# Multiple Biochemical and Morphological Factors Underlie the Production of Methylketones in Tomato Trichomes[1][W][OA]

**Imri Ben-Israel, Geng Yu, Michael B. Austin, Nazmul Bhuiyan, Michele Auldridge, Thuong Nguyen, Ines Schauvinhold, Joseph P. Noel, Eran Pichersky, and Eyal Fridman***

Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, Faculty of Agriculture, Food, and Environment, Hebrew University of Jerusalem, Rehovot 76100, Israel (I.B.-I., E.F.); Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109–1048 (G.Y., N.B., T.N., I.S., E.P.); and Jack H. Skirball Center for Chemical Biology and Proteomics, Salk Institute for Biological Studies, La Jolla, California 92037 (M.B.A., M.A., J.P.N.)

Genetic analysis of interspecific populations derived from crosses between the wild tomato species *Solanum habrochaites* f. sp. *glabratum*, which synthesizes and accumulates insecticidal methylketones (MK), mostly 2-undecanone and 2-tridecanone, in glandular trichomes, and cultivated tomato (*Solanum lycopersicum*), which does not, demonstrated that several genetic loci contribute to MK metabolism in the wild species. A strong correlation was found between the shape of the glandular trichomes and their MK content, and significant associations were seen between allelic states of three genes and the amount of MK produced by the plant. Two genes belong to the fatty acid biosynthetic pathway, and the third is the previously identified *Methylketone Synthase1* (*MKS1*) that mediates conversion to MK of β-ketoacyl intermediates. Comparative transcriptome analysis of the glandular trichomes of F2 progeny grouped into low- and high-MK-containing plants identified several additional genes whose transcripts were either more or less abundant in the high-MK bulk. In particular, a wild species-specific transcript for a gene that we named *MKS2*, encoding a protein with some similarity to a well-characterized bacterial thioesterase, was approximately 300-fold more highly expressed in F2 plants with high MK content than in those with low MK content. Genetic analysis in the segregating population showed that *MKS2*'s significant contribution to MK accumulation is mediated by an epistatic relationship with *MKS1*. Furthermore, heterologous expression of *MKS2* in *Escherichia coli* resulted in the production of methylketones in this host.

Plants exhibit a large range of chemical and morphological variation, reflecting different adaptations to mediating their interactions with the biotic and abiotic environment throughout their life cycle (Ehrlich and Raven, 1964). Some plant chemicals are lipophilic (oily) compounds that have high vapor pressure and therefore volatilize easily when exposed to air. Such volatiles can serve as signal molecules that either attract or repel animals. Many such compounds are also toxic and can damage a predatory organism through external or internal contact and are often synthesized in dedicated cells that also serve to store them (Wagner et al., 2004). In particular, such compounds may be synthesized and accumulated in small epidermal cell extensions on the surface of leaves, stems, and reproductive tissues called glandular trichomes (Schilmiller et al., 2008). Since the initial work on glandular trichomes in mint (*Mentha* spp.; Gershenzon et al., 1992), various studies involving transcriptomics, proteomics, and metabolomics have indicated that entire metabolic pathways responsible for the production of such compounds operate within the trichomes and that these unique cells require the import of only the basic building blocks to make these chemicals (Gang et al., 2001; Fridman et al., 2005; Nagel et al., 2008; Schilmiller et al., 2008; Xie et al., 2008).

The cultivated tomato *Solanum lycopersicum* and its wild relative *Solanum habrochaites* f. sp. *glabratum* represent two of the 12 main taxa found within the *Solanum* section *Lycopersicon* (Zuriaga et al., 2008). While only limited genetic diversity is found among the cultivated *S. lycopersicum* accessions, a wide range of variance is found in the wild relatives. This richness of genetic polymorphism is well reflected by the wide repertoire and quantity of specialized compounds

[1] This work was supported by the Binational Agricultural Research and Development Fund (grant no. IS–3869–06), by the U.S. Department of Agriculture Cooperative State Research, Education, and Extension Service (National Research Initiative grant no. 2004–35318–14874), by the National Science Foundation (grant no. EEC–0813570), and by the Alon Fellowship and the Birk Foundation (to E.F.).

* Corresponding author; e-mail fridmane@agri.huji.ac.il.

The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) are: Eyal Fridman (fridmane@agri.huji.ac.il) and Eran Pichersky (lelx@umich.edu).

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.109.146415

accumulated in their trichomes, including monoterpenes and sesquiterpenes (van Der Hoeven et al., 2000; Fridman et al., 2006; van Schie et al., 2007), acyl sugars (Ghangas and Steffens, 1993), and methylketones (MK; Williams et al., 1980; Fridman et al., 2005). Up to seven types of trichomes have been reported in the various *Solanum* species. One of these trichome types that has been investigated in some detail is the type VI glandular trichome, which is composed of a stalk cell with four cells at the top that form a mushroom-like shape; a cuticular sac wrapped around these cells allows accumulation of secreted compounds similar to an inflating balloon (Snyder and Carter, 1985; Werker, 2000). We have recently shown that in the wild species *S. habrochaites* f. sp. *glabratum*, the type VI glandular trichomes, which are present at high density on both the leaf surfaces and stems, contain two main MK compounds, 2-tridecanone (2TD; containing a 13-C backbone) and 2-undecanone (2UD; containing an 11-C backbone), as well as some 2-pentadecanone (containing a 15-C backbone) and a few other unidentified MK. These MK are synthesized and accumulated to very high levels in these trichomes, up to 5,500 $\mu g\ g^{-1}$ leaf fresh weight (Antonious, 2001; Fridman et al., 2005).

Analysis of a type VI-specific EST database from a MK-producing *S. habrochaites* f. sp. *glabratum* (accession PI126449 [hereafter referred to as PI]) showed that transcripts of genes encoding plastidic enzymes of fatty acid biosynthesis are highly represented, in contrast to their relatively low representation in another tomato wild species that does not make MK (accession LA1777; Fridman et al., 2005). The comparative analysis of the two EST databases also led to the isolation and characterization of a novel gene encoding a protein belonging to the $\alpha/\beta$-hydrolase family, which was specifically and exclusively expressed in type VI trichomes of MK-producing plants but not in nonproducers. Although the protein did not appear to have a transit peptide, the results of plastid import experiments indicated that it could be imported into the plastids.

Since 3-ketoacids are inherently unstable and undergo spontaneous decarboxylation (Kornberg et al., 1948), albeit at a low rate at ambient temperature, the evidence of elevated levels of fatty acid biosynthesis in these trichomes suggested that the observed straight-chain MK such as 2TD and 2UD could be derived from enzymatic or nonenzymatic decarboxylation of the respective $C_{n+1}$ 3-ketoacids. In plants, 3-ketoacyls of fatty acids mostly occur in plastids (as 3-ketoacyl-acyl-carrier proteins [ACPs]) as intermediates in the fatty acid biosynthesis pathway and in peroxisomes (as 3-ketoacyl-CoA) as intermediates in the fatty acid degradation pathway (http://lipids.plantbiology.msu.edu/?q=lipids/genesurvey/). The identification of a plastid-localized putative hydrolase led us to carry out in vitro assays with this enzyme, subsequently designated as METHYLKETONE SYNTHASE1 (MKS1), with the C12, C14, and C16 3-ketoacyl-ACPs as substrates (Fridman et al., 2005). In these assays, the respective C11, C13, and C15 MK were produced, suggesting that MKS1 is capable of both hydrolyzing the thioester bond and decarboxylating the resulting 3-ketoacid intermediate. However, it was noted that the turnover rate of the enzyme was unusually low (Fridman et al., 2005).

Crosses between MK-producing and nonproducing lines followed by segregation analysis have indicated that the ability to produce MK requires multiple quantitative trait loci in addition to *MKS1* (Zamir et al., 1984). Consequently, it has not been possible to breed cultivated tomato lines that produce high levels of MK in their glands. It is likely that the trait of MK production in *S. habrochaites* evolved through multiple morphological and biochemical changes that took place gradually during evolution. To uncover the additional factors influencing MK production, we took a quantitative genetic approach to identify quantitative trait loci that might affect MK production, including genes encoding biosynthetic enzymes, and tested the possible relationship between trichome characteristics and chemical content. In addition, comparative transcriptomic analysis was used to identify new genes whose differential expression is correlated with MK production in interspecific populations.

