). Multiple predicted transposable elements are also inserted near the promoters of these genes, although the effects of those insertions are much less predictable than the separation of *Solyc02g065270* from *SlMES3* (Supplemental Table S4).

In the cultivars used in this study, Brandywine Sudduth and Moneymaker, we found that *SlMES1* is a contributor to the methyl salicylate phenotype (Figure 7). In addition to the knockout plants, in vitro work shows that SlMES3 also demethylates methyl salicylate, although the affinity for the substrate is substantially lower than that of SlMES1 (Figure 8A). Generating and analyzing additional single gene knockouts in *slmes3* and *slmes4* would determine the relative importance of each gene and definitively conclude whether all the *SlMES* genes contribute to methyl salicylate conversion in ripe fruit, but will require making new sets of transgenic lines targeting single *SlMES* genes only, rather than the combinatorial approach in this work with triple *slmes1 slmes3 slmes4* mutants used to characterize the importance of the entire locus. Results with the single *slmes1* knockouts show that the plant is unable to compensate for the loss of a single *SlMES* with other functional *SlMES* genes, again demonstrating that *SlMES1* is a key player in regulating the amount of methyl salicylate in ripe fruit. The expression pattern of the *SlMES* genes also supports *SlMES1* as the primary fruit methyl esterase. *SlMES1* is most highly expressed in the immature fruit and ripe fruit, and *SlMES3* at ripening (Table 3). *SlMES4* is expressed in the largest range of tissues, with substantial expression in leaves, buds, and ripe fruit (Table 3). Together, these observations support the conclusion that the primary activity of *SlMES1* and *SlMES3* is during fruit development

Table 3 RNA expression levels of *SlMES1-4* in selected tissues from the reference tomato Heinz and Brandywine, one of the varieties used in this study

Tissue Type	Variety	SlMES1	SlMES2	SlMES3	SlMES4
Leaf	Heinz	18.305	2.135	0.515	59.21
	Brandywine	7.087	0.323	0.077	50.553
Bud	Heinz	18.375	2.185	0.17	76.92
Flower	Heinz	5.025	0.595	0	22.075
Root	Heinz	0.18	0	0.14	12.90
1-cm fruit	Heinz	112.88	23.66	1.91	14.9
3-cm fruit	Heinz	89.465	7.025	1.175	4.935
Mature Green Fruit	Heinz	9.1	0.865	0.16	0.81
	Brandywine	21.160	0	0.740	0.977
Breaker Fruit	Heinz	27.72	3.81	31.11	2.185
	Brandywine	107.53	0.623	111.84	7.240
Breaker Fruit + 7 days	Brandywine	121.633	0.830	408.427	47.780
Breaker Fruit + 10 days	Heinz	102.17	30.2	512.48	108.11

Expression values are from RNAseq of Heinz from The Tomato Gene Consortium (2012).

and not as major contributors to salicylic acid metabolism in vegetative tissues. In support of a limited role for the *SlMES* genes in volatile flavor production in ripening fruit, we observed no changes in plant architecture, ripening, pathogen resistance, nor fruit size or shape in the *slmes* knockout lines.

The *SlMES* locus was not identified through earlier GWAS analyses for tomato flavor (Tieman et al., 2017), probably because the previous study was composed of mainly heirloom and commercial SLL varieties, while this study analyzed mainly ancestral tomato varieties from Latin America. Discovery of a previously unidentified major locus controlling methyl salicylate accumulation suggests that there is a great deal of unexplored genetic diversity in available germplasm, particularly in the more genetically diverse SP and *S. lycopersicum* var. *cerasiforme* populations and highlights the importance of screening the broadest possible genetic pool.

The results presented here are a major step toward a more complete understanding of methyl salicylate synthesis in ripening tomato fruit. Loss of *SlMES1*, *SlMES3*, and *SlMES4* results in significant increases in methyl salicylate in ripe fruit, and SlMES1 and SlMES3 efficiently demethylate methyl salicylate. Selection of haplotypes with the highest expression of *MES* genes will provide breeders with new options for decreasing methyl salicylate flavor in ripe tomato fruit. In particular, selection of the most desirable haplotypes of the MES locus combined with the best haplotypes at the *NSGT* locus should result in fruits with greatly reduced methyl salicylate and improved flavor.

