

Role of an esterase in flavor volatile variation within the tomato clade

Charles Goulet, Melissa H. Mageroy, Nghi B. Lam, Abbye Floystad, Denise M. Tieman, and Harry J. Klee¹

Horticultural Sciences Department, University of Florida, Gainesville, FL 32611

Contributed by Harry J. Klee, September 21, 2012 (sent for review July 13, 2012)

Tomato flavor is dependent upon a complex mixture of volatiles including multiple acetate esters. Red-fruited species of the tomato clade accumulate a relatively low content of acetate esters in comparison with the green-fruited species. We show that the difference in volatile ester content between the red- and green-fruited species is associated with insertion of a retrotransposon adjacent to the most enzymatically active member of a family of esterases. This insertion causes higher expression of the esterase, resulting in the reduced levels of multiple esters that are negatively correlated with human preferences for tomato. The insertion was evolutionarily fixed in the red-fruited species, suggesting that high expression of the esterase and consequent low ester content may provide an adaptive advantage in the ancestor of the red-fruited species. These results illustrate at a molecular level how closely related species exhibit major differences in volatile production by altering a volatile-associated catabolic activity.

genome evolution | *Solanum lycopersicum* | SICXE1 | carboxylesterase

The flavor of a food involves integration of the information detected by taste and olfactory receptors (1). In the case of tomato (*Solanum lycopersicum*), flavor is the sum of the interactions between sugars, acids, and multiple volatile chemicals (2). Although sugars and acids are essential to our appreciation of tomato, the uniqueness and complexity of its flavor are dependent upon the blend of volatiles that can be detected (3, 4). Plants synthesize a vast array of volatile organic compounds. Derived from primary and secondary metabolites, they contribute to functions as diverse as defense against herbivores and pathogens, plant-to-plant interactions, and attraction of pollinators and seed dispersers (5–7). Most of the important tomato volatiles are derived from amino acids, fatty acids, and carotenoid precursors (2, 8). Despite the importance of these volatiles to fruit quality, regulation of their synthesis still remains poorly understood (4).

One approach to identification of genes involved in plant volatile production is based upon characterization of quantitative trait loci (QTL). This method exploits variation between cultivars of the same species or between closely related species (9–11). In tomato, populations of introgression lines (ILs) that contain only a fragment of the genome of a wild species have been useful in the discovery of numerous volatile-associated QTL (10, 12) as well as of QTL affecting yield, carotenoid production, and accumulation of primary metabolites (13, 14). From a human flavor perspective, it makes the most sense to concentrate on the volatiles that are most important to consumer preferences. Historically, the most important volatiles have been identified on the basis of odor units, that is, the concentration of a given volatile divided by its odor threshold (2, 15). Although this method has value, recent work has shown that the reality of taste preference is far more complex as interactions between volatiles and other flavor-associated chemicals influence perception. Consumer preference panels conducted with a large, diverse set of tomato varieties identified a set of volatiles that are positively or negatively correlated with preferences (16).

There is great diversity within the tomato clade as the different species have been shaped by the selective forces of different habitats, biotic stresses, pollinators, and seed-dispersal agents (17, 18). The clade can be divided into two subgroups based on fruit characteristics. One group consists of species characterized,

at least from a human perspective, by unpalatable green fruits. The second group, including *S. lycopersicum* and *Solanum pimpinellifolium*, produces comestible red fruits. *Solanum cheesmaniae* and *Solanum galapagense* can also be placed into this latter group as they have colored fruits and are closely genetically related to *S. lycopersicum*. *S. cheesmaniae* and *S. galapagense* bear yellow/orange fruits and are endemic to the isolated Galapagos Islands. The fruits of *S. cheesmaniae* and *S. galapagense*, although edible, have a sharp taste and are not as palatable as those of *S. lycopersicum* and *S. pimpinellifolium*. From an evolutionary point of view, the colored fruit species form a monophyletic group. The green-fruited species *Solanum chmielewskii* and *Solanum neorickii* are part of a different clade, and *Solanum pennellii* and *Solanum habrochaites* are further separated from *S. lycopersicum* (19).

Volatile esters are an important class of compounds that contribute to the characteristic aroma of many fruits and flowers. Acetate esters are, for example, key to the aroma of banana (3-methylbutyl acetate) (20), jasmine (benzyl acetate) (21), and apple cultivars (multiple esters) (22). Although the sensory attributes of each compound differ, acetate esters generally have fruity or floral-like aromas. Acetate ester volatiles are produced by alcohol acetyltransferases that condense an alcohol, also generally a volatile compound, and acetyl-CoA (23). In addition to being found in fruits and flowers, acetate esters also occur in vegetative tissues. *Cis*-3-hexenyl acetate, for example, is one of the most abundant compounds emitted from mechanically and herbivore-damaged plants, suggesting a role in plant defense (24). Moreover, this volatile was also able to prime a defense response in surrounding plants, suggesting a key role in plant-to-plant signaling (25, 26).

Here, we show that acetate esters negatively correlate with consumer liking of tomato fruits. Based on their potentially negative impact on consumer preferences as well as their potential contributions to plant defense, we were interested in understanding how the content of those compounds is regulated in fruit. To identify the most important regulators of ester accumulation, we used ILs derived from the wild relatives of tomato, *S. pennellii* and *S. habrochaites*. Using these IL populations, we identified a major QTL controlling volatile ester content.

