Supporting Information

Schilmiller et al. 10.1073/pnas.0904113106

SI Materials and Methods

GC-MS Analyses of Terpene and Prenyl Alcohols. Extracts were injected (2 μL) into a EC-WAX column (Grace Davison; 30-m length, 0.25-μm film thickness, and 0.32-mm I.D.) on a GC17-A (Shimadzu) coupled to a GCMS-QP5000. Injector temperature was 220 °C and working on splitless mode. Interface temperature was 280 °C. The temperature program was as follows: 44 °C for 3.5 min, 5 °C per min up to 200 °C, 70 °C per min up to 275 °C, hold for 1 min, and then back to initial conditions. The same program was followed for SPME injections (used in the enzyme assays). For prenols analysis, the injection port temperature was 250 °C and the temperature program was as follows: 50 °C for 3 min, 10 °C per min up to 275 °C, hold for 1 min, and then back to initial conditions. Identification of terpenes was done by comparison to authentic standards (Sigma-Aldrich) when possible or by comparison with available retention time and spectral information (1).

cDNA Library Construction and Sequencing. Total trichomes were collected from stems and petioles of 3-week-old *S. lycopersicum* cv. M82 plants by gently scraping tissue frozen in liquid nitrogen by using a flat-end plastic spatula. Frozen trichomes were ground to a powder in liquid nitrogen by using a mortar and pestle. RNA was extracted by using the RNeasy Plant mini kit (QIAGEN) according to the manufacturer's instructions including the on-column DNaseI treatment step and RNA quality assessment using the Agilent 2100 Bioanalyzer RNA chip (Agilent Technologies). Preparation of cDNA from total RNA was done with the SMART cDNA library construction kit (Clontech). First-strand cDNA was synthesized from 2 μg of total RNA and double-strand cDNA was prepared from 2 μL of first-strand product by PCR (20 cycles), both according to the manufacturer protocols. PCRs were treated with protease K and double-stranded cDNA digested with SfiI followed by size fractionation by using a CHROMA SPIN-400 column supplied with the kit. Four 100-μL PCRs were processed individually, and fractions from the sizing column containing cDNA pooled, precipitated, and dissolved in a volume of 15 μL. Preparation of trichome cDNA for sequencing by using the GS20 sequencer (Roche) was done by the Michigan State University Research Technology Support Facility according to the Roche protocols. Reads generated from the GS-20 were processed and trimmed to remove low quality and primer sequences by using SeqClean (2). The cleaned reads were initially assembled by using CAP3 (3) followed by a second round of CAP3 to cluster the first-pass contigs and remaining singletons. First-round CAP3 parameter settings for percent match, overlap length, maximum overhang percent, gap penalty, and base quality cutoff for clipping were -p 90 -o 50 -h 15 -g 2 -c 17, respectively. For the second-round CAP3 parameters, -o was changed to 100. A translated BLAST (BLASTX) search was then performed with the final contigs against the nonredundant and green plants databases at National Center for Biotechnology Information.

Gene Expression Analysis. Total trichomes were collected in liquid nitrogen from stems and petioles of 3-week-old plants (M82, IL8-1-1, IL1-4) or greenhouse grown plants (LA0716). After extraction, 0.5 μg of total RNA was used to synthesize first-strand cDNA by using SuperScript II (Invitrogen) and oligo(dT)$_{12–18}$ primer. The resulting cDNA was diluted 10-fold and 1 μL was used as template for PCR amplification in a 30-μL reaction by using Power SYBR green PCR master mix (Applied Biosystems) and gene-specific primers. Reactions were performed by using an Applied Biosystems 7300 Real-Time PCR System with a temperature program of 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s, 55 °C for 45 s, and 72 °C for 1 min. A final dissociation step was performed to assess the quality of amplified product. Relative expression levels of *NDPS1* and *PHS1* were calculated by using the standard curve method and normalizing to expression of elongation factor 1α (*EF-1*α; GenBank accession X14449). Primer sequences for each gene are as follows: *NDPS1* (forward, 5′-TGGCAATTGGCTTACACTGAA-3′; reverse, 5′-TTATTTTATTGAACTGGCATCGTG-3′), *PHS1* (forward, 5′-AAGGAAATCTTGGAAATGAATAGAA-3′; reverse, 5′-ATAGAAGGAAAGAACAAAAGTCATAA-3′), *EF-1*α (forward, 5′-AGCTTCACTGCCCAGGTCATCATC-3′; reverse, 5′-ACCATACCAGCATCACCGTTCTTCAA-3′).

Expression of NDPS1 and PHS1 in *E. coli* and Purification of the Proteins. The ORF of *NDPS1* from Ser-45 was amplified by using the forward primer 5′-ATGTCTGCTCGTGGACTCAACAA-GAT and the reverse primer 5′-ATATGTGTGTCCAC-CAAAACG. The amplified fragment was spliced into the vector pEXP5-CT/TOPO (Invitrogen), resulting in an ORF with a C-terminal His-tag extension. The plasmid was mobilized into *E. coli* BL21 cells, and protein production and purification were carried out by using procedures described by Koeduka et al. (4).