Materials and methods

Plant material

Tomato (*S. lycopersicum*) seeds were obtained from the Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana (Banco de Germoplasma Vegetal [BGV] lines), the Tomato Genetics Resource Center (Davis, CA, USA), or commercial suppliers (Supplemental Table S1). Plants in greenhouse experiments were grown in a heated greenhouse on the University of Florida campus according to recommended commercial practices. Plants in the field experiments were grown at the University of Florida's North Florida Research and Education Center-Suwannee Valley in Live Oak, FL. All fruit was harvested at the red-ripe stage.

Volatile collection

Tomato fruit volatiles were collected as described in Tieman et al. (2006). Briefly, whole fruits were chopped and loaded into glass tubes, then air was passed over the samples, and volatiles were collected on a SuperQ Resin for 1 h. Volatiles were eluted off the column with methylene chloride and run on a GC for analysis as described in Tieman et al. (2006). Replicates with fewer than two measurements in a single growing season were eliminated from the analysis.

Generation of slmes loss-of-function mutants

Brandywine Sudduth tomato was transformed with a construct containing two guide RNAs targeting *SlMES1*, *SlMES2*, and *SlMES3*. Moneymaker was transformed with a construct containing three guide RNAs targeting *SlMES1*, *SlMES3*, and *SlMES4*. A guide RNA was designed to target *SlMES1*, *SlMES2*, and *SlMES3*. This guide RNA had no mismatches to the *SlMES1* sequence, one mismatch to *SlMES3*, and two mismatches to *SlMES2*. An in vitro assay validated cleavage on three genes of interest (Takara, San Jose, CA, USA), and a BLAST search showed no other targets with fewer than five mismatches. The guide RNA was cloned into the kanamycin-resistant p201N vector (Jacobs et al., 2015). Transformation was performed at both the University of Nebraska's Plant Transformation Center or the University of Florida in Brandywine Sudduth cultivar and yielded 16 positive transformants. Of the 16 plants, 9 contained deletions in *SlMES1*, and 6 plants contained a deletion in *SlMES2*. No plants with deletions in *SlMES3* were recovered. Three different mutations were obtained in the *slmes1* knockout plants: a 1-bp deletion (*slmes1-1*), a 7-bp deletion (*slmes1-7*), and an 8-bp deletion (*slmes1-8*). Two mutations were obtained in the *slmes2* knockouts, a 3-bp deletion, named *slmes2-3*, and a 4-bp deletion, named *slmes2-4* (Table 2). Transformations in the Moneymaker background were done at University of Georgia using another set of guide RNAs targeting *SlMES1*, *SlMES3*, and *SlMES4* (Supplemental Table S2). No single mutants were generated in this transformation. However, one T1 carried a 5-bp frameshift mutation in only *MES1* and a 3-bp in-frame mutation in *MES4*. Two T1 families carried an out of frame deletions in *SlMES1* (5 bp) and *SlMES3* (1 bp) and were heterozygous for a 3- or 5-bp mutation in *SlMES4* (Table 2). In the T2 generation, mutants with the 5-bp deletion in *MES4* were selected. Primary transformants were allowed to self-fertilize and backcrossed to Brandywine Sudduth or Moneymaker. The backcrossed lines were screened to select mutants without the Cas9 T-DNA. Plants for analysis were selected from these self-pollinated and backcrossed lines.

Expression of SlMES1 and SlMES3 in *E. coli*

SlMES1 and SlMES3 ORFs were cloned from RNA extracted from ripe Heinz 1706 (reference) fruit. The ORF was then cloned into the 6X-His tag vector pET28a (Invitrogen, Waltham, MA, USA). SlMES1 was expressed in BL21(DE3) cells, and SlMES3 was expressed in Rosetta (DE3) cells (NEB). Cells were grown at 37°C until they reached mid-exponential growth phase, then induced with 100-μM IPTG (Sigma, St Louis, MO, USA), and moved to room temperature overnight. His-SlMES1 and His-SlMES3 were purified with a nickel-resin column (Qiagen Hilden, Germany). Purified protein was quantified by Nanospectrometer, and purity evaluated on a Bis–Tris gel.