Results

Ester Volatiles in the Tomato Clade. We previously screened a large, chemically diverse population of heirloom tomato varieties, permitting us to correlate the concentrations of flavor-associated chemicals with consumer liking (16). Correlation of consumer-liking scores with content of individual acetate esters as well as total acetate esters as a percentage of total volatiles indicated a significant negative correlation value (Table S1). This negative

Author contributions: C.G., D.M.T., and H.J.K. designed research; C.G., M.H.M., N.B.L., A.F., and D.M.T. performed research; C.G., M.H.M., and H.J.K. analyzed data; and C.G. and H.J.K. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. JX847651–JX847658).

¹To whom correspondence should be addressed. E-mail: hjklee@ufl.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1216515109/-DCSupplemental.

correlation with liking indicates that it is desirable to reduce the contents of acetate esters in tomato fruits.

To gain insight into the regulation of acetate ester accumulation in tomato, we examined ester variation in wild accessions of multiple species within the tomato clade. The volatile profiles of the wild accessions showed wide variation in the total acetate ester content (Fig. 1). All of the *S. lycopersicum* and *S. pimpinellifolium* accessions had a low content of acetate esters relative to the total amount of volatiles emitted by the fruits. In contrast, the green-fruited species accumulated much higher levels of acetate esters. In some cases, such as *S. pennellii*, they formed one of the most abundant groups of volatiles detected. Not surprisingly, the proportion of each ester volatile varies between the different accessions and species (Table S2). This suggests multiple factors determining the content of each ester, including substrate availability, acetyltransferase content and specificity, and esterase activity. Despite these variations in specific ester levels, the total content of acetate esters in red-fruited species is substantially lower than in the green-fruited species. The overall aromas of the green-fruited species have a distinct note of ester volatiles that can be described as fruity and banana-like. *S. cheesmaniae* and *S. galapagense* are somewhat intermediate, having more acetate esters than *S. lycopersicum* and *S. pimpinellifolium*.

To identify the cause of the variation in acetate ester content between the red- and green-fruited species, two IL populations derived from crosses between tomato and *S. pennellii* (LA0716) or *S. habrochaites* (LA1777) were used (27, 28). Several lines with much higher ester contents contained a common segment derived from the wild species at the bottom of chromosome 1 (Table 1 and Fig. S1), indicative of a QTL affecting ester content at this position. The most abundant esters in those lines are *cis*-3-hexenyl acetate, hexyl acetate, 2-methylbutyl acetate, isobutyl acetate, and 3-methylbutyl acetate. Although the increase in the amount of each compound varied, the QTL affected all of the acetate esters that were detected. Further fine mapping allowed us to localize the QTL to a region of ~700 kb. Analysis of the tomato genome sequence indicated the presence of a group of

five homologous, tandemly arranged esterase genes (88–92% DNA identity) within the introgressed segment. On the basis of homology, these esterase enzymes could be classified as carboxylesterases. Carboxylesterases hydrolyze carboxylic esters to produce an alcohol and a carboxylate (Fig. 2). Therefore, the enzymes were annotated SICXE1–SICXE5. Quantitative PCR analysis revealed that only *SICXE1* is highly expressed in ripe fruit with total *SICXE2–5* transcript levels less than 0.3% of total *SLCXE* mRNA (Table S3). The expression of *SICXE1* was also analyzed in different tissues as well as different stages of fruit development. The transcript was mostly found in the fruit, and its expression increased during ripening (Fig. S2).

Silencing of *SICXE1* in Tomato. To determine whether this candidate esterase is responsible for the higher ester levels associated with the QTL, its function was assessed by gene silencing in transgenic plants. Several lines with reduced expression of *SICXE1* were obtained in the cultivar (cv.) Flora-Dade background. Due to the high homology of *SICXE1* with *SICXE2–SICXE5*, these lines also had a reduced transcript content of the four other tandem genes, although not to the same extent as *SICXE1* (Table S4). Volatiles from the ripe fruits of three lines with low *SICXE1* expression (less than 6% of the control) and the control were measured to determine changes in volatiles content (Fig. 3). The silenced lines accumulated high amounts of acetate esters, similar to those observed in the QTL-containing ILs. This result confirms the role of *SICXE1* in regulating ester content of the tomato fruit.

Enzymatic Activity of *SICXE1*. On the basis of the tomato genome sequence (29), *SICXE1* is unique compared with the four other esterases (*SICXE2–5*) due to an insertion of 27 amino acids in the middle of the protein (Fig. 4 and Fig. S3). This insertion event is present in all of the *CXE1* orthologs in the species investigated within the genus, including *Solanum tuberosum* (Fig. S4). However, we were not able to find this insertion in any other esterase in the National Center for Biotechnology Information database (30). The presence of this insertion in an acetyl esterase seems to be limited to the Solanaceae family and probably only to the *Solanum* genus.

The increase in all of the acetate esters in the ILs derived from green-fruited species suggested a broad activity of *SICXE1*. To confirm that point, we examined *in vitro* activity of recombinant *SICXE1*. Because the presence of an insertion in the middle of the enzyme could impact the activity, a truncated form of *SICXE1* was produced (Fig. 4). The truncated form (*SICXE1del*) lacks the unique sequence of *SICXE1* and is therefore more similar to *SICXE2–5*. *SICXE1*, *SICXE1del*, and *SICXE5* were expressed in *Escherichia coli* BL21 and purified. *SICXE5* is the most highly expressed (albeit at a level far below *SICXE1*) of the other *SICXEs* in the gene cluster (Table S3). The affinity constant (K_m) and turnover number (k_{cat}) were determined for the three enzymes using the acetate esters detected in tomato (Table 2 and Table S5). As expected, *SICXE1* exhibited activity against all of the tested acetate esters although the enzyme was more active with some substrates, including 3-methylbutyl acetate and hexyl acetate. The kinetic parameters are similar to the published values for other carboxylesterases (www.brenda-enzymes.info). Interestingly, the truncated form of *SICXE1* retained only a fraction of its activity against the various substrates and had activity comparable to *SICXE5*.