The ORF of *PHS1* from Met 45 was amplified by using the forward primer 5′-CCATATGAATGGTTTCGAAGATG-CAA (which introduces an NdeI site that encompasses the initiating ATG codon) and the reverse primer 3′-AGGATCCT-TAATGATTGAGTGGTTTGT (which introduces a BamH1 site after the stop codon). The amplified fragment was digested with NdeI and BamH1 and spliced into the vector pET28, creating a fusion ORF with an N-terminal His tag. The plasmid was next mobilized into *E. coli* BL21 cells, and protein production and purification were carried out by using procedures described by Koeduka et al. (4), with the following modification: the buffer used for lysis and purification was 50 mM Hepes buffer, with 100 mM KCl, 7.5 mM $MgCl_2$, 20 mM imidazole, 5% glycerol, and 5 mM DTT.

1. Adams RP (2001) *Identification of Essential Oil Components by Gas Chromatography/ Quadropole Mass Spectroscopy* (Allured Publishing, Carol Stream, IL).
2. Pertea G, et al. (2003) TIGR Gene Indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. *Bioinformatics* 19:651–652.
3. Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. *Genome Res* 9:868–877.
4. Koeduka T, et al. (2006) Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester. *Proc Natl Acad Sci USA* 103:10128–10133.

agcATGagttctttggttcttcaatgttggaaattatcatctccatctctgattttacaacaaaatacatcaatatccat
gggtgcattcaaaggtattcataaacttcaaatcccaaattcgcctctgacagtgtctgctcgtggactcaacaagattt
catgctcactcaacttacaaaccgaaaagctttgttatgaggataatgataatgatcttgatgaagaacttatgcctaaa
cacattgctttgataatggatggtaataggagatgggcaaaggataagggtttagaagtatatgaaggtcacaaacatat
tattccaaaattaaaagagatttgtgacatttcttctaaattgggaatacaaattatcactgcttttgcattctctactg
aaaattggaaacgatccaaggaggaggttgatttcttgttgcaaatgttcgaagaaatctatgatgagttttcgaggtct
ggagtaagagtgtctattataggttgtaaatccgacctcccaatgacattacaaaaatgcatagcattaacagaagagac
tacaaagggcaacaaaggacttcaccttgtgattgcactaaactatggtggatattatgacatattgcaagcaacaaaaa
gcattgttaataaagcaatgaatggtttattagatgtagaagatatcaacaagaatttatttgatcaagaacttgaaagc
aagtgtccaaatcctgatttacttataaggacaggaggtgaacaaagagttagtaacttttgttgtggcaattggctta
cactgaattttacttcaccaacacattgtttcctgattttggagaggaagatcttaaagaggcaataatgaactttcaac
aaaggcatagacgttttggtggacacacatatTGAggaacaaaaagaagtttctagtatctcttgtttaatgtgtttttt
tttctttctaaaaaaataatgttatgtcgatttgttaaatcatacgagcatcattatgcacgatgccagttcaataaaat
aa

Fig. S1. *NDPS1* cDNA sequence (accession no. FJ797956). The initiating ATG codon and the stop codon of the reading frame are in capital letters.