## RESULTS

### Morphological and Chemical Analyses of Interspecific Populations Derived from Crosses between the Cultivated Tomato and *S. habrochaites* f. sp. *glabratum*

The chemical profiles of leaves of the cultivated tomato *S. lycopersicum* (var M82) and the wild species *S. habrochaites* f. sp. *glabratum* (PI) differ in their shape and chemical content. In particular, leaves of the cultivated tomato contain little or no MK, while leaves of the wild species contain high levels of 2UD and 2TD, which are synthesized and stored in the type VI glandular trichomes on the leaf surface (Fridman et al., 2005). A series of crosses were conducted between these accessions to genetically dissect the contribution of candidate genes to MK content. Tomato plants of different genetic backgrounds were then evaluated, including the two parental lines: *S. habrochaites* f. sp. *glabratum* (PI) and *S. lycopersicum* var M82 (14 plants of each), F1 hybrids of these parents (14 plants), an F2 segregating population derived from self-pollinated F1 (245 plants), progeny derived from the first and second backcrossing of F1 with M82 (82 and 72 plants, respectively), and progeny derived from the first backcrossing of F1 with PI (22 plants). All plants were randomly planted, and from each, six young leaflets were removed for chemical characterization and 2TD level determination, since 2TD is the major MK produced in the parental wild species (Fridman et al., 2005). Overall, the 2TD levels of most F2 progeny were more similar to the cultivated tomato parent (Fig. 1). This, combined with the observation of very low

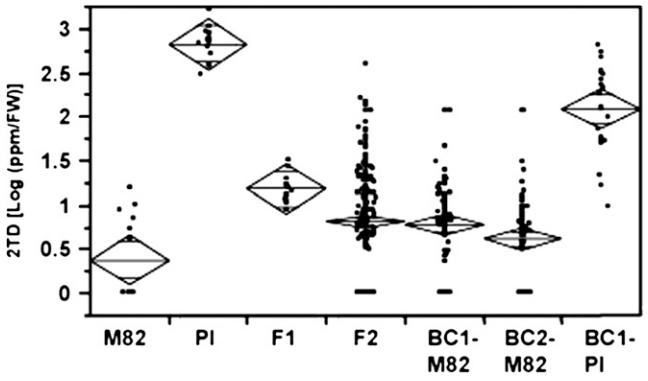


**Figure 1.** Distribution plot of 2TD levels in the cultivated tomato *S. lycopersicum* (var M82), the wild species *S. habrochaites* f. sp. *glabratum* (PI), the F1 hybrid of these parents, the F2 segregating population derived from self-pollinated F1, and progeny derived from the first and second backcrossing of the F1 with M82 (BC1-M82 and BC2-M82) and the first backcrossing with PI (BC1-PI). 2TD levels were log transformed. The line across each diamond represents the group mean. The vertical span of each diamond represents the 95% confidence interval for each group. FW, Fresh weight.

values in the backcrossed generations, indicated polygenic inheritance of this trait and suggested the recessive characteristic of the wild species alleles that participate in this pathway.

Digital images of leaflet surfaces were taken to determine trichome density and its association with MK accumulation. While analyzing these images, we noticed that the F2 population segregates not only for trichome number but also for trichome shape. This observation is in agreement with previously described distinctions in trichome shape between cultivated and wild species of tomato (Snyder and Carter, 1985; Antonious, 2001). While none of the F2 plants showed clear separation of the cells at the tip of the trichomes (as the trichomes of M82 show), 31% of the population had type VI trichomes with partial separation of these cells (M82-like; Fig. 2, A and B), 18% of the F2 progeny had round type VI trichomes, basically identical in shape to those of the wild species (PI shape; Fig. 2, A and B), and in 51% of the plants the cells of the type VI trichomes were not separated similar to the M82 parent but the trichome appeared more square than round. The latter morphology was designated as intermediate (Fig. 2, A and B). Interestingly, on average, plants with PI-shaped trichomes accumulated the highest levels of MK, plants with the intermediate trichomes accumulated intermediate levels of MK, and plants with M82-like trichomes accumulated the lowest levels of MK (Fig. 2C). The mean MK values of these three groups differed significantly from each other (Tukey's honestly significant difference; $\alpha$ = 0.005).

Since *MKS1* has been previously identified as involved in the MK pathway (Fridman et al., 2005), we looked for a possible relationship between the *MKS1* genotype and the shape of the trichomes in the interspecific segregating F2 population. There were significant differences (Pearson test, $P < 0.003$) in the frequencies of the three groups of plants with different type VI trichome shapes among the three genotypes of *MKS1*. In particular, no F2 progeny exhibited PI-shaped trichomes among homozygotes for the cultivated allele of *MKS1* (C/C), and the trichomes of most of the plants in this group bore an M82-like shape (Fig. 3).

### Association between Candidate Genes, Trichome Characteristics, and MK Content

The association between variation in candidate structural genes and 2TD content was examined in genetic mapping experiments employing these genes as simple PCR markers, cleaved amplified polymorphism sequences, or single-nucleotide polymorphism

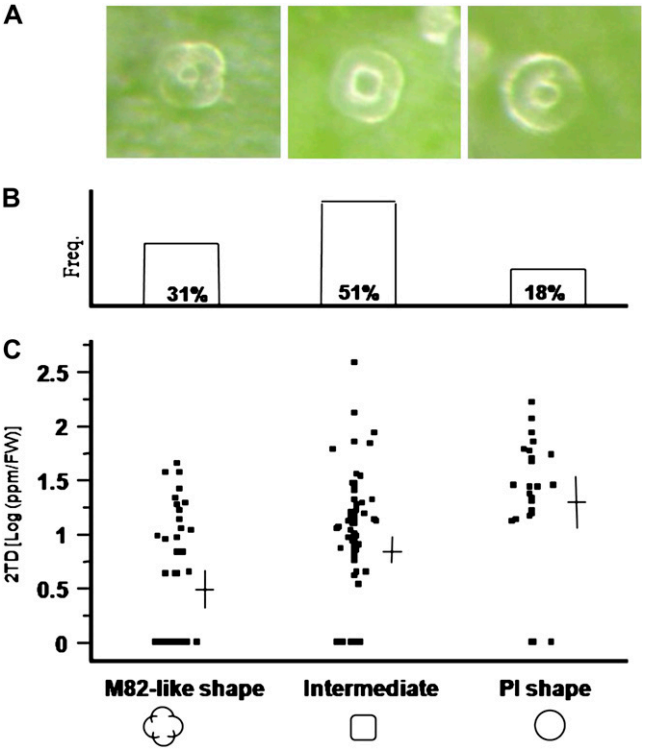


**Figure 2.** Variation and distribution of the gland shape and 2TD content in an interspecific F2 population originated from a cross between the cultivated and wild species. A, A representative binocular image of a type VI glandular trichome on the abaxial surface of a young leaflet from segregating progeny. Plants were categorized as having three types of trichomes based on six independent photographs that were taken from different leaves: M82-like shape, in which the top cells of the trichome are partially separated (left); intermediate shape, in which the cells are merged into a square-like shape (middle); and PI shape, in which the cells are merged into a globular shape (right). B, Distribution of the F2 plants among the three trichome categories. C, Distribution of the 2TD content in the three trichome categories. Horizontal and vertical lines (black crosses) represent the average and 95% confidence limits, respectively. FW, Fresh weight.

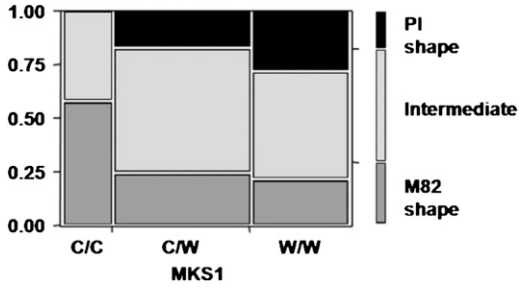


**Figure 3.** Mosaic plot of trichome shape in each of the genotypic classes of *MKS1*. C, Cultivated (M82) allele; W, wild species (PI) allele.

(SNP) markers, after aligning the open reading frame (ORF) of the alleles from both species and designing amplicons flanking the insertion/deletion, SNP, or restriction site (see "Materials and Methods"). The latter approach used high-resolution melt technology (HRM Assay Design and Analysis, CorProtocol 6000, 2006; Supplemental Fig. S1).

Since MK are derived from fatty acids, we examined the following genes from the fatty acid biosynthetic pathway (http://lipids.plantbiology.msu.edu/?q=lipids/genesurvey/) as genetic markers: acetyl-CoA carboxylase (*ACC*), malonyl-CoA:ACP transacylase (*MaCoA-ACP trans*), 3-ketoacyl-ACP synthase III (*KASIII*), 2,3-trans-enoyl-ACP reductase, 3-ketoacyl-ACP synthase I (*KASI*), *ACP1*, and *ACP2*. The *MKS1* locus was also included in the genetic screening. Association test was conducted by multiple regression analysis using the allelic state of the different genes in each F2 progeny and trichome characters as a predictor of 2TD level. This analysis (power = 0.999) showed that segregating progeny carrying the wild allele in two loci (*MKS1* and *ACC*) contain significantly ($P < 0.0003$ and $P < 0.003$, respectively) higher amounts of 2TD (Fig. 4A), while for the *MaCoA-ACP trans* locus, the opposite trend was observed: plants carrying the wild species allele had, on average, significantly ($P < 0.039$) less 2TD content. In addition, a significant positive correlation between the density and shape of the trichomes and the amount of MK in the leaves was found (Fig. 4A). This multiple regression reinforced the previous results, indicating an association between trichome morphology and 2TD levels (Fig. 2), and overall this model explained approximately one-third of the total 2TD phenotypic variation in the F2 population ($r^2 = 0.333$; Fig. 4B). To test whether the three candidate genes that are significantly associated with 2TD levels in the segregating population (*MKS1*, *ACC*, and *MaCoA-ACP trans*) exhibit differential expression between the wild and cultivated species, a quantitative reverse transcription (qRT)-PCR approach was taken. Primer pairs that fully matched both alleles were designed for each gene, and qRT-PCR was conducted using RNA from trichomes of both accessions (see "Materials and Methods"). *MKS1*, *ACC*, and *MaCoA-ACP trans* showed 355-, 2.7-, and 7.7-fold higher expression, respectively, in the trichomes of the PI parent versus those of the M82 parent (Fig. 4C).