SlMES1 substrate determination

Purified SlMES1 and SlMES3 (2.4 mg mL^{-1} $reaction^{-1}$) were incubated with 300-μM methyl salicylate (Sigma) for 0–30 min in a 100-μL reaction containing 50-mM Tris–HCl, pH 6.0, 10% (v/v) glycerol, 5-mM DTT, then each reaction stopped with a 10× reaction volume of 1-M HCl (Fisher, Waltham, MA, USA). Methyl salicylate was quantified on an Agilent GC–MS with a Polydimethylsiloxane/Divinylbenzene SPME fiber and DB-5 column. Peaks corresponding to methyl salicylate were quantified. Disappearance of the substrate compared to the negative control, containing no protein, indicates successful hydrolysis by each SlMES.

SlMES1 and SlMES3 kinetic parameter determination

Purified SlMES1 was diluted to 3 μg mL^{-1} per reaction and incubated for 0, 5, 10, and 15 min with 250–800 μM methyl salicylate before the reaction was stopped with 1-M HCl. The amount of substrate remaining was quantified on an Agilent GC–MS and used for determination of V_{max} and K_m by fitting to the Michaelis Menten function in JMP. SlMES3 kinetic parameters were determined in a similar manner but with a 100–500 μM range of substrate concentrations.

SlMES1 reaction product determination

Purified SlMES1 was incubated with methyl salicylate ranging from 0 to 750 μM, along with 0–10 μM salicylic acid (Sigma). Reactions were run for 15 or 30 min, stopped with 1-M HCl (Fisher), then run on an Agilent GC–MS as described above. Methyl salicylate was quantified by peak area as described above.

GWAS analysis

The raw ILLUMINA read files of 166 tomato accessions from the Varitome collection (Razifard et al., 2020) were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/, SRA: SRP150040, BioProject: PRJNA454805). Raw read evaluation, filtering, alignment, and variant calling for SNPs were performed as described previously using the SL4.0 genome build (Pereira et al., 2021). For the association analysis, the genome-wide kinship matrix was calculated using Centered IBS method in TASSEL 5.2.44 (Bradbury et al., 2007). Hapmap files were generated from the extracted SNPs using TASSEL. Associations between the genotype and methyl salicylate levels were calculated using FarmCPU model in GAPIT (version 3) (Wang and Zhang, 2021). Minor allele frequency and SNP FDR value was set to 2% and 5%, respectively, for the analysis. The significance thresholds for the association were set at a significant level of 5%, respectively, after FDR multiple test correction. The association results were plotted using ggplot2 package in R (Wickham, 2011).

Linkage mapping in the biparental population

The F_2 population were genotyped for SNPs using the Kompetitive allele-specific PCR (KASP) genotyping assay (Semagn et al., 2014). The primer sequences and their positions in the genome are shown in Supplemental Table S5. Methyl salicylate accumulation was associated with the markers by QTL mapping using "R/qtl" package (Broman et al., 2003) in R version 4.0.2 (Team R Core, 2013) with RStudio version 1.3.1056 interface (Studio, 2012). QTLs were detected via composite interval mapping analysis using the Kosambi mapping function and Extended Haley Knott regression. LOD thresholds were calculated from 1,000 permutations to determine the significance level. QTL with LOD > 2.91 was considered significant. The "fitqtl" function was used to estimate the effect of identified QTL. QTL maps were plotted using "ggplot2" package in R (Wickham, 2011).

Phylogenetic tree construction

Phylogenetic tree was built using Neighbor-Joining method using the cDNA sequences of all the genes. The cDNA sequences were extracted using in house pipeline. The sequences were exported in Geneious Prime version 2021.2.2 (https://www.geneious.com/) and aligned using Clustal alignment option (Jeanmougin et al., 1998). Phylogenetic tree was then constructed using Neighbor-Joining method and Tamura-Nei genetic distance model. The *methylesterase* gene from *Solanum lycopersicoides* (*Solyd02g058110*) was used as an outgroup.