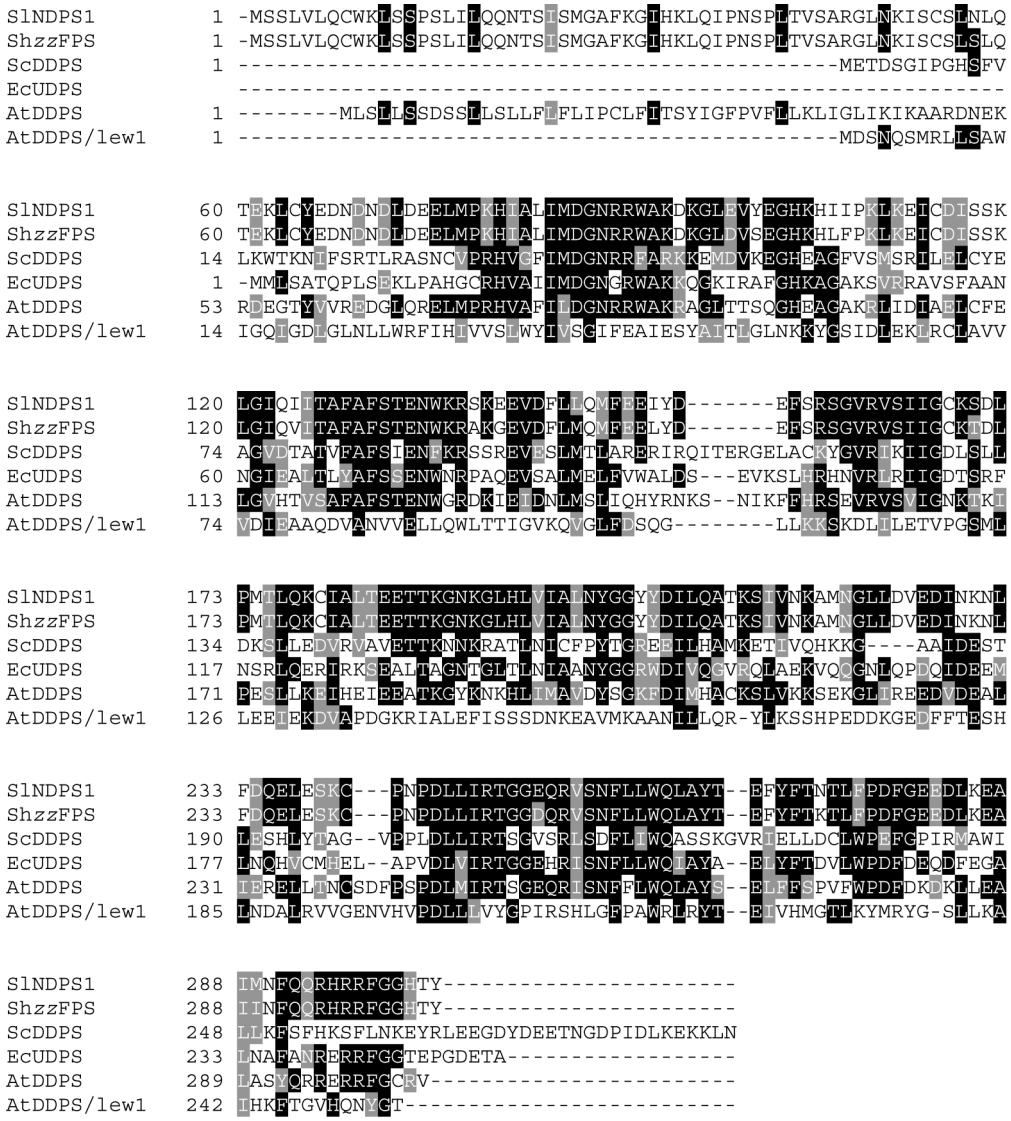

Fig. S2. Protein sequence alignment of NDPS1 and similar sequences. NDPS1 is most similar to *Z,Z*-farnesyl diphosphate synthase from *S. habrochaites* LA1777, and to *cis*-prenyltransferases from various species. Residues are shaded if identical (black) or similar (gray) in 50% or more of the sequences. SlNDPS1, *S. lycopersicum* neryl diphosphate synthase1 (FJ797956); ShzzFPS, *S. habrochaites* Z,Z-farnesyl diphosphate synthase (ACJ38408); ScDDPS, *S. cerevisiae* dolichol diphosphate synthase (BAA36577); EcUDPS, *E. coli* undecaprenyl diphosphate synthase (P60472); AtDDPS, *A. thaliana* dolichol diphosphate synthase (At2g23410); AtDDPS/lew1, *A. thaliana* dolichol diphosphate synthase (At1g11755).

ggaaaataggagaaatttcttcatttttgacctctccactcccaaaaacaacacacaatattcaaggATGatagttggct
atagaagcacaatcataacccttttctcatcctaagctaggcaatgggaaaacaatttcatccaatgcaattttccagaga
tcatgtagagtaagatgcagccacagtaccacttcatcaatgaatggtttcgaagatgcaagggatagaataagggaaag
ttttgggaaattagagttatctccttcttcctatgacacagcatgggtagctatggtcccttcaagacattcactaaatg
agccatgttttccacaatgtttggattggattattgaaaatcaaagagaagatggatcttggggactaaaccctacccat
ccattgcttctaaaggactcactttcttccactcttgcatgtttgcttgcactaaccaaatggagagttggagatgagca
aatcaaaagaggtcttggcttcattgaaacgtatggttgggcagtagataacaaggatcaaatttcacctttaggatttg
aagttatattttctagtatgatcaaatctgcagagaaattagatttaaatttgcctttgaatcttcatcttgtaaatttg
gtgaaatgcaaaagagattcaacaattaaaaggaatgttgaatatatgggtgaaggagttggtgaattatgtgattggaa
ggaaatgataaagttacatcaaagacaaaatggttcattatttgattcaccagccactactgcagctgccttgatttatc
atcaacatgatcaaaaatgctatcaatatcttaattcaatcttccaacaacacaaaaattgggttcccactatgtatcca
acaaaggtacattcattgctttgcttggttgatacacttcaaaatcttggagtacatcggcattttaaatcagaaataaa
gaaagctctagatgaaatatacaggctatggcaacaaaagaatgaacaaattttctcaaatgtcacccattgtgctatgg
cttttcgacttctaaggatgagctactatgatgtctcctcagatgaactagcagaatttgtggatgaagaacatttcttt
gcaacaaatgggaaatataaaagtcatgttgaaattcttgaactccacaaagcatcacaattggctattgatcatgagaa
agatgacattttggataaaataaacaattggacaagagcttttatggagcaaaaactcttaaacaatggcttcatagata
ggatgtcaaagaaagaggtggaacttgctttgaggaagttttataccacatctcatctagcagaaaatagaagatatata
aagtcatacgaagagaacaatttttaaaatcttaaaagcagcttataggtcacccaacattaacaataaggacttgttagc
attttcaatacacgactttgaattatgccaagctcaacaccgagaagaacttcaacaactcaagaggtggtttgaagatt
atagattggaccaactcggacttgcagaacgatatatacatgctagttacttatttggtgttactgttatccccgagcct
gaattatccgatgctcgcctcatgtacgcgaaatacgtcatgctcctgactattgtcgatgatcatttcgagagttttgc
atctaaagatgaatgtttcaacatcattgaattagtagaaaggtgggatgactatgcaagtgtaggttataaatctgaga
aggttaaagtttttttttctgttttctataaatcaatagaggagcttgcaacaattgctgaaattaaacaaggacgatcc
gtcaaaaatcaccttattaatttgtggcttgaattgatgaagttgatgttgatggagcgagtagagtggtgttctggcaa
gacaataccaagcatagaagagtacttgtatgttacatctataacattttgtgcaaaattgattcctctctcaacacaat
attttcttggaataaaaatatccaaagatctactagaaagtgatgaaatatgtggcctatggaattgtagcggtagagtg
atgcgaatccttaatgatttacaagattccaagagagaacaaaaggaggtctcaataaatttagtcacattactaatgaa
aagtatgtctgaggaagaagctataatgaagataaaggaaatcttggaaatgaatagaagagagttattgaaaatggttt
tagttcaaaaaaagggaagccaattgcctcaattatgcaaagatatattttggaggacaagcaaatgggctcatttcact
tattcacaaactgatggatatagaattgcagaggaaatgaagaatcacattgatgaagtcttttacaaaccactcaatca
tTAAtcccttattttgaatttatgacttttgttctttccttctatattccaactaataaacttgttcgcgcgaatccaga
taataaaaaacctctataaacctgtggtttaatcttttatctaataccttaatattataaaaatatatatatatctcttt
atttttttgttcaaaaatataatacaa