Mapping these three genes on the tomato genome using the *Solanum pennellii* introgression line population (Eshed and Zamir, 1995; Liu and Zamir, 1999) identified genes *ACC* and *MaCoA-ACP trans* on chromosome 1, bin 1-B. Based on the F2 mapping population, the two loci are 8.8 centimorgan (cM) apart (39 recombination events in 220 F2 progeny successfully scored for both loci). *MKS1* was localized to bin 1-I, more than 50 cM away (Supplemental Fig. S2). There were no significant linkages with any of the other loci tested.

A

| Factor | Effect | Prob>\|t\| |
|---|---|---|
| MKS1 | 0.51 | 0.0003 |
| ACC | 0.37 | 0.003 |
| MaCoA:ACP trans | -0.25 | 0.039 |
| Shape | 0.23 | 0.003 |
| Density | 0.04 | 0.013 |

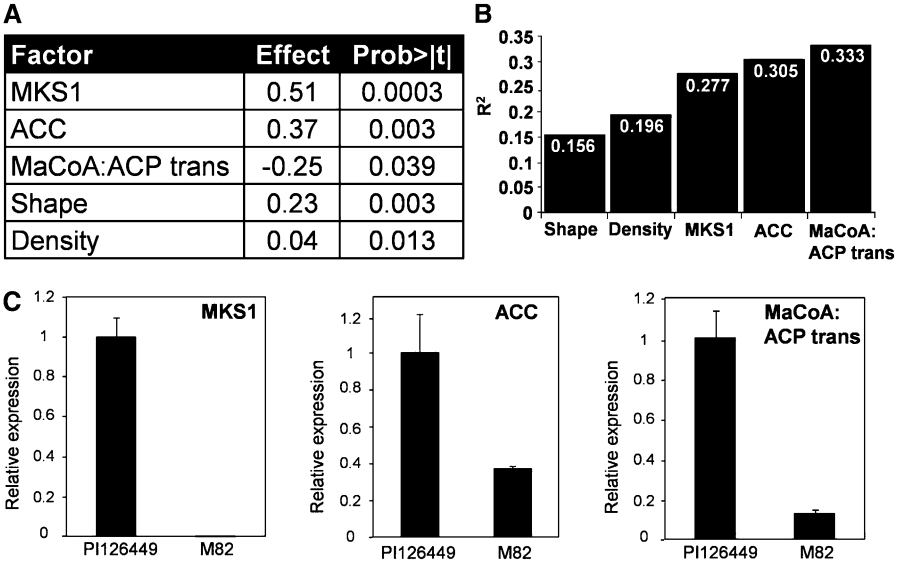


**Figure 4.** A, Association analysis of candidate genes and trichome characteristics with leaf 2TD content. The multiple regression model includes all of the tested factors for which a significant effect was found. The effect is defined as the contribution of one level/allele of the tested factor to the 2TD content. B, Accumulated variation explained by the model ($r^2$) with each additional factor. C, Expression of *MKS1*, *ACC*, and *MaCoA-ACP trans* in trichomes of the wild species *S. habrochaites* f. sp. *glabratum* (PI) and the cultivated tomato *S. lycopersicum* (var M82). mRNA abundance in isolated trichomes was determined by qRT-PCR. Expression levels for the various samples were normalized to the expression of Actin. Data are averages of three biological replicates, and error bars represent SE.

### Transcriptome Analysis of Trichomes from Bulked Segregants

Bulked segregant analysis (Michelmore et al., 1991) was used to compare transcriptomes of plants with low and high MK contents. Five plants with high levels of 2TD and five plants with no detectable 2TD were selected from the segregating F2 population (total of 245 plants) and propagated for this analysis. RNA from the trichomes of each of these two groups of plants (with high and low MK contents) was extracted, reverse transcribed, labeled, and hybridized to a custom-made microarray containing tomato genes (see "Materials and Methods"). A comparison of the hybridization results revealed a number of genes whose transcripts were present at either higher or lower levels in the high-MK-containing plants relative to their low-MK counterparts (Table I). In particular, one wild species-specific transcript of a gene that we subsequently designated *MKS2* (see below) was 337-fold more highly expressed in F2 plants with high versus low MK content, while a similar transcript, derived from the cultivated species, was 7.5-fold more highly expressed in the F2 plants with low versus high MK content (Table I).

### MKS2 Shares Sequence Identity with Hotdog-Fold Thioesterases

The MKS2 protein is 52% to 70% identical to several plant proteins with no proven functions encoded by

**Table I.** *Microarray analysis of genes differentially expressed in high- and low-MK bulks*

All genes listed showed significant difference ($P < 0.05$) in hybridization intensity between high- and low-MK bulks. cDNA sequences from the customized microarray that were up-regulated or down-regulated by more than 7-fold are shown.

| Gene Code[a] | Annotation | Ratio[b] |
|---|---|---|
| DN167657 | A protein related to a *Pseudomonas* thioesterase (Sh allele) | +336.6 |
| AI779239 | rRNA-16S ribosomal RNA | +62.6 |
| AF230371 | Allene oxide synthase | +46.7 |
| BI925004 | Plasma membrane intrinsic protein | +39 |
| BI931228 | Unknown | +27.3 |
| AW616884 | Dehydrodolichyl diphosphate synthase | +24.1 |
| DN169296 | DNA repair protein RAD23 | +18.3 |
| DB719610 | Calcium-binding EF hand family protein | +15.5 |
| DN168712 | Rubisco small subunit 1A | +14.7 |
| DN169129 | Major latex protein-related (Sh allele) | +12.9 |
| AW039905 | Peroxisomal protein involved in the activation of fatty acids | +11.8 |
| AI777019 | Unknown | +11.2 |
| AW615872 | Glycosyltransferase family 14 protein | +11.1 |
| BF097749 | Mitochondrial 26S ribosomal RNA protein | +9.2 |
| BW688217 | Unknown | +9.1 |
| BM412813 | Methyltransferase family 2 protein | +8.8 |
| DN170232 | Protein kinase | +8.4 |
| DN171038 | Casein kinase 1 protein family | +7.2 |
| DB683900 | X-Pro dipeptidase | −7.1 |
| BG131749 | A protein related to a *Pseudomonas* thioesterase (Sl allele) | −7.5 |
| AI772024 | Unknown | −7.6 |
| DB722221 | Unknown | −7.7 |
| DN168641 | PSII oxygen-evolving complex 23 | −7.8 |
| ES893822 | Cell wall protein precursor | −7.9 |
| BG643000 | Phospholipase A2$\beta$ | −8.0 |
| BW690350 | GRAM domain-containing protein/abscisic acid-responsive protein-related | −8.3 |
| AW034502 | Cytochrome P450, putative | −8.9 |
| BG1237766 | Aldo/keto-reductase family protein | −9.5 |
| BI928231 | NAD-dependent epimerase/dehydratase family protein | −10.1 |
| BI932160 | UDP-glucoronosyl/UDP-glucosyl transferase family protein | −10.6 |
| BG128416 | Unknown | −10.9 |
| AW624755 | Major latex protein-related | −11.9 |
| ES896328 | Chaperonin | −12.5 |
| BE434841 | 3-Ketoacyl-ACP synthetase 2 nuclear gene | −15.2 |

[a]Corresponding GenBank accession numbers with the highest similarity are shown. [b]Ratio depicts the difference in average ratios of high-MK over low-MK bulks in four hybridizations for all of the probes that represent the same sequence.