Differences in the 5' Proximal Sequences of *CXE1* Within the Tomato Clade. To determine whether the high content of esters observed in the green-fruited species was due to differential expression of *CXE1*, mRNA levels in LA3916, Il-1–4 and their respective tomato parents LA4024 and M82 were determined. The *S. pennellii* and *S. habrochaites* ILs both had low *CXE1* RNA in comparison with their parental controls (Fig. 5). That decreased RNA content correlates with the higher content of acetate esters observed in the ILs. We also wanted to know if the expression of *CXE2–*

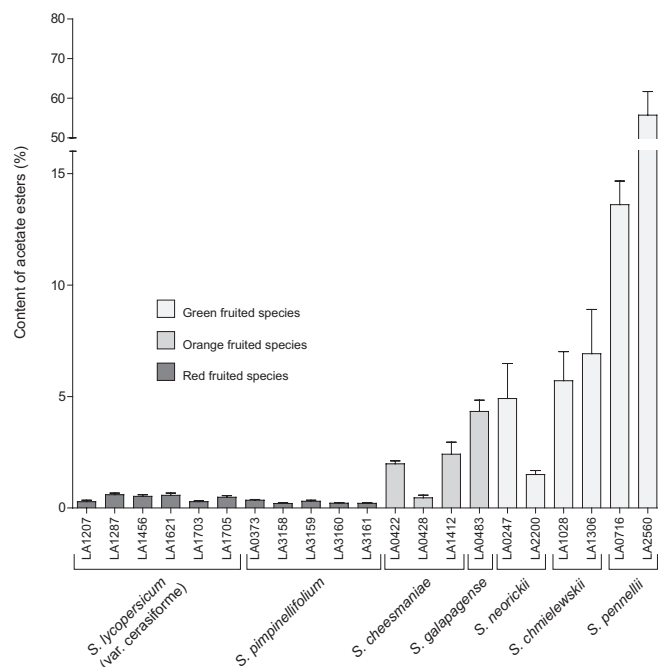


Fig. 1. Ester volatile content in wild accessions of species belonging to the tomato clade. Content of acetate esters relative to the total amount of volatiles (fresh mass standardized) was measured in ripe fruits of green-, orange-, and red-fruited species (\pm SE). Below each bar is the relevant accession number from the Tomato Genetics Resource Center.

Table 1. Acetate ester content of ILs

Volatile	<i>S. pennellii</i> IL (ng/gfw/h)			<i>S. habrochaites</i> IL (ng/gfw/h)		
	IL1-4	M82 (Ctrl)	Fold change*	LA3916	LA4024 (Ctrl)	Fold change*
Propyl acetate	0.82	0.27	3.0	4.03	0.59	6.9
sec-Butyl acetate	1.78	0.51	3.5	16.00	2.05	7.8
Isobutyl acetate	5.56	0.66	8.5	23.51	1.44	16.4
Butyl acetate	0.93	0.13	7.0	3.89	0.28	14.1
3-Methylbutyl acetate	1.85	0.10	18.5	16.46	0.09	186.8
2-Methylbutyl acetate	5.61	0.49	11.6	22.95	1.05	21.8
Pentyl acetate	1.02	0.22	4.7	3.59	0.31	11.5
cis-3-Hexenyl acetate	18.69	2.31	8.1	48.08	3.57	13.5
Hexyl acetate	13.93	1.05	13.3	29.46	1.77	16.6
Benzyl acetate	0.06	0.03	1.9	0.38	0.03	14.1
Phenylethyl acetate	0.08	0.01	6.5	0.64	0.05	12.4
Total acetate esters	50.31	5.78	8.7	168.98	11.22	15.06

gfw, gram of fresh weight.

*P value <0.01 for all of the volatiles.

CXE5 was similarly reduced in the green-fruited species in comparison with tomato. In contrast to *CXE1*, the levels of expression of *CXE2–CXE5* for *S. pennellii* are not substantially reduced and are, as in tomato, weakly expressed in the fruit (Table S3). Therefore, we concluded that the large reduction in *S. lycopersicum* ester content is specifically correlated with higher levels of *CXE1* transcript.

The higher level of *CXE1* expression in *S. lycopersicum* suggested a difference in the transcriptional promoter between tomato and the ILs derived from *S. pennellii* and *S. habrochaites*. Therefore, the DNA sequences of the promoter regions of tomato and the two parents of the ILs were determined. The promoter region of tomato contains an insertion of ~2,300 bp in proximity (~110 bp) to the start codon of the predicted ORF (Fig. 5). Alignment against the Plant Repeat Databases (31) allowed us to identify the insertion as a copia-like retrotransposon. Sequences of a copia-like retrotransposon are abundant in the genome of tomato although substitutions in the long terminal repeat sequences (LTR) indicate that they are now mostly inactive (29). Analysis of the *SICXE1* ESTs in the Sol Genomics Network Database (32) indicates that the 5' UTR of the gene extends inside the LTR of the retrotransposon, indicating that the retrotransposon contains promoter elements that result in a high expression in the fruit of tomato.