Fig. S3. *PHS1* cDNA sequence (accession no. FJ797957). The initiating ATG codon and the stop codon of the reading frame are in capital letters.

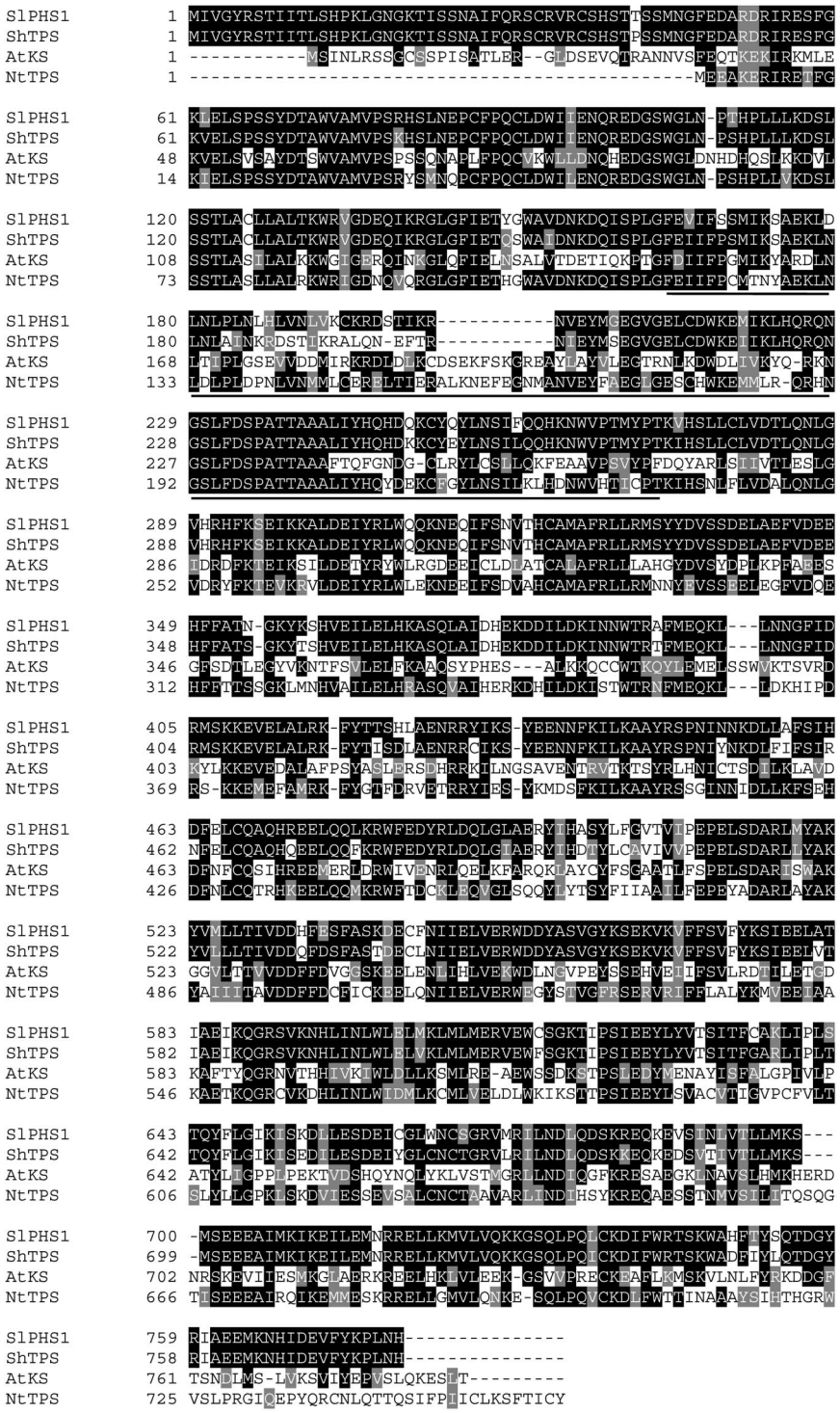

Fig. S4. Protein sequence alignment of PHS1 and similar sequences. PHS1 is most similar to a sesquiterpene synthase from *S. habrochaites* LA1777 that produces santalene and bergamotene isomers, as well as an unknown tobacco TPS, and an Arabidopsis diterpene synthase (kaurene synthase). Residues are shaded if identical (black) or similar (gray) in 50% or more of the sequences. The underlined sequence is part of an additional ancestral internal element generally found in diterpene synthases. SlPHS1, *S. lycopersicum* phellandrene synthase1 (FJ797957); ShTPS, *S. habrochaites* santalene/bergamotene synthase (ACJ38409); AtKS, *A. thaliana* kaurene synthase (AF034774); NtTPS, *N. tabacum* terpene synthase (AY528645).