genes in the Arabidopsis (*Arabidopsis thaliana*) and rice (*Oryza sativa*) genomes, by many ESTs from various plant species of the angiosperm family, as well as from white spruce (*Picea glauca*; Fig. 5). Sequence similarity established that the 149-residue protein encoded by *MKS2* is a member of the 4-hydroxybenzoyl-CoA thioesterase (4HBT) subfamily of hotdog-fold enzymes (Dillon and Bateman, 2004). Although only recently discovered, hotdog domains constitute a broad superfamily that is evolutionarily unrelated to the vast $\alpha/\beta$-hydrolase-fold superfamily of (thio)esterases, of which MKS1 is a member (Hotelier et al., 2004). Notably, several hotdog-fold subfamilies required for typical fatty acid metabolism, including the FabA/FabZ 3-hydroxy-acyl-ACP dehydratases and the FatA/FatB saturated acyl-ACP thioesterases, occur instead as longer sequences that represent a tandem duplication of the hotdog fold (Dillon and Bateman, 2004; Mayer and Shanklin, 2005). BLAST search revealed the MKS2 amino acid sequence to also

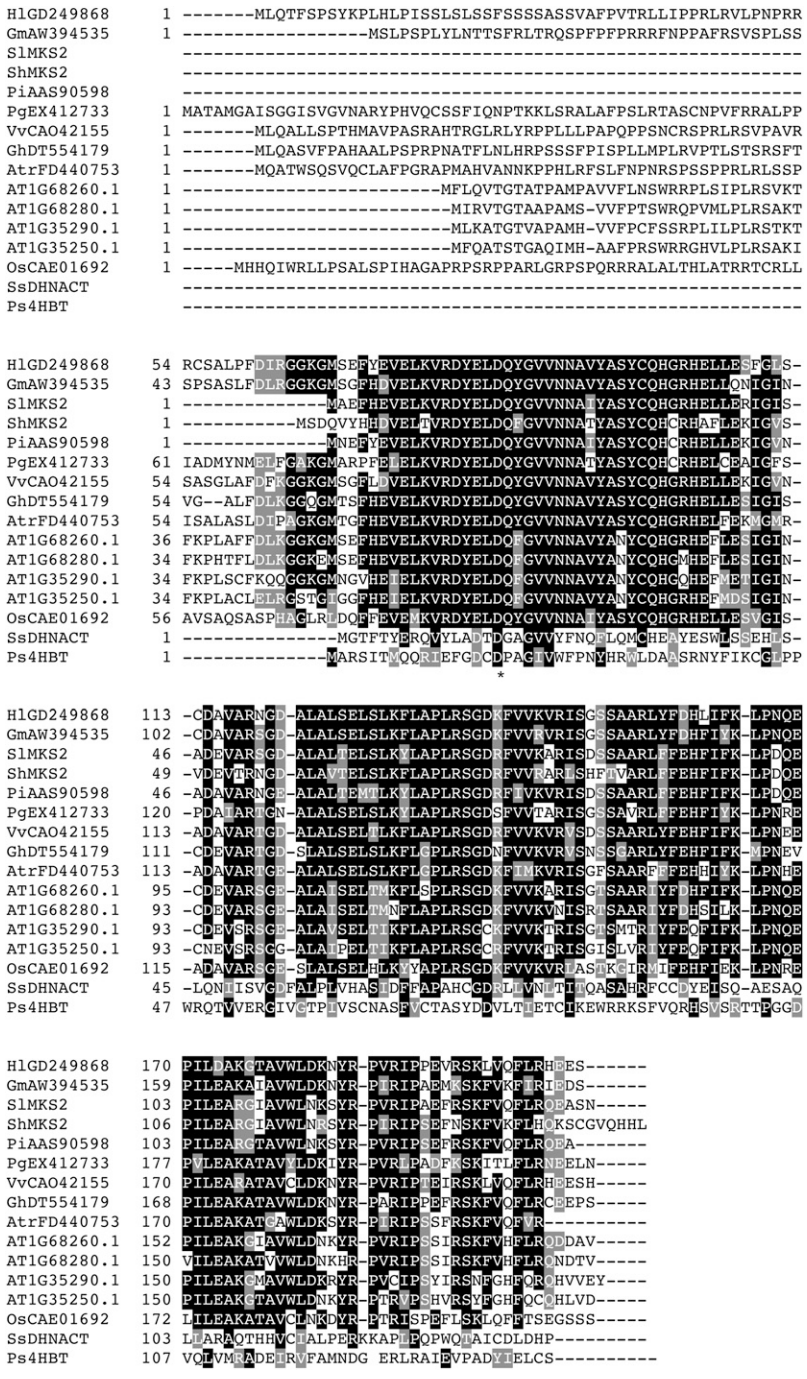

**Figure 5.** Amino acid sequence alignment of SlMKS2, ShMKS2, and related proteins. White letters on a black background indicate identical amino acids in the majority (eight or more) of the sequences. White letters on a gray background indicate conserved amino acid substitutions. The asterisk indicates the catalytic Asp residue identified in the enzyme 4HBT. The complete cDNA sequences of *SlMKS2* and *ShMKS2* indicate that the ORFs begin as indicated in this figure. At, *Arabidopsis thaliana*; Atr, *Amborella trichopoda*; Gh, *Gossypium hirsutum*; Gm, *Glycine max*; Hl, *Humulus lupulus* (cv Phoenix); Os, *Oryza sativa*; Pa, *Prunus armeniaca*; Pi, *Petunia integrifolia* subsp. *inflata*; Pg, *Picea glauca*; Ps, *Pseudomonas* sp. (strain CBS-3); Sh, *Solanum habrochaites*; Sl, *Solanum lycopersicum*; SsDHNACT, *Synechocystis* sp. PCC6803 1,4-dihydroxy-2-naphthoyl-CoA thioesterase; Vv, *Vitis vinifera*. Accession numbers are as follows: Atr, FD440753; Gh, DT554179; Gm, AW394535; Hl, EX521228; Os, CAE01692; Pi, AAS90598; Pg, EX412733; SsDHNACT, NP442358; Vv, CAO42155.

be 34% identical (and 48% similar) over 131 aligned residues to a structurally characterized putative thioesterase from *Thermus thermophilus* (Protein Data Bank [PDB] code 1Z54), which in turn shares 26% sequence identity and 51% similarity with 4HBT over 91 aligned residues. The crystallized 1Z54 protein backbone nearly perfectly overlays with known 4HBT structure (PDB code 1L09; Supplemental Fig. S3). This "bridging" 1Z54 sequence and crystal structure firmly establish the homology of MKS2 to the well-characterized 4HBT, and 1Z54 facilitates a structural homology-based examination of the MKS2 sequence. In addition, MKS2 conserves the catalytic Asp-17 of 4HBT, although our model predicts extensive substitution of juxtaposed residues in the MKS2 active-site cavity relative to any characterized 4HBT subfamily member.

### *MKS2* Is Associated with MK Content and Reveals an Epistatic Interaction with *MKS1*

Nucleotide differences were used to employ the *MKS2* gene as a DNA HRM marker (Fig. 6A) and to investigate the association between the allelic state in this locus and the 2TD content variation in the segregating population. The allelic variation in *MKS2* was significantly associated with 2TD content ($P < 0.0001$) and ranked as the second-most contributing factor (after *MKS1*) among the loci thus far identified in this quantitative analysis (Fig. 6B). Moreover, inclusion of this locus in the multiple regression analysis increased the $r^2$ of the model from 0.333 to 0.485 (Fig. 6C). Expression analysis by qRT-PCR showed that *MKS2* is 980-fold more highly expressed in the trichomes of the high-MK accumulator PI parent than in those of the M82 parent (Fig. 6D), similar to *MKS1*, *ACC*, and *MaCoA-ACP trans* (Fig. 4C).

In an attempt to define possible epistatic interactions between the different genetic components of the MK network, the genetic factors that significantly contribute to MK variation in the test population were evaluated for possible two-way interactions. This analysis identified a single significant interaction between the *MKS2* and *MKS1* loci (Fig. 7A). The data showed that to achieve high levels of these compounds, the plant has to carry at least one wild species allele in each of these two interacting loci. While *MKS1* shows a dominant mode of inheritance, that of *MKS2* is only partially so: all three genotypic classes differ significantly. Indeed, incorporation of the interaction between the two loci into a new regression model by grouping the plants according to the two-locus haplotypes (Fig. 7B) increased the $r^2$ of the model from 0.485 to 0.545 (Fig. 7C).