We were interested in determining whether the presence of the retrotransposon correlates with ester levels across species. Therefore, the regions 5' to *CXE1* from eight species within the clade were sequenced. The promoter region of *S. pimpinellifolium*

also contains the copia-like retrotransposon. The same insertion was also found in the orange-fruited species from the Galapagos Islands, *S. galapagense* and *S. cheesmaniae*. However, none of the green-fruited species (*S. pennellii*, *S. habrochaites*, *S. neorickii*, and *S. chmielewskii*) contain the retrotransposon. This observation suggests an insertion time more recent than the divergence between the green- and the red-fruited species. By assuming that the two LTR of the retrotransposon were identical at the time of the insertion, it is possible to estimate that the insertion event occurred ~2.2 million years ago. This estimate places the retrotransposon insertion after the divergence between the red- and the green-fruited species, but before the speciation inside the group of the red-fruited species (33, 34).

Discussion

Identification of genes that impact flavor has been an important facet of the effort to improve the quality of domesticated fruits and vegetables. Commercial cultivars of several crops including tomato, melon, and strawberries are often viewed by consumers as having poor flavor. Modern breeding programs have emphasized characters such as yield, disease resistance, appearance, and postharvest shelf life. The large number of genes that impact volatiles production and the lack of knowledge about them present a major challenge to the breeder (4). Characterization of volatiles pathways will provide valuable tools to include flavor as an essential part of the breeding process. In this study, we identified *SICXE1* as a key regulator of ester volatile content in tomato fruits. Relatively high *SICXE1* expression in tomato results in low acetate ester content, affecting the volatile profile and therefore the overall aroma of the fruit. Although *SICXE1* plays a definite role in the overall content of esters, other factors are likely to influence the content of each acetate ester. In particular, the substrate pools as well as the substrate specificities of the alcohol acetyltransferase(s) are likely to affect the accumulation of the different esters in tomato and in the different wild species. Because none of the known alcohol acetyltransferases map to the bottom of chromosome 1 in the published tomato genome sequence (29), the large increase in ester content associated with the chromosome 1 QTL indicates a major role for *SICXE1* in differentiating the green- and red-fruited species. This hypothesis is consistent with the ester changes observed in the transgenic plants. Together, the results highlight the importance of catabolism in determining the overall content of plant volatiles.

The unique sequence of 27 amino acids inserted into the center of *SICXE1* significantly increased the affinity and turnover number of the enzyme. Removing that sequence decreases the affinity and the turnover number of *SICXE1* against all of the acetate esters that were tested. This difference in activity is

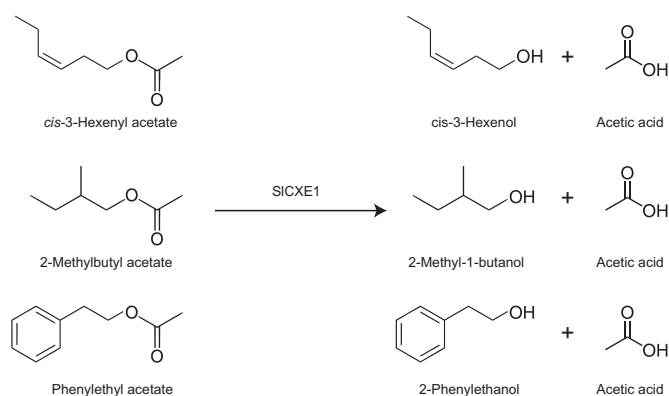


Fig. 2. *SICXE1* reaction with ester volatiles in tomato. *SICXE1* catalyzes the conversion of acetate esters to their corresponding alcohol plus acetic acid. The enzyme is active against a diversity of acetate esters as illustrated here.

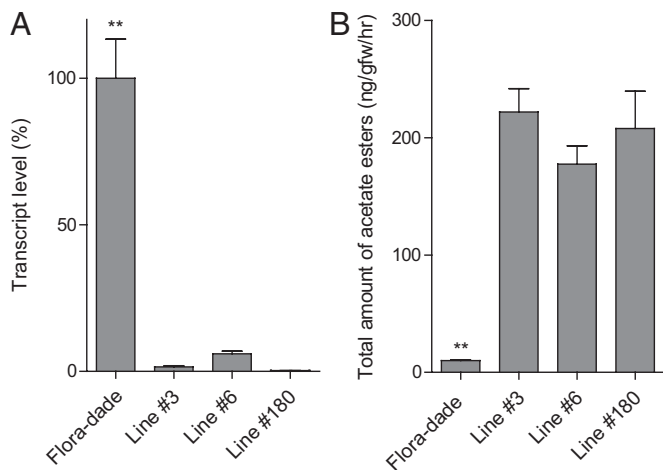


Fig. 3. Silencing *SICXE1* in tomato increases the emission of acetate esters. (A) Transcript levels of *SICXE1* in ripe fruits of three silenced transgenic lines relative to their control, Flora-Dade (\pm SE). (B) Emission rate of acetate ester volatiles from the cut tomato fruits of transgenics plants and the control (\pm SE). ** $P < 0.01$.

particularly intriguing because we did not find this inserted sequence in any esterase outside of the *Solanum* genus. Computer modeling of *SICXE1* and its paralog *SICXE5* (Fig. 4) indicates that this insertion is located adjacent to the active site of the enzyme. This insertion, therefore, likely facilitates more efficient binding and/or turnover of the substrate.