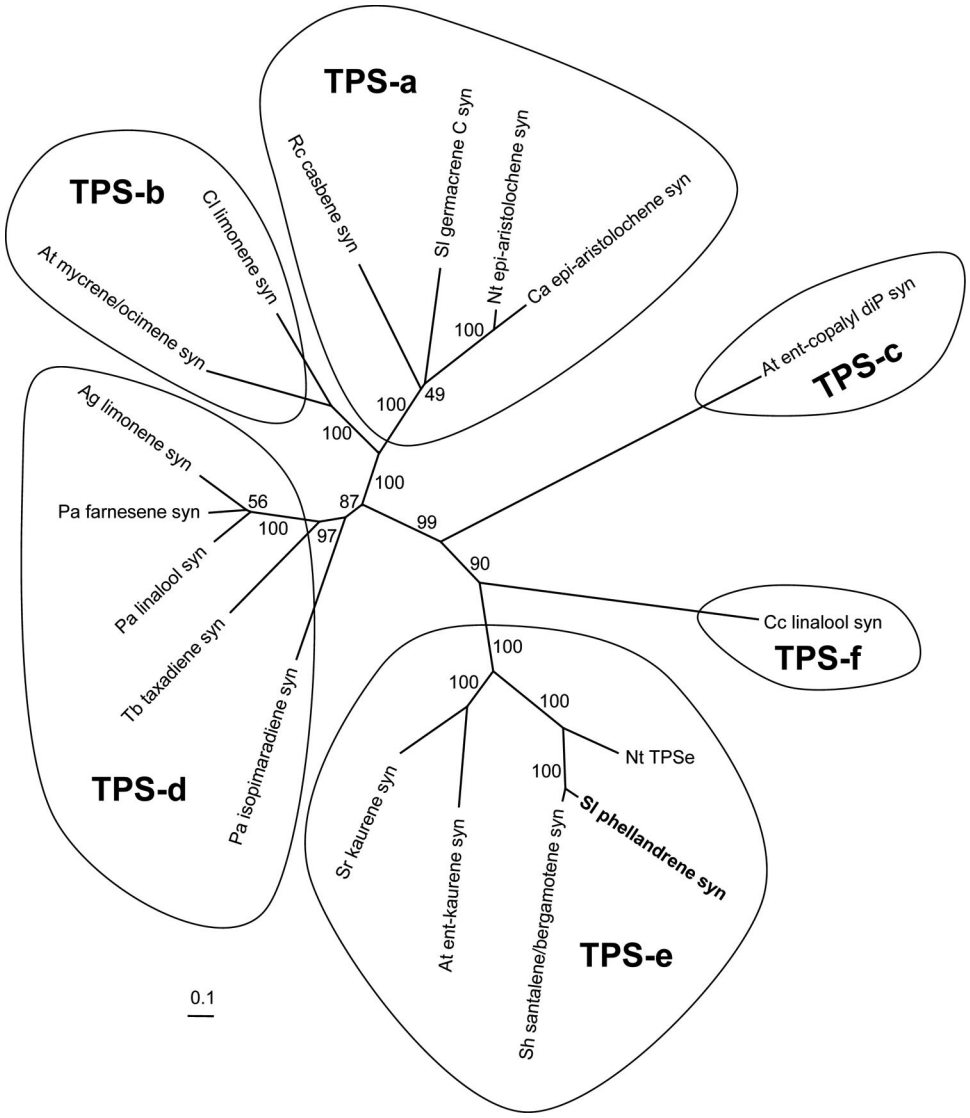

Fig. S5. Unrooted neighbor-joining phylogeny showing the position of PHS1 in the TPS-e subfamily. TPS protein sequences were aligned with ClustalX and the phylogeny was inferred by using PHYLIP (http://evolution.genetics.washington.edu/phylip.html). Sl phellandrene syn (FJ797957), Sh santalene/bergamotene syn (ACJ38409), Nt TPSe (AAS98912), At ent-kaurene syn (AAC39443), Sr kaurene syn (AAD34294), Cc linalool syn (AAD19839), At ent-copalyl diP syn (NP_192187), Ca epi-aristolochene syn (CAA06614), Nt epi-aristolochene syn (AAA19216), Sl germacrene syn (AF035631), Rc casbene syn (P59287), Cl limonene syn (AAM53944), At myrcene/ocimene syn (AAG09310), Ag limonene syn (AAB70907), Pa farnesene syn (AAS47697), Pa linalool syn (AAS47693), Tb taxadiene syn (AAK83586), Pa isopimaradiene syn (AAS47690).