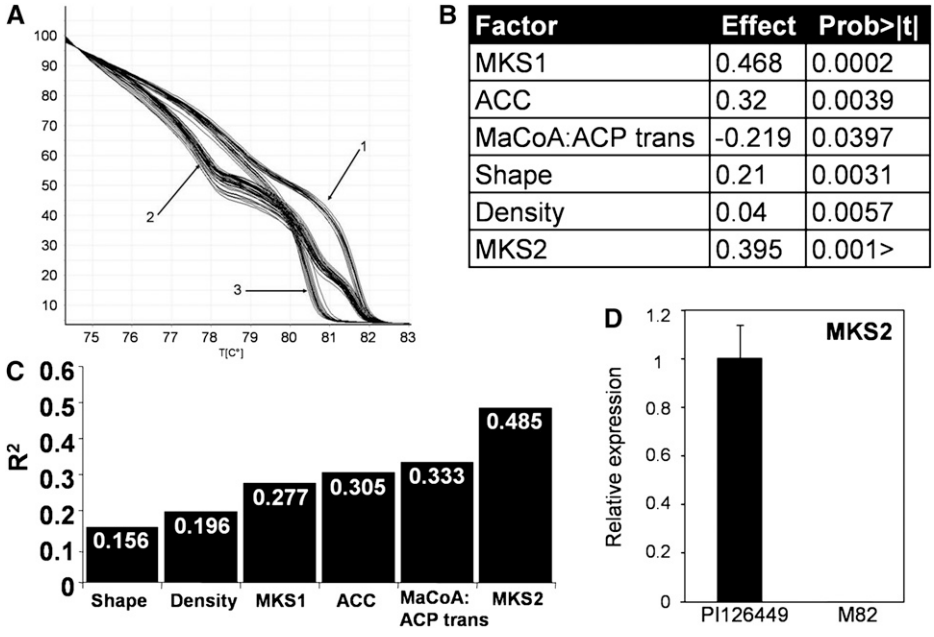


**Figure 6.** Association analysis of *MKS2* and other MK-modulating loci with 2TD levels. A, Allelic distribution at the *MKS2* locus in the F2 population using HRM marker: 1, homozygous for the M82 allele; 2, heterozygous; 3, homozygous for the PI allele. B, Multiple regression analysis for testing the association of candidate genes and trichome characteristics to the 2TD content in the leaves. The model includes all of the tested factors for which a significant effect was found. The effect is defined as the contribution of one level/allele of the tested factor to the 2TD content. C, Accumulated variation explained by the model ($r^2$) with each additional factor. D, Expression of *MKS2* in trichomes of the wild species *S. habrochaites* f. sp. *glabratum* (PI) and the cultivated tomato *S. lycopersicum* (var M82). mRNA abundance in isolated trichomes was determined by qRT-PCR. Expression levels were normalized to the expression of Actin. Data are averages of three biological replicates, and error bars represent SE.

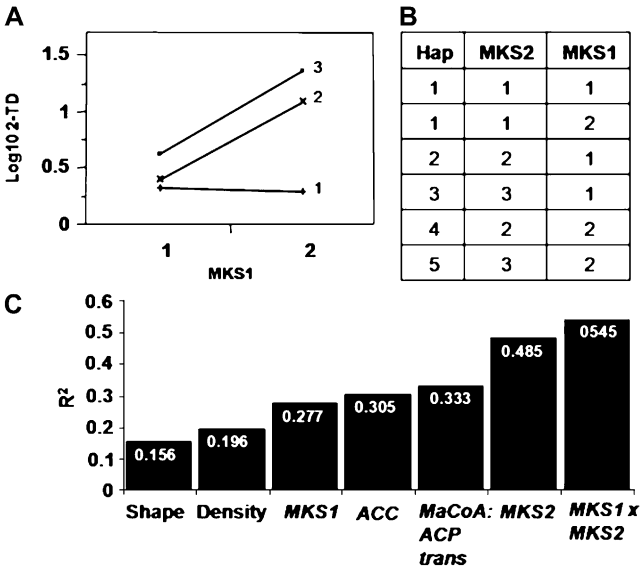


**Figure 7.** Genetic interaction of the *MKS1* and *MKS2* loci. A, 2TD least-square means plot. The *x* axis represents the genotypes of the *MKS1* locus: 1, homozygous for the M82 allele; 2, heterozygous or homozygous for the PI allele. The lines represent different genotypes of the *MKS2* locus: 1, homozygous for the M82 allele; 2, heterozygous; 3, homozygous for the PI allele. B, Levels of the haplotype (Hap) factor. *MKS1*: 1, homozygous for the M82 allele; 2, heterozygous or homozygous for the PI allele. *MKS2*: 1, homozygous for the cultivated allele; 2, heterozygous; 3, homozygous for the PI allele. C, Accumulated variation explained by the model ($r^2$) with each additional factor.

### Heterologous Expression of *MKS2* in *Escherichia coli*

To investigate the biochemical activity of *MKS2*, the full ORFs of the wild species allele (*ShMKS2*) and the cultivated allele (*SlMKS2*) were amplified and ligated into an *E. coli* expression vector (see "Materials and Methods"). These vectors were introduced into *E. coli* BL21 cells, and *ShMKS2* or *SlMKS2* expression was induced by the addition of isopropylthio-$\beta$-galactoside (IPTG; see "Materials and Methods"). After induction and overnight growth, the culture was analyzed by solid-phase microextraction of its headspace followed by gas chromatography-mass spectrometry (GC-MS; see "Materials and Methods"). The major compound in the headspace of the *E. coli* cells expressing *ShMKS2* was identified as 2TD (Fig. 8A). Lower amounts of 2UD and 2-pentadecanone were also detected, as well as the reduced alcohol forms of 2UD and 2TD (i.e. 2-undecanol and 2-tridecanol). The headspace of the *E. coli* cells expressing *SlMKS2* contained 2UD as well as 2-nonanone as the two main MK and only trace amounts of 2TD. The headspace also contained 2-nonanol and 2-undecanol (Fig. 8B). However, the major headspace compound produced by *SlMKS2*-expressing cells eluted slightly later than 2TD (peak 1 in Fig. 8B and also present at lower levels in the chromatograph in Fig. 8A). MS analysis suggested that it is a 2-tridecenone, but the position of the double bond has not yet been determined.

## DISCUSSION

### Developmental and Biochemical Connection in MK Synthesis

One of the most surprising findings of this study was the tight relationship between the shape of the trichomes and MK content (Fig. 2). The round and globular trichome shape of the wild species and its progeny was significantly associated with higher MK content. While this observation suggests that morphology constitutes a general barrier to accumulation of volatile compounds, analysis of other volatile compounds in the F2 population did not support this. For example, the distribution of one of the other major volatiles in the glandular trichomes of PI, $\beta$-caryophellene, was not correlated with trichome shape. Another possible explanation is that since cuticular waxes are complex mixtures of C20-C34 straight-chain aliphatics derived from very-long-chain fatty acids (Jetter and Kunst, 2008), the diversion of fatty acid pool toward MK comes at the expense of cuticle biosynthesis. This is also supported by the three-way relationship between MK content, trichome shape, and the genotype of *MKS1* in the segregating population (Figs. 2 and 3). However, we cannot reject the possibility that the connection between *MKS1* variation and trichome shape might be due to genetic linkage with a gene(s) that modulates the development of this specialized organ. The globular shape of the wild species trichome may be comparable to the "fused" organ morphology seen in mutants with defective cuticles (Sinha and Lynch, 1998). These fusion phenotypes have been associated with defects in several genes that modulate the biosynthesis and deposition of very-long-chain fatty acids, including enzymes (Efremova et al., 2004) and transporters (Bird et al., 2007). Together, these observations suggest a model in which enhanced activity of MK biosynthesis in the wild species may underlie the diversion of the fatty acids to MK at the expense of the synthesis of very-long-chain fatty acids, hence changing the morphology of the trichomes.

### The Genetic Basis for MK Biosynthesis in *S. habrochaites* f. sp. *glabratum*

Although MK have been found in several plant lineages, their occurrence in *S. habrochaites* f. sp. *glabratum* is unique in the *Solanum* genus, suggesting a monophyletic evolution of the specialized metabolic pathway in this subspecies. This study was aimed at identifying the genetic network required for the operation of this pathway within a single cell type, the glandular trichome. The quantitative mode of inheritance of MK in the F2 population (Fig. 1) clearly indicates that several genes are involved in the biosynthesis of these compounds and that most of the wild species alleles are recessive, as reflected in the backcrossed population (Fig. 1).

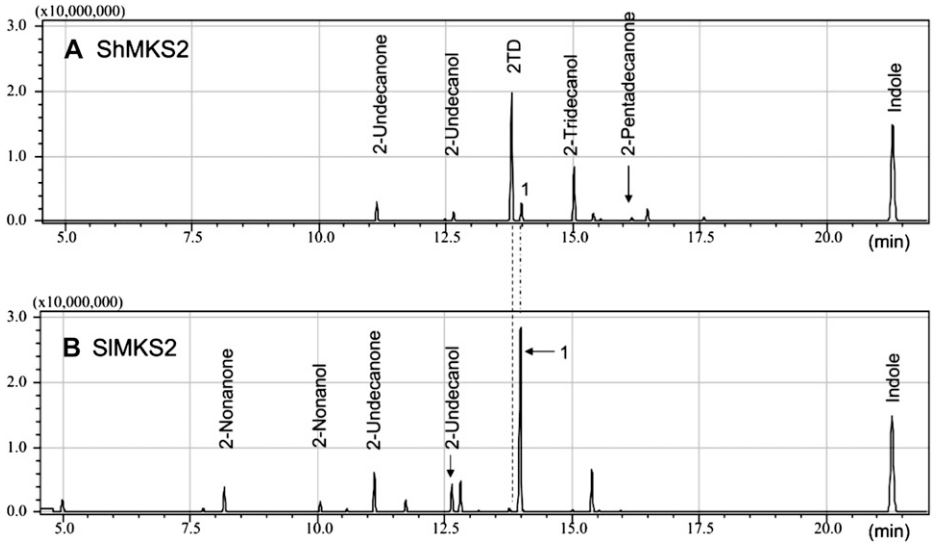


**Figure 8.** GC-MS analysis of volatile compounds produced in *E. coli* when *ShMKS2* (A) or *SlMKS2* (B) is expressed. Peak 1 is 2-tridecenone, but the position of the double bond has not yet been determined.