The unique 27-amino-acid insertion into *SICXE1* results in a substantial increase in enzyme activity relative to its paralogs. In turn, expression of *SICXE1* in ripening fruits is much higher than that of any of its paralogs. Taken together, these results confirm an important role for *SICXE1* in determining the flavor of tomato. The elevated expression of *SICXE1* is associated with the presence of a copia-like retrotransposon in the promoter region of the gene. Copia-like retrotransposons are highly abundant in the tomato genome (29), and they were likely an important source of species diversification. Insertion of a retrotransposon in the transcriptional promoter of a gene can have a detrimental effect on its expression as illustrated by the inactivation in some white grape cultivars of a transcription factor controlling the

accumulation of red anthocyanin pigments (35). Retrotransposons can also positively affect gene expression as recently demonstrated in blood oranges (36). The insertion of a copia-like retrotransposon in front of Ruby, a MYB transcriptional activator of anthocyanin production, is responsible for activation of the gene and subsequent accumulation of the anthocyanin pigment. As with *SICXE1*, the 5' UTR region of Ruby starts within the LTR of the retrotransposon, indicating the presence of promoter elements inside the retrotransposon.

Insertion of the retrotransposon in front of *CXE1* most likely occurred after the divergence of the green- and red-fruited species within the tomato clade. This time frame is supported by the absence of the retrotransposon in the green-fruited species and their high level of ester volatiles. The fact that the insertion was evolutionarily fixed in the red-fruited species suggests that high expression of the esterase and consequent low ester content provided an evolutionary advantage. It is also interesting that, despite having the retrotransposon insertion, *S. galapagense* and *S. cheesmaniae* accumulate more ester volatiles than *S. lycopersicum* and *S. pimpinellifolium*. *S. galapagense* and *S. cheesmaniae* are endemic to the Galapagos Islands, and they face different selection pressures than *S. lycopersicum* and *S. pimpinellifolium*. Indeed, there are no documented cases of *S. cheesmanii* frugivory by any species endemic to the Galapagos whereas there are examples of frugivory for the introduced *S. lycopersicum* (37). The precise factor(s) driving genetic fixation of the insertion and the subsequent expression of the esterase remain to be determined. Volatiles, including esters, perform multiple functions from plant defense to attraction of specific seed-dispersal agents. Fruits of different species in the tomato clade vary considerably in their biochemical composition and appearance (38). Although it is likely that multiple factors influence the fitness of each species in their environment, our results are consistent with an important role for acetate esters in fitness.

The difference in the ester volatile content between the green- and the red-fruited species within the tomato clade demonstrates how closely related species have evolved to accumulate a different balance of volatiles. Given the complexity of volatile production in plants, it is particularly interesting that a single insertion event in front of a gene is responsible for so much variability in the volatiles profile. The high activity of *SICXE1* caused by its unique 27-amino-acid insertion makes that enzyme a good candidate for examining the functions of acetate esters in fruits and flowers in situ. The negative correlation of acetate ester volatiles in tomato

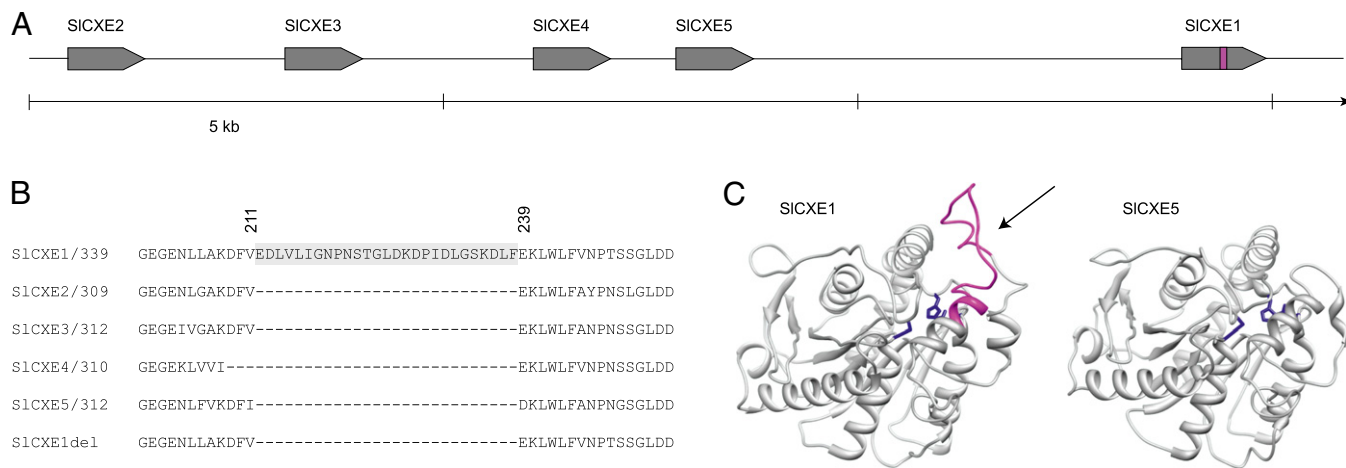


Fig. 4. *SICXE1* contains a unique amino acid insertion in the middle of its sequence. (A) Scheme of the five tandem esterases on *S. lycopersicum* chromosome 1. (B) Alignment of part of the protein sequence of the five esterases and the truncated version of *SICXE1*. The extra 27 amino acids of *SICXE1* are between positions 211 and 239 of the 339-amino-acid sequence. (C) Structure modeling of *SICXE1* and *SICXE5*. The unique sequence in *SICXE1* (in magenta) forms a loop near the active site of the protein. The residues in blue are thought to be the active site based on homology with other carboxylesterases.