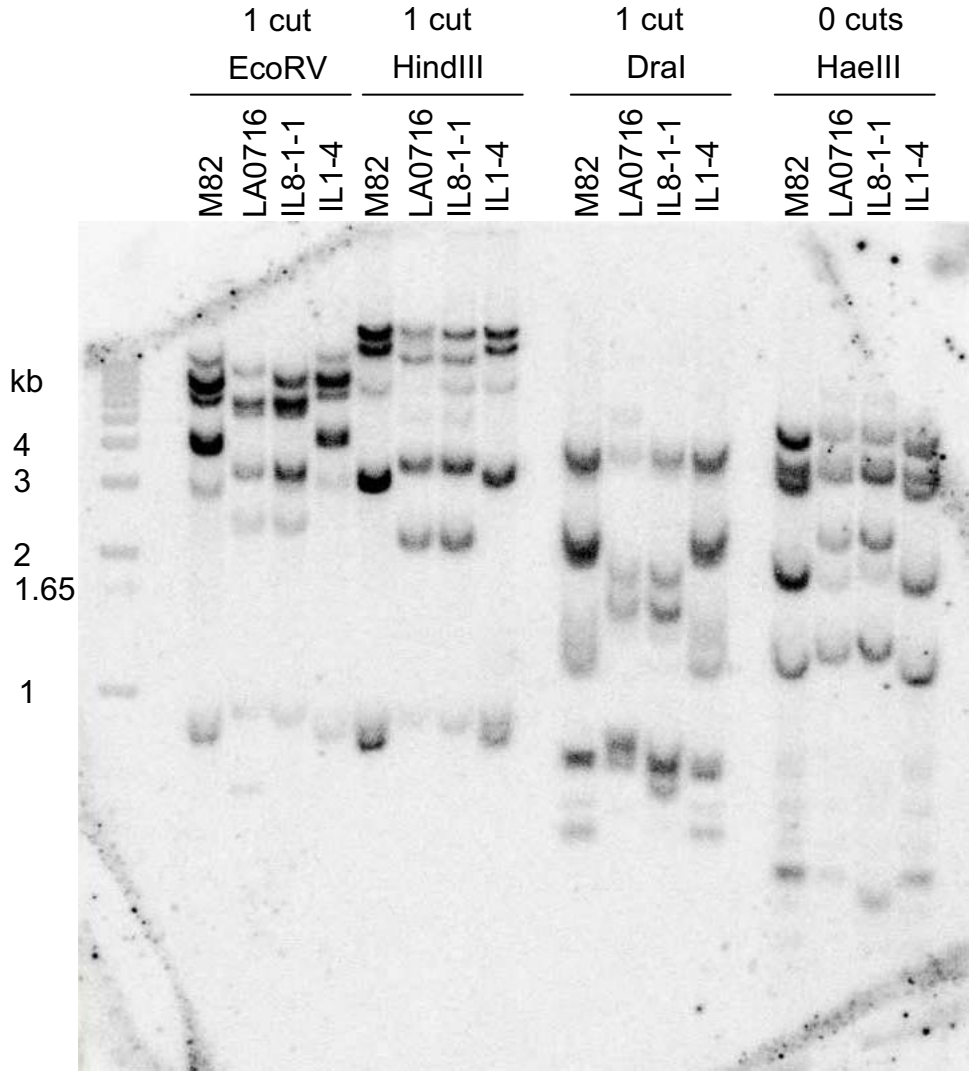

Fig. S6. Mapping of the *NDPS1* locus by restriction fragment length polymorphisms. Genomic DNA from *S. lycopersicum* cv M82, *S. pennellIii* LA0716, IL8-1-1, and IL1-4 was digested with the indicated restriction enzymes, separated on 1% agarose gel, blotted, and the blot probed with labeled full length *DNPS1* cDNA. The *NDPS1* gene has no introns, and the number of cuts that occur in the coding regions by each enzyme is indicated above the name of the enzyme. IL8-1-1 shows the same restriction fragment pattern as LA0716 indicating that *NDPS1* is located on the top of chromosome 8.

```
          1                                                                                 80
    M82   GAAATCTTGG AAATGAATAG AAGAGAGTTA TTGAAAATGG TTTTAGTTCA AAAAAAGGGA AGCCAATTGC CTCAATTATG
  IL1-4   GAAATCTTGG AAATGAATAG AAGAGAGTTA TTGAAAATGG TTTTAGTTCA AAAAAAGGGA AGCCAATTGC CTCAATTATG
 LA0716   GAAATCTTGG AAATGAATAG AAGAGAGTTA TTGAAAATGG TTTTAGTTCA AAAAAAGGGA AGCCAATTGC CTCAATTATG
IL8-1-1   GAAATCTTGG AAATGAATAG AAGAGAGTTA TTGAAAATGG TTTTAGTTCA AAAAAAGGGA AGCCAATTGC CTCAATTATG

          81                                                                              160
    M82   CAAAGATATA TTTTGGAGGA CAAGCAAATG GGCTCATTTC ACTTATTCAC AAACTGATGG ATATAGAATT GCAGAGGAAA
  IL1-4   CAAAGATATA TTTTGGAGGA CAAGCAAATG GGCTCATTTC ACTTATTCAC AAACTGATGG ATATAGAATT GCAGAGGAAA
 LA0716   CAAAGATATA TTTTGGAGGA CAAGCAAATG GACTCATTTC ACTTATTCAC AAACTGATGG ATTTAGAATT GAAGAGGAAA
IL8-1-1   CAAAGATATA TTTTGGAGGA CAAGCAAATG GACTCATTTC ACTTATTCAC AAACTGATGG ATTTAGAATT GAAGAGGAAA

          161                                                                        234
    M82   TGAAGAATCA CATTGATGAA GTCTTTTACA AACCACTCAA TCATTAATCC CTTATTTTGA ATTTATGACT TTTG
  IL1-4   TGAAGAATCA CATTGATGAA GTCTTTTACA AACCACTCAA TCATTAATCC CTTATTTTGA ATTTATGACT TTTG
 LA0716   TGAAGAATCA CATTGATGAA GTCTTTTACA AACCACTCAA TCATTAATCC CTCATTTTGA ATTTATGACT TTTG
IL8-1-1   TGAAGAATCA CATTGATGAA GTCTTTTACA AACCACTCAA TCATTAATCC CTCATTTTGA ATTTATGACT TTTG
```