Annotation of conserved protein motifs

The online tool MEME (Bailey and Elkan, 1994; Bailey et al., 2009) was used to identify conserved protein motifs likely involved in protein function. Methyl esterase sequences in other *Solanaceae* species (potato, pepper, eggplant, and *N. benthamiana*) and *A. thaliana* were obtained using Blast. Predictions of catalytic domains and pocket residues were done with Phyre2.0 (Kelley and Sternberg, 2009).

SlMES3/4 transcript from SP accessions

True leaves from BGV006775, PAS014479, Brandywine Sudduth, and Heinz 1706 were frozen and ground in liquid nitrogen before extraction via an RNAeasy kit (Qiagen) and conversion to cDNA (LunaScript RT, NEB). We amplified transcript using each combination of four primers: one forward primer starting at the annotated *SlMES3* start, one forward primer starting at the alternate upstream *SlMES3* start

codon, one reverse primer at the annotated end of *SlMES3*, and a reverse primer at the end of *SlMES4* (Supplemental). All resulting DNA bands were excised and sequenced (Genewiz) to determine transcribed gene regions. Bands with strong signal were cloned into pENTR.

RT–qPCR

Gene-specific primers for the quantitative expression analyses were designed using NCBI Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi). Only one SNP toward the 5′-end of the primer was permitted (Supplemental Table S6). The primer sets were separated by one intron in the corresponding genomic DNA. *ACTIN4* (*Solyc04g011500*) was used as the control to standardize relative expression levels (Supplemental Table S6). RNA from ripe fruits was extracted by using ISOLATE II RNA Plant Kit (Bioline, BIO-52077) following the specifications by the manufacturer. The mRNA was transcribed into cDNA by reverse transcriptase (Invitrogen SuperScript III First-Strand Synthesis System for RT–qPCR, 18080-051). Quantitative PCR was setup using the SsoAdvanced Universal SYBR Green Supermix kit (BIO-RAD, 1725271) with two technical replicates. The amplification was conducted on a CFX Real-Time PCR Instrument (BIO-RAD, Hercules, CA, USA; CFX96 real time system). The cycling protocol for the quantitative PCR was: 95°C for 3 min, followed by 40 cycles of 95°C for 10 s, 60°C for 30 s. The plate was read using the melt curve 65°C to 95°C with 0.5°C increment for 5 s/step. The relative expression was calculated using the formula of ΔCq method using a reference gene: $R = 2^{Cq(reference)-Cq(target)}$, where Cq is quantification cycle, reference is *ACT4*, and targets are *MES1*, *MES2*, *MES3*, or *MES4*. The expression analysis figure was generated using "ggplot2" package in R (Wickham, 2011).

RNA-seq analysis

RNA expression of Brandywine fruit was performed as described in Shi et al. (2021).

Accession numbers

Sequence data from this article can be found in the GenBank data library under the following NCBI Gene IDs: *SlMES1*-101261578, *SlMES2*-101261293, *SlMES3*-112940921, and *SlMES4*-101260990.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. Alignment of two parental accessions against the SL4.0 build at the *MES* locus.

Supplemental Table S1. List and brief description of major varieties used within this study.

Supplemental Table S2. Deletions in the *SlMES* genes induced by gene editing.

Supplemental Table S3. Difference in length between cultivated and wild tomato.

Supplemental Table S4. Transposable elements inserted <1 kb from *SlMES1* contain promoter elements.

Supplemental Table S5. List of KASP markers used for genotyping the F2 population.

Supplemental Table S6. RT-qPCR primers for the expression analysis of SlMES1-4 and actin.

Acknowledgments

We thank Ben Broughton and his team at University of Florida's Suwannee County Extension field site for their support in field preparation and plant care; Leila Wolf, C. Brooke Mclaren, Amber Newsome, and the undergraduates in the Klee lab for their help with genotyping, harvest, and fruit processing.

Funding

This work was supported by funding from U. S. Department of Agriculture NIFA AFRI (Award No. 2019-67012-29533 to E.M.F.), the National Science Foundation IOS 1564366 (E.V.K. and D.T.) and IOS 1855585 (H.K., D.T., and J.G.).

Conflict of interest statement. None declared.