Table 2. Enzymatic activity of SICXE1, SICXE5, and SICXE1del with multiple tomato volatile esters

Volatile	K_m (mM)			k_{cat} (s^{-1})			k_{cat}/K_m		
	SICXE1	SICXE1del	SICXE5	SICXE1	SICXE1del	SICXE5	SICXE1	SICXE1del	SICXE5
2-Methylbutyl acetate	0.65	2.12	1.82	84.89	1.19	1.34	131.45	0.56	0.74
3-Methylbutyl acetate	0.62	3.22	3.89	88.71	1.17	1.82	144.24	0.36	0.47
<i>cis</i> -3-hexenyl acetate	0.47	1.57	3.38	68.97	3.05	3.35	145.60	1.95	0.99
Hexyl acetate	0.26	1.45	1.54	65.05	1.94	2.94	250.87	1.34	1.91
Isobutyl acetate	0.95	2.14	6.69	34.48	0.40	0.32	36.19	0.19	0.05

taste panels indicates that SICXE1 is a target in breeding programs for improving the overall flavor of tomato.

In summary, we have shown that the difference in volatile ester content between the red- and green-fruited varieties of tomato is linked to the insertion of a transposable element immediately 5' of the most enzymatically active member of a family of esterases. The consequence of that insertion event is to reduce the levels of multiple esters that are negatively correlated with human preferences.

Materials and Methods

Volatiles Collection and Analysis. Volatiles were collected from chopped ripe fruits during a 1-h period as previously described (10). Volatiles collected on the SuperQ resin were eluted with methylene chloride using nonyl acetate as an internal control. The samples were separated on a DB-5 column (Agilent, www.agilent.com) and analyzed on an Agilent 6890N gas chromatograph. Retention times compared with known standards and identities of volatile peaks were confirmed by gas chromatography/mass spectrometry (GC/MS) (Agilent 5975 GC/MS, www.agilent.com).

Transgenic Plants. RNA interference constructs were made by cloning the full-length *CXE1* isolated from *S. lycopersicum* cv. M82 into the hairpin RNA expression plasmid pB7GWIWG2(I) (39). Expression of the hairpin RNA is under the control of the constitutive 35S promoter. *S. lycopersicum* cv. Flora-

Dade cotyledons were transformed by *Agrobacterium*-mediated transformation (40) using phosphinothricin as a selective agent. Volatiles from the fruit of transgenic plants were collected as described above. The phenotypes were heritable across multiple generations.

Protein Purification and Enzymatic Assay. *SICXE1*, *SICXE5*, and *SICXE1del* were cloned into pGEX-3x (GE Healthcare, www.gelifesciences.com) and transformed into *E. coli* BL21-DE3. Cells were grown at 37° to an OD₆₀₀ of 0.6. Induction of protein expression was done using isopropyl β -D-1-thiogalactopyranoside at a final concentration of 0.5 mM. Induced cells were grown for 4 h at 37° before being harvested by centrifugation. Following freeze-thaw cell disruptions, proteins were purified according to the supplier recommendation using Glutathione Sepharose 4B (GE Healthcare, www.gelifesciences.com).

Enzyme activity was determined by measuring acetic acid production in a spectrophotometric coupled assay (acetic acid assay; R-Biopharm, www.r-biopharm.com). Enzymes were assayed at 22 °C. The ester volatile substrates were diluted at different concentrations in dimethyl sulfoxide. Acetic acid concentration was measured using a standard curve. The esterase enzymatic reaction was confirmed by analyzing the products on GC/MS with solid-phase microextraction (Agilent 5975 GC/MS, GC sampler 80, www.agilent.com).

Protein Modeling. Computer modeling of the protein structure was performed using the I-TASSER server (41). Protein visualization and rendering