The multiple regression analysis showed that the *S. habrochaites* f. sp. *glabratum* alleles of genes encoding the first enzyme in the fatty acid biosynthesis pathway, ACC, as well as the enzyme MKS1 are both positively associated with MK biosynthesis. In contrast, other genes encoding enzymes that catalyze intermediate steps did not show this positive correlation. The analysis used in this study for associating candidate genes with MK variation has a few shortcomings that stem from the population structure and the lack of additional genetic markers. The availability of additional markers would have strengthened the association of the variation of MK with the candidate genes and reduced the possibility that such associations are the result of linkage disequilibrium with other causative linked and nonlinked loci. The facts that all of the genes that were included in the multiple regression analysis were also differentially expressed in the two species (Figs. 4C and 6D) and that the F2 population genotyped and phenotyped in this study is relatively large support the conclusion that a major portion of the MK variation observed in this interspecific population can indeed be attributed to diversity in these genes rather than to other genes that may be in linkage disequilibrium with these candidates. Overall, it appears that the flux in the MK pathway is controlled at the gene expression level and that the alleles from both species encode almost identical proteins that are therefore likely to be equally active.

Interestingly, the wild-type allele of the gene encoding MaCoA-ACP trans, the enzyme that acts immediately after ACC, was inversely associated with MK content. The relative contribution of this locus to the chemical variation was very low (Fig. 6), and our genetic analysis indicated that the genes encoding these two enzymes are tightly linked (8.8 cM) on chromosome 1. Since these loci act in repulsion (i.e. in the first locus the wild species allele increases MK content and in the other locus the wild species allele reduces it), the results depicted in Figure 6 are somewhat biased. The magnitude of the positive and negative additive effects of *ACC* and *MaCoA-ACP trans* loci on MK content is likely to be higher due to linkage drag, which is not included in a single-point analysis such as that conducted in this study.

The combination of a classical genetic approach (bulked segregant analysis; Michelmore et al., 1991) and transcriptome analysis of the glandular trichomes led to the discovery of a new participant in MK accumulation in these specialized cells (MKS2). Interestingly, the microarray transcriptome analysis did not detect differences in *MKS1* expression levels between bulked high- and low-MK-containing F2 plants. Genotyping the individual members of the two groups of five plants explains this unexpected result: four plants from the low-MK bulk carried the wild allele at this locus (*ShMKS1*) and three of them were homozygous, giving a total dosage of seven *ShMKS1* alleles. Similar dosages of *ShMKS1* alleles were found in the high-MK bulk in two heterozygous and three homozygous plants, giving a total of eight alleles and leading to equivalent transcripts in both bulks. Similarly, plants that accumulate no MK showed high levels of MKS1 protein in immunoblot tests of an F2 population segregating for MK content (Fridman et al., 2006). This indicates that the regulation of *MKS1* is in cis rather than in trans (expression controlled by the locus itself and not by other unlinked factors). This conclusion is also supported by the fact that the wild species' promoter can drive high levels of GFP and GUS expression in glands of the cultivated tomato (E. Fridman and R.C. Schuurink, unpublished data). In addition, these plants provided a specific and strong demonstration of the epistatic relationship occurring between *MKS1* and *MKS2* (Fig. 7). Four out of five low-MK plants that carried the *ShMKS1* allele were found to be homozygous for the cultivated allele at the *MKS2* locus (*SlMKS2*), thus lacking the wild species allele

*ShMKS2*. The fifth plant, on the other hand, presented the opposite pattern (i.e. heterozygous at the *MKS2* locus but homozygous for the cultivated allele at the *MKS1* locus, *SlMKS1*), thus lacking the wild species allele *ShMKS1*. Conversely, analysis of the two-locus haplotype in the high-MK bulk plants showed that they carried at least one wild allele at both the *MKS1* and *MKS2* loci (*ShMKS1* and *ShMKS2*).

### The Role of MKS1 and MKS2 in MK Biosynthesis

The protein with the highest sequence similarity to MKS2 that has established enzymatic activity is 4-hydroxybenzoyl-CoA thioesterase from *Pseudomonas* sp. strain CBS3. Indeed, the 1Z54 and 4HBT crystal structures reveal similar homotetrameric assemblies, which our MKS2 model also reflects (Supplemental Fig. S3). The conservation of the 4HBT catalytic Asp-17 by ShMKS2 and SlMKS2 suggests that they are likely also thioesterases. Moreover, the production of MK in *E. coli* cells expressing either allele of *MKS2* is a strong indication that the heterologous MKS2 enzyme may be capable of hydrolyzing (Fig. 9, step I) and perhaps also decarboxylating (Fig. 9, step II) 3-ketoacyl intermediates, analogous to the reaction catalyzed by MKS1. However, the production of MK in *E. coli* expressing MKS2 is not informative with regard to the specific substrates 3-ketoacyl-ACPs or 3-ketoacyl-CoA, because *E. coli* cells produce both types of substrates.

Proteins with high levels of identity to tomato MKS2 are found throughout the plant kingdom, but interestingly, all such sequences outside Solanaceae contain an N-terminal extension predicted to be a plastid or mitochondrial transit sequence (Fig. 5). SlMKS2 and ShMKS2 (and also a petunia [*Petunia integrifolia*] MKS2 homolog) lack such a transit peptide, raising the possibility that these Solanaceae proteins are not localized in the plastids and that their substrates may therefore not be 3-ketoacyl-ACPs but rather 3-ketoacyl-CoAs. The MKS2 proteins, however, do not contain any other obvious subcellular targeting sequences (e.g. no obvious PTS1 or PTS2 sequences that would target the protein to the peroxisomes).

The presence of two distinct enzymes that contribute to the production of the same compound in the same organ, and even in the same cell, is not unprecedented, and such functional redundancy was recently reported for eugenol biosynthesis in *Clarkia* (Koeduka et al., 2009). In the case of MKS1 and MKS2, however, genetic evidence for epistatic interactions between the two loci suggests that they do not act independently of each other and therefore raises some questions. Do MKS1 and MKS2 form a complex? If such a complex is formed, does each type of subunit carry out both reactions of thioester bond hydrolysis and decarboxylation (Fig. 9, steps I and II), or does each subunit catalyze only one of these reactions?

Alternatively, epistatic interactions may indicate not a physical interaction but that they act sequentially in the pathway from 3-ketoacyl intermediates to MK. A closer analysis of the genetic data reveals that although the two wild species alleles in *MKS1* and *MKS2* loci are required for the accumulation of high 2TD levels in tomato, some levels are nevertheless found in plants that carry only the *MKS2* wild allele (*ShMKS2*) but not vice versa (Fig. 7). These results suggest a model for MK biosynthesis in the trichomes in which MKS2 works upstream of MKS1. By this model, MKS2 hydrolyzes the 3-ketoacyl intermediates (Fig. 9, step I) and a low level of spontaneous decarboxylation (Fig. 9, step II) can occur to produce MK (Kornberg et al., 1948), a step that is sped up by MKS1 when present.

A resolution between these competing hypotheses will require a determination of the subcellular localization of the MKS2 protein, whether it physically interacts with MKS1 and the substrate it acts on, either independently or in complex with MKS1. The results of these experiments may in turn require a reassessment of the subcellular localization of MKS1 as well as its substrate specificity.

## CONCLUSION

The above results present the complex monophyletic evolution of a specialized pathway and highlight the power of incorporating morphological and chemical data for a detailed understanding of pathways that appear to be isolated in specialized cells. The combined data provide a framework for determining the molecular and biochemical bases for the unexpected relationships between shape and content of the glandular trichomes. Moreover, the genetic and biochemical relationship between *MKS1* and newly identified *MKS2* loci highlights the major role of epistasis interactions in determining phenotypic variation among populations and emphasizes the importance of taking it into account when dissecting the genetic basis of complex phenotypes.