References

Agricultural Marketing Resource Center (2018) https://www.agmrc.org/commodities-products/vegetables/tomatoes

Alonge M, Wang X, Benoit, M, Soyk S, Pereira L, Zhang L, Suresh H, Ramakrishnan S, Maumus F, Ciren D, et al. (2020) Major impacts of widespread structural variation on gene expression and crop improvement in tomato. Cell **182**: 145–161

Bailey T, Bodén M, Buske F, Frith M, Grant C, Clementi L, Ren J, Li W, Noble W (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res **37**: W202–W208

Bailey T, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, CA, pp 28–36

Bradbury P, Zhang Z, Kroon D, Casstevens T, Ramdoss Y, Buckler E (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics **23**: 2633–2635

Broman K, Wu H, Sen S, Churchill G (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics **19**: 889–890

Bruhn C, Feldman N, Garlitz C, Harwood J, Ivans E, Marshall M, Riley A, Thurber D, Williamson E (1991) Consumer perceptions of quality: apricots, cantaloupes, peaches, pears, strawberries, and tomatoes. J Food Qual **14**: 187–195

Cao X, Duan W, Wei C, Chen K, Grierson D, Zhang B (2019) Genome-wide identification and functional analysis of carboxylesterase and methylesterase gene families in peach (*Prunus persica* L. Batsch). Front Plant Sci **10**: 1511

Chen F, D'Auria J, Tholl D, Ross J, Gershenzon J, Noel J, Pichersky E (2003) An *Arabidopsis thaliana* gene for methylsalicylate biosynthesis, identified by a biochemical genomics approach, has a role in defense: benzoic/salicylic acid methyltransferase in Arabidopsis. Plant J **36**: 577–588

Cna'ani A, Reut R, Ravid J, Aravena-Calvo J, Skaliter O, Masci T, Vainstein A (2017) Phenylpropanoid scent compounds in petunia x hybrida are glycosylated and accumulate in vacuoles. Front Plant Sci **8**: 1898

Deng W, Wang R, Yang T, Jiang L, Zhang Z (2017) Functional characterization of salicylic acid carboxyl methyltransferase from

Camellia sinensis, providing the aroma compound of methyl salicylate during the withering process of white tea. J Agric Food Chem **65**: 11036–11045

Domínguez M, Dugas E, Benchouaia M, Leduque B, J Jiménez-Gómez J, Colot V, Quadrana L (2020) The impact of transposable elements on tomato diversity. Nat Commun **11**: 4058

Forouhar F, Yang Y, Kumar D, Chen Y, Fridman E, Park S, Chiang Y, Acton T, Montelione G, Pichersky E, et al. (2005) Structural and biochemical studies identify tobacco SABP2 as a methyl salicylate esterase and implicate it in plant innate immunity. Proc Natl Acad Sci USA **102**: 1773–1778

Fushan A, Simons C, Slack J, Manichaikul A, Drayna D (2009) Allelic polymorphism within the TAS1R3 promoter is associated with human taste sensitivity to sucrose. Curr Biol **19**: 1288–1293

Huang W, Wang Y, Li X, Zhang Y (2020) Biosynthesis and regulation of salicylic acid and N-hydroxypipecolic acid in plant immunity. Mol Plant **13**: 31–41

Jacobs T, LaFayette P, Schmitz R, Parrott W (2015) Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnol **15**: 16

Jeanmougin F, Thompson J, Gouy M, Higgins D, Gibson T (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci **23**: 403–405

Kelley L., Sternberg M (2009) Protein structure prediction on the web: a case study using the phyre server. Nat Protocol **4**: 363–371

Koo Y, Kim M, Kim E, Song J, Jung C, Moon J, Kim J, Seo H, Song S, Kim J, et al. (2007) Overexpression of salicylic acid carboxyl methyltransferase reduces salicylic acid-mediated pathogen resistance in Arabidopsis thaliana. Plant Mol Biol **64**: 1–15

Krumbein A, Auerswald H (1998) Characterization of aroma volatiles in tomatoes by sensory analyses. Food/Nahrung **42**: 395–399