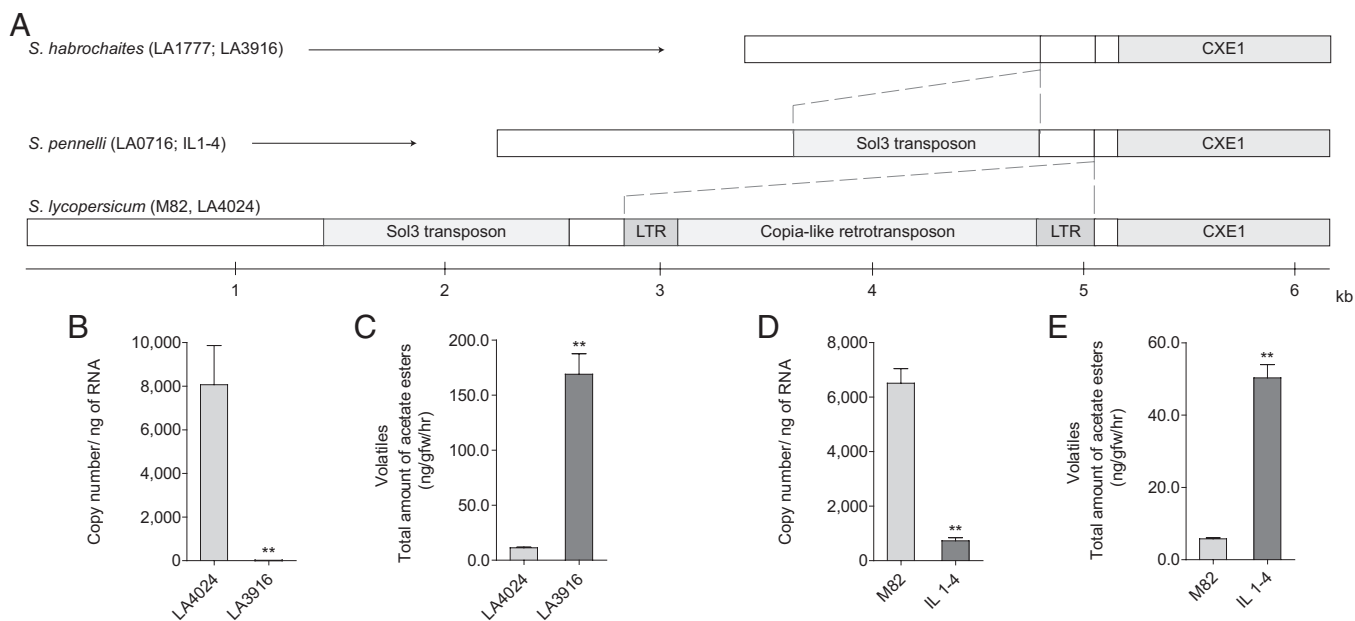


Fig. 5. High expression of *SICXE1* in tomato is associated with the insertion of a retrotransposon. (A) Scheme of the genomic region in front of *CXE1* on chromosome 1. The green-fruited species and their respective ILs lack the copia-like retrotransposon in the promoter region of the esterase. The promoter regions of all species examined also contain a Sol3 transposon, with the exception of *S. habrochaites*. (B and D) Transcript levels (\pm SE) of *CXE1* in the ILs of the green-fruited species *S. habrochaites* (B) and *S. pennellii* (D) with their tomato control (\pm SE). The esterase is highly expressed in the tomato control but only weakly expressed in the ILs of the green-fruited species. (C and E) Impact of *CXE1* expression on the acetate ester content. The ILs of *S. habrochaites* (C) and *S. pennellii* (E) accumulate more acetate esters than their respective control. ** $P < 0.01$.

was done using the USCF Chimera package (University of California, San Francisco) (42).

Quantitative PCR. Samples were quickly frozen in liquid nitrogen and kept at -80°C until further use. RNA was extracted using Plant RNA reagent (Invitrogen, www.invitrogen.com). Possible genomic DNA contamination was removed by DNase treatment (Qiagen, www.qiagen.com), and RNA was purified using GeneJET Plant RNA Purification (Fermentas, www.fermentas.com). Absolute quantification was generated using a standard curve. Quantitative PCR was performed on a StepOnePlus Real-time PCR system using Power SYBR Green RNA-to-CT 1-Step Kit (Applied Biosystems, www.appliedbiosystems.com).

Sequence Analysis. Sequencing of *CXE1* and the upstream genomic region was performed on the following accessions: *S. lycopersicum* (M82), *S. pennellii* (LA0716), *S. habrochaites* (LA1777), *S. neorickii* (LA0247), *S. chmielewskii* (LA1306), *S. cheesmaniae* (LA0428), *S. galapagensis* (LA0483), and *S. pimpinellifolium* (LA1589). Seeds of the wild accession were obtained from the Tomato Genetic Resource Center (<http://tgrc.ucdavis.edu>). Potato genomic sequence was retrieved from the SGN database (32). Multiple sequence alignment was performed with MUSCLE (www.ebi.ac.uk/tools/msa/muscle). Identification of the transposable elements was done with the Plant Repeat Databases (plantrepeats.plantbiology.msu.edu) (31). The insertion date of the copia-like retrotransposon was calculated with the equation $T = d/2r$, where T represents

the insertion time, d the LTR divergence, and r the substitution rate per site per year. The molecular clock of transposable elements is approximately twofold more rapid than synonymous base substitutions within genes (43). A substitution rate of 1.9×10^{-8} was therefore derived from the synonymous substitution rate of Solanaceae (44).

Statistical Analysis. Unpaired Student's t test was used for two-sample comparisons. For multiple comparisons, an ANOVA was performed followed by a Newman–Keuls test. The level of significance is indicated in each figure.

Accessions. Sequence data from this article can be found in GenBank under accession nos. JX847651–JX847658, sequences of *CXE1* in the different species of the tomato clade. Sequences of *SICXE2* (Soly01g108540), *SICXE3* (Soly01g108560), *SICXE4* (Soly01g108570), and *SICXE5* (Soly01g108580) are accessible in the SGN database (32). Sequences of *SICXE2* (Soly01g108540), *SICXE3* (Soly01g108560), *SICXE4* (Soly01g108570), and *SICXE5* (Soly01g108580) are accessible in the SGN database (32).

ACKNOWLEDGMENTS. The authors thank Dawn Bies, Dr. Yusuke Kamiyoshihara, Dr. Mark Taylor, and Peter Bliss for their helpful discussion and their technical assistance. This work was supported in part by a grant from the National Science Foundation to H.J.K. (IOS-0923312). C.G. was the recipient of a postdoctoral fellowship from Fonds Québécois de la Recherche sur la Nature et les Technologies.