Fig. S7. Mapping of the *PHS1* locus by PCR of the 3′ end of the gene. A 234-nucleotide fragment of the *PHS1* gene was amplified from *S. lycopersicum* cv M82, *S.pennellii* LA0716, IL1-4, and IL8-1-1 genomic DNA, and the PCR products sequenced. Nucleotide positions in which the same nucleotide occurred in all four PCR products are shown in red. Positions in which nucleotides are not identical in all sequences are show in black or blue. The sequences of M82 and IL1-4 are identical to each othe but different from those of LA0716 and IL8-1-1, and the sequences of LA0716 and IL8-1-1 are identical to each other but different from those of M82 and IL1-4, indicating that *PHS1* is located on the top of chromosome 8.

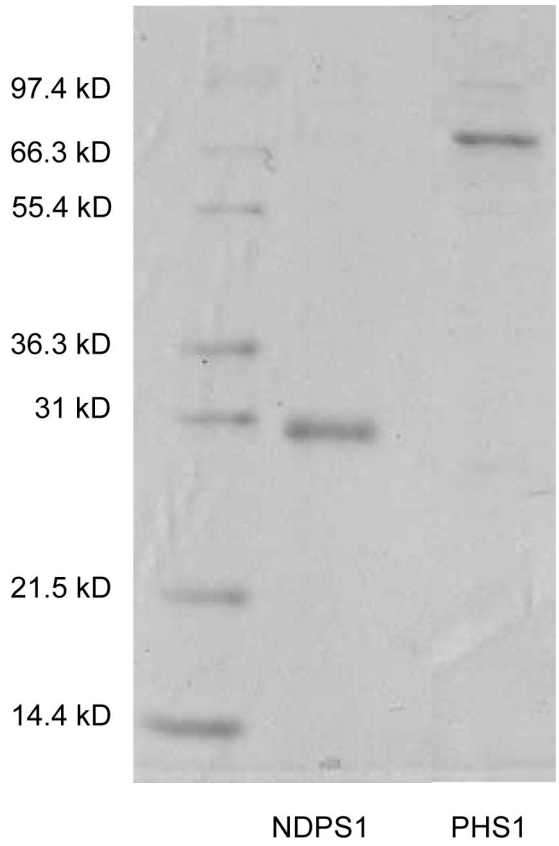

Fig. S8. SDS/PAGE analysis of purified recombinant His-tagged NDPS1 and PHS1.