Lee H, León J, Raskin I (1995) Biosynthesis and metabolism of salicylic acid. Proc Natl Acad Sci USA **92**: 4076–4079

Lin J, Mazarei M, Zhao N, Zhu J, Zhuang X, Liu W, Pantalone V, Arelli P, Stewart C, Chen F (2013) Overexpression of a soybean salicylic acid methyltransferase gene confers resistance to soybean cyst nematode. Plant Biotechnol J **11**: 1135–1145

Louveau T, Leitao C, Green S, Hamiaux C, van der Rest B, Dechy-Cabaret O, Atkinson R, Chervin C (2011) Predicting the substrate specificity of a glycosyltransferase implicated in the production of phenolic volatiles in tomato fruit. FEBS J **278**: 390–400

MacLeod A, Gonzales de Troconis N (1982) Volatile flavor components of sapodilla fruit (Achras sapota L). J Agric Food Chem **30**: 515–517

Mageroy M, Tieman D, Floystad A, Taylor M, Klee H (2012) A *Solanum lycopersicum* catechol-O-methyltransferase involved in synthesis of the flavor molecule guaiacol. Plant J **69**: 1043–1051

Manosalva P, Park S, Forouhar F, Tong L, Fry W, Klessig D (2010) Methyl esterase 1 (StMES1) is required for systemic acquired resistance in potato. Mol Plant-Microbe Interact **23**: 1151–1163

Martina, M, Tikunov Y, Portis E, Bovy A (2021) The genetic basis of tomato aroma. Genes **12**: 226

Métraux J (2002) Recent breakthroughs in the study of salicylic acid biosynthesis. Trends Plant Sci **7**: 332–334

Min D, Fujun L, Zhang X, Shu P, Cui X, Dong L, Ren C, Meng D, Li J (2018) Effect of methyl salicylate in combination with 1-methylcyclopropene on postharvest quality and decay caused by *Botrytis cinerea* in tomato fruit. J Sci Food Agric **98**: 3815–3822

Negre F, Kolosova N, Knoll J, Kish C, Dudareva N (2002) Novel S-adenosyl-L-methionine: salicylic acid carboxyl methyltransferase, an enzyme responsible for biosynthesis of methyl salicylate and methyl benzoate, is not involved in floral scent production in snapdragon flowers. Arch Biochem Biophys **406**: 261–270

Park S, Kaimoyo E, Kumar D, Mosher S, Klessig D (2007) Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. Science **318**: 113–116

Pereira L, Sapkota M, Alonge M, Zheng Y, Zhang Y, Razifard H, Taitano N, Schatz M, Fernie A, et al. (2021) Natural genetic diversity in tomato flavor genes. Front Plant Sci **12**: 914

Razifard H, Ramos A, Della Valle A, Bodary C, Goetz E, Manser E, Li X, Zhang L, Visa SD, van der Knaap E, Caicedo A (2020) Genomic evidence for complex domestication history of the cultivated tomato in Latin America. Mol Biol Evol **37**: 1118–1132

Rigano M, Raiola A, Docimo T, Ruggieri V, Calafiore R, Vitaglione P, Ferracane R, Frusciante L, Barone A (2016) Metabolic and molecular changes of the phenylpropanoid pathway in tomato (*Solanum lycopersicum*) lines carrying different *Solanum pennellii* wild chromosomal regions. Front Plant Sci **7**: 1484

Ross J, Nam K, D'Auria J, Pichersky E (1999) S-adenosyl-L-methionine: salicylic acid carboxyl methyltransferase, an enzyme involved in floral scent production and plant defense, represents a new class of plant methyltransferases. Arch Biochem Biophys **367**: 9–16

Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol Breed **33**: 1–14

Shi Y, Vrebalov J, Zheng H, Xu Y, Yin X, Liu W, Liu Z, Sorensen I, Su G, Ma Q, et al. (2021) A tomato LATERAL ORGAN BOUNDARIES transcription factor, SlLOB1, predominantly regulates cell wall and softening components of ripening. Proc Natl Acad Sci USA **18**: e2102486118

Studio R (2012) RStudio: Integrated Development Environment for R. RStudio Inc, Boston, MA, p 74