- Shepherd GM (2006) Smell images and the flavour system in the human brain. *Nature* 444(7117):316–321.
- Goff SA, Klee HJ (2006) Plant volatile compounds: Sensory cues for health and nutritional value? *Science* 311(5762):815–819.
- Baldwin EA, Goodner K, Plotto A (2008) Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. *J Food Sci* 73(6):S294–S307.
- Klee HJ (2010) Improving the flavor of fresh fruits: Genomics, biochemistry, and biotechnology. *New Phytol* 187(1):44–56.
- Kessler A, Baldwin IT (2001) Defensive function of herbivore-induced plant volatile emissions in nature. *Science* 291(5511):2141–2144.
- Dudareva N, Pichersky E (2008) Metabolic engineering of plant volatiles. *Curr Opin Biotechnol* 19(2):181–189.
- Raguso RA (2008) Wake up and smell the roses: The ecology and evolution of floral scent. *Annu Rev Ecol Syst* 39(1):549–569.
- Buttery RG, Teranishi R, Ling LC, Flath RA, Stern DJ (1988) Quantitative studies on origins of fresh tomato aroma volatiles. *J Agric Food Chem* 36(6):1247–1250.
- Saliba-Colombani V, Causse M, Langlois D, Philouze J, Buret M (2001) Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theor Appl Genet* 102(2):259–272.
- Tieman DM, et al. (2006) Identification of loci affecting flavour volatile emissions in tomato fruits. *J Exp Bot* 57(4):887–896.
- Ulrich D, Komes D, Olbricht K, Hoberg E (2007) Diversity of aroma patterns in wild and cultivated *Fragaria* accessions. *Genet Resour Crop Evol* 54(6):1185–1196.
- Mathieu S, et al. (2009) Flavour compounds in tomato fruits: Identification of loci and potential pathways affecting volatile composition. *J Exp Bot* 60(1):325–337.
- Fridman E, Carrari F, Liu Y-S, Fernie AR, Zamir D (2004) Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* 305(5691):1786–1789.
- Schauer N, et al. (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat Biotechnol* 24(4):447–454.
- Buttery RG, Teranishi R, Ling LC (1987) Fresh tomato aroma volatiles: A quantitative study. *J Agric Food Chem* 35(4):540–544.
- Tieman D, et al. (2012) The chemical interactions underlying tomato flavor preferences. *Curr Biol* 22(11):1035–1039.
- Moyle LC (2008) Ecological and evolutionary genomics in the wild tomatoes (*Solanum* sect. *Lycopersicon*). *Evolution* 62(12):2995–3013.
- Peralta IE, Spooner DM, Knapp S (2008) Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, and sect. *Lycopersicon*; Solanaceae). *Syst Bot Monogr* 84:1–186.
- Spooner DM, Peralta IE, Knapp S (2005) Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes. *Taxon* 54(1):43–61.
- Macku C, Jennings WG (1987) Production of volatiles by ripening bananas. *J Agric Food Chem* 35(5):845–848.
- Reverchon E, Della Porta G, Gorgoglione D (1995) Supercritical CO₂ fractionation of jasmine concrete. *J Supercrit Fluids* 8(1):60–65.
- Young H, Gilbert JM, Murray SH, Ball RD (1996) Causal effects of aroma compounds on royal gala apple flavours. *J Sci Food Agric* 71(3):329–336.
- Beekwilder J, et al. (2004) Functional characterization of enzymes forming volatile esters from strawberry and banana. *Plant Physiol* 135(4):1865–1878.
- D'Auria JC, Pichersky E, Schaub A, Hansel A, Gershenzon J (2007) Characterization of a BAHD acyltransferase responsible for producing the green leaf volatile (Z)-3-hexenyl acetate in *Arabidopsis thaliana*. *Plant J* 49(2):194–207.
- Engelberth J, Alborn HT, Schmelz EA, Tumlinson JH (2004) Airborne signals prime plants against insect herbivore attack. *Proc Natl Acad Sci USA* 101(6):1781–1785.
- Frost CJ, et al. (2008) Priming defense genes and metabolites in hybrid poplar by the green leaf volatile cis-3-hexenyl acetate. *New Phytol* 180(3):722–734.
- Eshed Y, Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* 141(3):1147–1162.
- Monforte AJ, Tanksley SD (2000) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: A tool for gene mapping and gene discovery. *Genome* 43(5):803–813.
- Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485(7400):635–641.
- Geer LY, et al. (2010) The NCBI BioSystems database. *Nucleic Acids Res* 38(Database issue):D492–D496.
- Ouyang S, Buell CR (2004) The TIGR Plant Repeat Databases: A collective resource for the identification of repetitive sequences in plants. *Nucleic Acids Res* 32(Database issue):D360–D363.
- Bombarely A, et al. (2011) The Sol Genomics Network (solgenomics.net): Growing tomatoes using Perl. *Nucleic Acids Res* 39(Database issue):D1149–D1155.
- Kamenetzky L, et al. (2010) Genomic analysis of wild tomato introgressions determining metabolism- and yield-associated traits. *Plant Physiol* 152(4):1772–1786.
- Nesbitt TC, Tanksley SD (2002) Comparative sequencing in the genus *Lycopersicon*: Implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics* 162(1):365–379.
- Fournier-Level A, Lacombe T, Le Cunff L, Boursiquot JM, This P (2010) Evolution of the VvMybA gene family, the major determinant of berry colour in cultivated grapevine (*Vitis vinifera* L.). *Heredity (Edinb)* 104(4):351–362.
- Butelli E, et al. (2012) Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* 24(3):1242–1255.
- Heleno R, et al. (2011) Frugivory and seed dispersal in the Galápagos: What is the state of the art? *Integr Zool* 6(2):110–129.
- Schauer N, Zamir D, Fernie AR (2005) Metabolic profiling of leaves and fruit of wild species tomato: A survey of the *Solanum lycopersicum* complex. *J Exp Bot* 56(410):297–307.
- Karimi M, Inzé D, Depicker A (2002) GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci* 7(5):193–195.
- Sun H-J, Uchii S, Watanabe S, Ezura H (2006) A highly efficient transformation protocol for Micro-Tom, a model cultivar for tomato functional genomics. *Plant Cell Physiol* 47(3):426–431.
- Roy A, Kucukural A, Zhang Y (2010) I-TASSER: A unified platform for automated protein structure and function prediction. *Nat Protoc* 5(4):725–738.
- Pettersen EF, et al. (2004) UCSF Chimera: A visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
- Ma J, Bennetzen JL (2004) Rapid recent growth and divergence of rice nuclear genomes. *Proc Natl Acad Sci USA* 101(34):12404–12410.
- Wang Y, et al. (2008) Sequencing and comparative analysis of a conserved syntenic segment in the Solanaceae. *Genetics* 180(1):391–408.