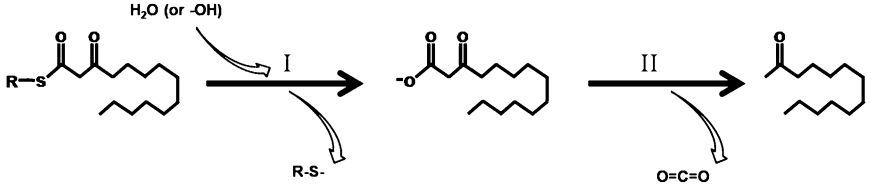


**Figure 9.** Illustration of the hydrolysis (I) and decarboxylation (II) steps that mediate MK biosynthesis from 3-ketoacyl intermediates. R represents either ACP or CoA.

## MATERIALS AND METHODS

### Plant Material, and Interspecific F2 and Backcross Populations

*Solanum lycopersicum* var M82 indeterminate and *Solanum habrochaites* f. sp. *glabratum* (PI) were obtained from the Tomato Seed Stock Center at the University of California and from the U.S. Department of Agriculture Agricultural Research Service. A single PI plant served as the male to fertilize *S. lycopersicum*. The hybrids (1) were selfed to obtain the F2 population, (2) served as the male for fertilizing M82 to obtain BC1-M82, or (3) served as the female for a PI male to obtain BC1-PI. Seeds were sprouted in trays for 2 d in a closed room at 25°C and 95% humidity and were then grown in an open greenhouse for 3 weeks. Seedlings were transplanted to the greenhouse, trellised with ropes, and grown in red loam soil with 1 $m^3$ of water and 50 mL of fertilizer (Shefer; ICL Fertilizer) per day. For bulk analysis, F2 plants were propagated by cuttings using rooting powder with 0.3% indole-3-butyric acid. Cuttings were rooted in germination trays held under spraying water for 0.5 h twice a day.

### Volatile Analysis

Six young leaflets (the first, second, and third from the first or second leaves) were sampled into scintillation vials on ice, and volatiles were extracted and analyzed as described by Fridman et al. (2005).

### Morphology Indexes

Six young leaflets (opposite those taken for volatile analysis) were sampled into scintillation vials, and a digital photograph of the central upper surface was taken. Mean trichome number per square millimeter was calculated, and trichome shape was classified as follows: wild shape (PI shape), intermediate shape (intermediate), and cultivated-like shape (M82-like shape).

### Genotyping

DNA samples were extracted from approximately 100 mg of fresh young tomato leaves and buds following the protocol described by Murray and Thompson (1980). See Supplemental Table S1 for PCR conditions and primers. *KASI* PCR products (15 $\mu$L) were digested with 1 $\mu$L of Taq restriction enzyme (New England Biolabs) for 1 h at 65°C in a reaction that included 2 $\mu$L of 10× buffer and 10 $\mu$g of bovine serum albumin (20 $\mu$L total).

### HRM Genotyping

Sequences were aligned using the Align function in the Vector NTI software package (Invitrogen) to identify SNPs between the sequences of the *S. lycopersicum* and *S. habrochaites* alleles. The *S. habrochaites* sequences were taken from an EST library produced from the glandular trichomes of PI, and the *S. lycopersicum* alleles were retrieved from the total tomato EST repository (SOL database; http://www.sgn.cornell.edu/index.pl). Three different pairs of primers flanking the identified SNPs (amplicon size varied from 60 to 100 bp per SNP) were selected for each gene using the Primer3 software (http://primer3.sourceforge.net). First, PCR was conducted with a test panel that included the parental lines M82 and PI and their hybrid (F1). PCR products were analyzed on an agarose gel (3%), and reactions that produced a single product with no primer dimers were selected for HRM analysis on a Rotor-Gene 6000 (Corbett Research). Primers that showed the best allelic discrimination by HRM examination were selected to score the genotype of the F2 population. HRM was performed immediately after the PCR cycles as a single run following the manufacturer's default parameters (see Supplemental Table S2).

### Transcriptome Analysis

Trichome isolation was performed as described by Fridman et al. (2005). The tomato microarray design, cDNA synthesis, hybridization, and analysis were performed by Genotypic Technology. The microarray was a complex of 44,000 probes of 25 bp each representing all of the tomato ESTs (Tomato Gene Index; http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=tomato). Total trichome RNA (5 $\mu$g) was labeled with Cy3 and Cy5, and hybridization was repeated four times (two repeats for each dye swap between RNA samples) in a 4 × 44 format following the version 5.5.X protocol for a two-color array. Results were analyzed following the GE-v5_95_Feb07 protocol (Agilent). More details can be found at the National Center for Biotechnology Information Gene Expression Omnibus under GSE16431.

### qRT-PCR

Total RNA was isolated from isolated glandular trichomes as described previously (Fridman et al., 2005). The RNA was subjected to DNase treatment using a DNA-free kit (Ambion), and first-strand cDNA was synthesized by SuperScript II reverse transcriptase (Invitrogen) with poly-T primers in parallel with a negative control reaction in which no SuperScript II reverse transcriptase was added. qPCR utilizing power SYBR-Green PCR master mix (Applied Biosystems), gene-specific primers, and a dilution series of each cDNA was performed as described previously (Varbanova et al., 2007). qPCR was performed using the StepOnePlus Real-Time PCR System (Applied Biosystems), and the conditions were as follows: 95°C for 3 min, 50 cycles of 95°C for 15 s, 60°C for 30 s, and 72°C for 30 s, followed by a melting cycle of 55°C to 95°C with an increasing gradient of 0.5°C and a 10-s pause at each temperature. All reactions were performed in triplicate, and each experiment was repeated twice. *ShMKS2* and *SlMKS2* allele-specific primers were designed as follows: *ShMKS2* forward, 5′-GCCTATATTGGAGGCAAGAGGA-3′, and *ShMKS2* reverse, 5′-TGTACACCGCAACTCTTCTGGT-3′; *SlMKS2* forward, 5′-ATGCAAGTTATTGCCAACATGG-3′, and *SlMKS2* reverse, 5′-GAAAAACAAACGAGCAGCTGAA-3′; *ACC* forward, 5′-CTGCTAGGAAAGCTCATCGTATGG-3′, and *ACC* reverse, 5′-GTGGTAGGAACTCCAGTGATAACG-3′; *MaCoA-ACP trans* forward, 5′-GAATGACGGTACGTCTAGCTGTTG-3′, and *MaCoA-ACP trans* reverse, 5′-GGTGAAGTCACCTGGCTAGCTAAT-3′. Actin transcript amplification was used as an internal control, with the forward primer 5′-AACACCCTGTTCTCCTGACTGA-3′ and the reverse primer 5′-AACACCATCACCAGAGTCCAAC-3′.

### Sequence Analysis

Alignment of multiple protein sequences was performed using the ClustalW program (Thompson et al., 1997).

### Statistical Analysis

Statistical analyses were conducted with JMP software (SAS Institute). Since phenotypic data of the segregating F2 and backcrossed plants did not fit the normal distribution, they were log transformed. A nonparametric test (Wilcoxon) was used to test the *MKS1* genotype effect on 2TD levels of 221 plants under the Fit Y by X function (because of unequal variances). The association between trichome shape and the 2TD levels was tested for 164 plants by ANOVA under the Fit Y by X function. For these two tests, the tested factor was set as a character and the 2TD levels were continuous. Association between the *MKS1* locus genotype and trichome shape was tested in 134 plants by Pearson test for category parameters under the Fit Y by X function. Data of 122 individuals were used for multiple regression analysis that performed by choosing the "stepwise" option in the Fit model function, and the "forward" direction was used for building the final regression model. All factors included in the analysis were set as continuous. Power was calculated with G*Power 3.1.0 software (Faul et al. 2007). Regression analysis, including the *MKS1* and *MKS2* interaction, was performed by replacing these two singular factors with a new factor representing the haplotype at those loci. Interactions between genes were tested by two-way ANOVA under the Fit model function.

### Isolation of Full-Length *ShMKS2* and *SlMKS2* cDNAs and Expression in *Escherichia coli*

The following primers were used to amplify the full ORF of MKS2 from PI (*ShMKS2*) or M82 (*SlMKS2*) leaf cDNA into the TA cloning vector (pCRT7/CT TOPO-TA; Invitrogen): forward, 5′-ATGAGTGATCAGGTCTATCACC-3′; reverse, 5′-CTCTTGATCTGGAAGCTTGA-3′. The sequences of these cDNAs were verified and transferred into the *E. coli* expression vector pET28 (Novagene). The pET28 vectors carrying ShMKS2 or SlMKS2 were mobilized into *E. coli* BL21(DES) cells, and gene expression was induced by the addition of 2 mM IPTG after the culture optical density at 595 nm value had reached 0.6

to 0.7. After IPTG addition, ShMKS2- and SlMKS2-expressing bacterial cells were grown at 30°C or 18°C overnight.