Sudheeran P, Ovadia R, Galsarker O, Maoz I, Sela N, Maurer D, Feygenberg O, Shamir M, Alkan N (2020) Glycosylated flavonoids: fruit's concealed antifungal arsenal. New Phytologist **225**: 1788–1798

Team R Core (2013) R: A Language and Environment for Statistical Computing. Team R Core, Vienna, Austria

Tieman D, Taylor M, Schauer N, Fernie A, Hanson A, Klee H (2006) 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, Zeigler M, Schmelz E, Taylor M, Rushing S, Jones J, Klee H (2010) Functional analysis of a tomato salicylic acid methyl transferase and its role in synthesis of the flavor volatile methyl salicylate. Plant J **62**: 113–123

Tieman D, Bliss P, McIntyre L, Blandon-Ubeda, A, Bies D, Odabasi A, Rodríguez G, van der Knaap E, Taylor M, Goulet C, et al. (2012) The chemical interactions underlying tomato flavor preferences. Curr Biol **22**: 1035–1039

Tieman D, Zhu G, Resende Jr, M, Lin T, Nguyen C, Bies D, Rambla J, Beltran K, Taylor M, Zhang B, et al. (2017) A chemical genetic roadmap to improved tomato flavor. Science **355**: 391–394

Tikunov Y, de Vos R, Gonzalez Paramas A, Hall R, Bovy A (2010) A role for differential glycoconjugation in the emission of phenylpropanoid volatiles from tomato fruit discovered using a metabolic data fusion approach. Plant Physiol **152**: 55–70

Tikunov Y, Molthoff J, de Vos R, Beekwilder J, van Houwelingen A, van der Hooft J, Nijenhuis-de Vries M, Labrie C, Verkerke W, van de Geest H, et al. (2013) Non-smoky glycosyltransferase1 prevents the release of smoky aroma from tomato fruit. Plant Cell **25**: 3067–3078

Trimmer C, Keller A, Murphy A, Snyder L, Willer J, Nagai M, Katsanis N, Vosshall L, Matsunami H, Mainland J (2019) Genetic variation across the human olfactory receptor repertoire alters odor perception. Proc Natl Acad Sci **116**: 9475–9480

Wang J, Zhang Z (2021) GAPIT version 3: boosting power and accuracy for genomic association and prediction. Genomics Proteomics Bioinformatics **19**: 629–640

Wang H, Sun M, Li L, Xie X, Zhang Q (2015) Cloning and characterization of a benzoic acid/salicylic acid carboxyl methyltransferase gene involved in floral scent production from lily (Lilium 'Yelloween'). Genet Mol Res **14**: 14510–14521

Wang Y, Li B, Qin G, Li L, Tian S (2011) Defense response of tomato fruit at different maturity stages to salicylic acid and ethephon. Sci Hortic **129**: 183–188

Wickham H (2011) ggplot2. Wiley Interdisciplin Rev Comput Statist **3**: 180–185

Zanor M, Rambla J, Chaïb J, Steppa A, Medina A, Granell A, Fernie A, Causse M (2009) Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. J Exp Bot **60**: 2139–2154

Zhang X, Shen L, Li F, Meng D, Sheng J (2011) Methyl salicylate-induced arginine catabolism is associated with up-regulation of polyamine and nitric oxide levels and improves chilling tolerance in cherry tomato fruit. J Agric Food Chem **59**: 9351–9357

Zhang B, Tieman D, Jiao C, Xu Y, Chen K, Fei Z, Giovannoni J, Klee H (2016) Chilling-induced tomato flavor loss is associated with altered volatile synthesis and transient changes in DNA methylation. Proc Natl Acad Sci USA **113**: 12580–12585

Zhang X, Min D, Li F, Ji N, Meng D, Li L (2017) Synergistic effects of L-arginine and methyl salicylate on alleviating postharvest disease caused by Botrysis cinerea in tomato fruit. J Agric Food Chem **65**: 4890–4896

Zhao N, Guan J, Farad Forouhar F, Tschaplinski T, Cheng Z, Tong L, Chen F (2009) Two poplar methyl salicylate esterases display comparable biochemical properties but divergent expression patterns. Phytochemistry **70**: 32–39