### Headspace Analysis of Spent Media of *E. coli* Cultures Expressing ShMKS2 and SlMKS2

After induction with IPTG and growth overnight, 1 mL of culture was placed in a glass vial at 42°C. The vial was capped with a screw cap in which a small hole had been bored. The needle of a solid-phase microextraction device was inserted into the vial through the hole in the cap and the fiber was extended for 30 min for volatile collection, after which the fiber was withdrawn and then injected into the GC device. GC-MS analysis was performed as described previously (Fridman et al., 2005). Labeled peaks in Figure 8 were identified by comparison of retention time and MS of authentic standards (MK) or MS and Kovac indices (alcohols).

### Homology Modeling

The MKS2 homology model was constructed using MODELLER (Sali and Blundell, 1993), and the illustration was prepared using MOLSCRIPT (Kraulis, 1991), with final rendering by POV-Ray (Persistence of Vision Ray tracer; available at www.povray.org).

Sequence data from this article have been deposited with the GenBank data library under the following accession numbers: ShMKS2 from *S. habrochaites* f. sp. *glabratum* (accession PI126449), EU883793; ShMKS2 from *S. lycopersicum* (var M82), EU908050. The Gene Expression Omnibus accession number for raw microarray data and platform description is GSE16431.

### Supplemental Data

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

**Supplemental Figure S1.** Markers developed for the candidate genes in the MK pathway and genotyping of an interspecific F2 population.

**Supplemental Figure S2.** Genetic mapping using *Solanum pennellii* introgression lines (ILs) with the HRM technology.

**Supplemental Figure S3.** Homology model of tomato MKS2 templated on the structure of a putative thioesterase from *Thermus thermophilus* (PDB code 1Z54).

**Supplemental Table S1.** Primers and PCR conditions.

**Supplemental Table S2.** Primers and HRM conditions.

## ACKNOWLEDGMENTS

We thank Dr. Raja Mugasimangalam and Rohit Shukla (Genotypic Technology, Bangalore, India) for their valuable help in designing and performing the microarray experiments. The technical help of Orly Oren, Leah Zucker, Dorit Fink, and Shula Blum (Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, Hebrew University of Jerusalem, Rehovot) is greatly appreciated.

Received August 20, 2009; accepted September 29, 2009; published October 2, 2009.

## LITERATURE CITED

**Antonious GF** (2001) Production and quantification of methyl ketones in wild tomato accessions. J Environ Sci Health B **36:** 835–848

**Bird D, Beisson F, Brigham A, Shin J, Greer S, Jetter R, Kunst L, Wu XW, Yephremov A, Samuels L** (2007) Characterization of Arabidopsis ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. Plant J **52:** 485–498

**Dillon SC, Bateman A** (2004) The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. BMC Bioinformatics **5:** 109

**Efremova N, Schreiber L, Bar S, Heidmann I, Huijser P, Wellesen K, Schwarz-Sommer Z, Saedler H, Yephremov A** (2004) Functional conservation and maintenance of expression pattern of FIDDLEHEAD-like genes in Arabidopsis and Antirrhinum. Plant Mol Biol **56:** 821–837

**Ehrlich PR, Raven PH** (1964) Butterflies and plants: a study in coevolution. Evolution **18:** 586–608

**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:** 1147–1162

**Faul F, Erdfelder E, Lang AG, Buchner A** (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods **39:** 175–191

**Fridman E, Koezuka T, Auldridge M, Austin MB, Noel JP, Pichersky E** (2006) Tomato glandular trichomes as a model system for exploring evolution of specialized metabolism in a single cell. Recent Adv Phytochem **40:** 115–130

**Fridman E, Wang J, Iijima Y, Froehlich JE, Gang DR, Ohlrogge J, Pichersky E** (2005) Metabolic, genomic, and biochemical analyses of glandular trichomes from the wild tomato species *Lycopersicon hirsutum* identify a key enzyme in the biosynthesis of methylketones. Plant Cell **17:** 1252–1267

**Gang DR, Wang JH, Dudareva N, Nam KH, Simon JE, Lewinsohn E, Pichersky E** (2001) An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. Plant Physiol **125:** 539–555

**Gershenzon J, McCaskill D, Rajaonarivony JI, Mihaliak C, Karp F, Croteau R** (1992) Isolation of secretory cells from plant glandular trichomes and their use in biosynthetic studies of monoterpenes and other gland products. Anal Biochem **200:** 130–138

**Ghangas GS, Steffens JC** (1993) UDPglucose: fatty acid transglucosylation and transacylation in triacylglucose biosynthesis. Proc Natl Acad Sci USA **90:** 9911–9915

**Hotelier T, Renault L, Cousin X, Negre V, Marchot P, Chatonnet A** (2004) ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins. Nucleic Acids Res **32:** D145–D147

**Jetter R, Kunst L** (2008) Plant surface lipid biosynthetic pathways and their utility for metabolic engineering of waxes and hydrocarbon biofuels. Plant J **54:** 670–683

**Koeduka T, Orlova I, Baiga TJ, Noel JP, Dudareva N, Pichersky E** (2009) The lack of floral synthesis and emission of isoeugenol in *Petunia axillaris* subsp. *parodii* is due to a mutation in the isoeugenol synthase gene. Plant J **58:** 961–969

**Kornberg A, Ochoa S, Mehler AH** (1948) Spectrophotometric studies on the decarboxylation of beta-keto acids. J Biol Chem **174:** 159–172

**Kraulis PJ** (1991) *MOLSCRIPT*: a program to produce both detailed and schematic plots of protein structures. J Appl Cryst **24:** 946–950

**Liu YS, Zamir D** (1999) Second generation *L. pennellii* introgression lines and the concept of bin mapping. Tomato Genet Coop **49:** 26–30

**Mayer KM, Shanklin J** (2005) A structural model of the plant acyl-acyl carrier protein thioesterase FatB comprises two helix/4-stranded sheet domains, the N-terminal domain containing residues that affect specificity and the C-terminal domain containing catalytic residues. J Biol Chem **280:** 3621–3627

**Michelmore RW, Paran I, Kesseli RV** (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA **88:** 9828–9832

**Murray MG, Thompson WF** (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res **8:** 4321–4325

**Nagel J, Culley LK, Lu YP, Liu EW, Matthews PD, Stevens JF, Page JE** (2008) EST analysis of hop glandular trichomes identifies an *O*-methyltransferase that catalyzes the biosynthesis of xanthohumol. Plant Cell **20:** 186–200

**Sali A, Blundell TL** (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol **234:** 779–815

**Schilmiller AL, Last RL, Pichersky E** (2008) Harnessing plant trichome biochemistry for the production of useful compounds. Plant J **54:** 702–711

**Sinha N, Lynch M** (1998) Fused organs in the adherent1 mutation in maize show altered epidermal walls with no perturbations in tissue identities. Planta **206:** 184–195

**Snyder JC, Carter CD** (1985) Trichomes on leaves of *Lycopersicon hirsutum*, *Lycopersicon esculentum* and their hybrids. Euphytica **34:** 53–64

**Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG** (1997) The CLUSTAL_X Windows interface: flexible strategies for multiple

sequence alignment aided by quality analysis tools. Nucleic Acids Res **25:** 4876–4882

**van Der Hoeven RS, Monforte AJ, Breeden D, Tanksley SD, Steffens JC** (2000) Genetic control and evolution of sesquiterpene biosynthesis in *Lycopersicon esculentum* and *L. hirsutum*. Plant Cell **12:** 2283–2294

**van Schie CCN, Haring MA, Schuurink RC** (2007) Tomato linalool synthase is induced in trichomes by jasmonic acid. Plant Mol Biol **64:** 251–263

**Varbanova M, Yamaguchi S, Yang Y, McKelvey K, Hanada A, Borochov R, Yu F, Jikumaru Y, Ross J, Cortes D, et al** (2007) Methylation of gibberellins by Arabidopsis GAMT1 and GAMT2. Plant Cell **19:** 32–45

**Wagner GJ, Wang E, Shepherd RW** (2004) New approaches for studying and exploiting an old protuberance, the plant trichome. Ann Bot (Lond) **93:** 3–11

**Werker E** (2000) Trichome diversity and development. Adv Bot Res **31:** 1–35

**Williams WG, Kennedy GG, Yamamoto RT, Thacker JD, Bordner J** (1980) 2-Tridecanone: naturally-occurring insecticide from the wild tomato *Lycopersicon hirsutum* f *glabratum*. Science **207:** 888–889

**Xie Z, Kapteyn J, Gang DR** (2008) A systems biology investigation of the MEP/terpenoid and shikimate/phenylpropanoid pathways points to multiple levels of metabolic control in sweet basil glandular trichomes. Plant J **54:** 349–361

**Zamir D, Selilaben-David T, Rudich J, Juvik JA** (1984) Frequency-distributions and linkage relationships of 2-tridecadone in interspecific segregating generations of tomato. Euphytica **33:** 481–488

**Zuriaga E, Blanca J, Nuez F** (2008) Classification and phylogenetic relationships in *Solanum* section *Lycopersicon* based on AFLP and two nuclear gene sequences. Genet Resour Crop Evol **56:** 663–678