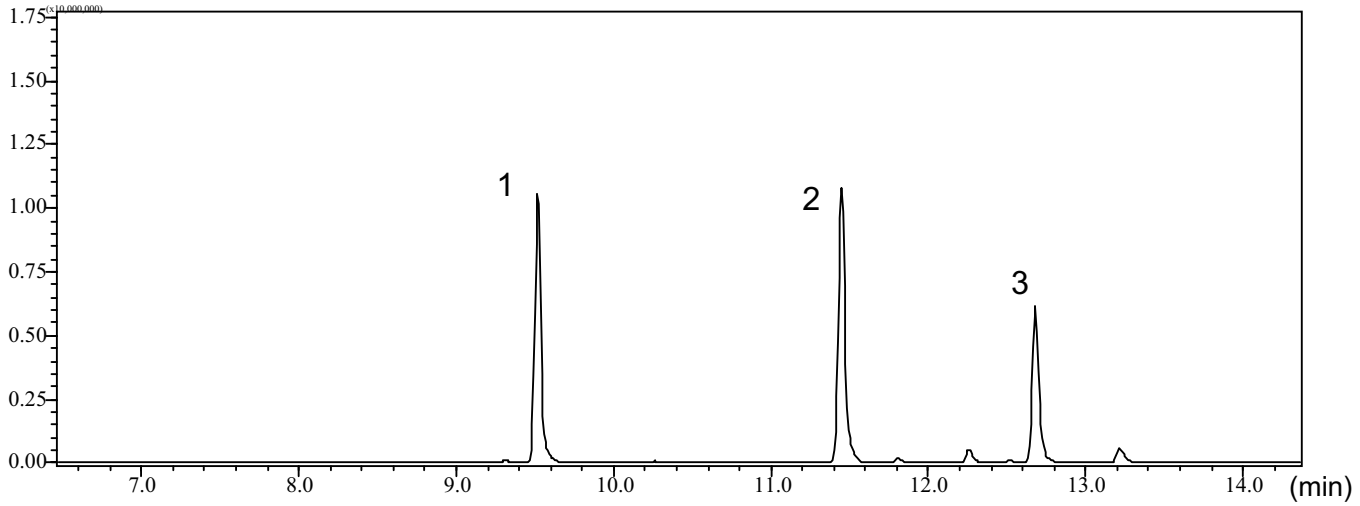

Fig. S9. A gas chromatograph showing the peaks obtained after acid hydrolysis of the product of the IPP and DMAPP condensation reaction catalyzed by NDPS. 1, linalool; 2, α-terpineol; 3, nerol. Peaks were identified by MS and comparisons to the retention times of authentic standards.

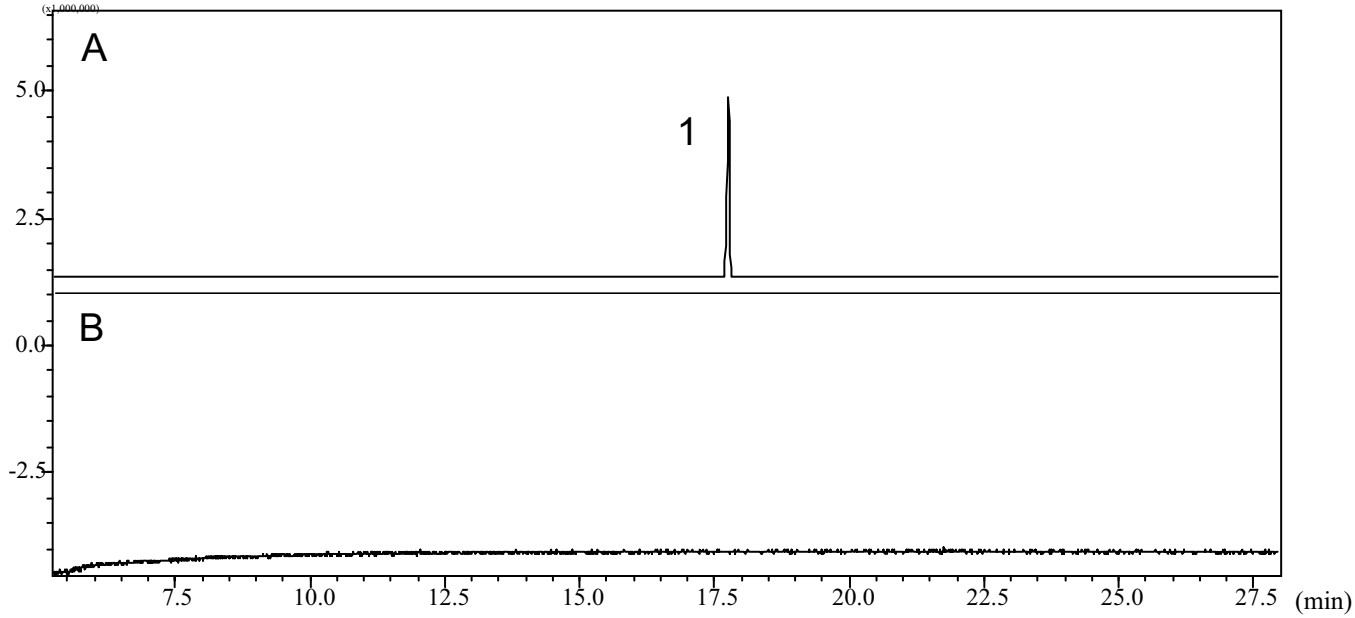

Fig. S10. Gas chromatographs of the products obtained in reactions catalyzed by MTPS1 with GPP (*A*) and NPP (*B*) as substrates. Peak 1 in *A* is linalool, identified by MS and comparison with the retention time of an authentic linalool standard. No products were obtained in *B*.

Table S1. Terpenes found in leaves briefly dipped in MTBE and in isolated glands of tomato accessions

		Leaf dips in MTBE, μg/gdw				Type VI glands, ng/200 glands			
Peak	Compounds	M82	IL1-4	IL8-1-1	LA0716	M82	IL1-4	IL8-1-1	LA0716
1	δ-2-Carene	130.0 (0.3)*	186.9 (12.2)	n.d.†	n.d.	97.1 (2.0)	112.7 (26.0)	n.d.	n.d.
2	α-Phellandrene	74.8 (1.4)	103.4 (11.7)	956.6 (5.9)	254.5 (132.0)	58.4 (1.3)	52.4 (11.7)	578.9 (201.0)	114.7 (4.3)
3	α-Terpinene	14.1 (1.0)	20.2 (2.3)	39.2 (4.5)	11.4 (6.6)	12.8 (0.7)	12.1 (1.5)	23.7 (10.2)	5.3 (1.3)
4	Limonene	202.2 (3.9)	261.0 (25.2)	52.8 (3.6)	n.d.	134.2 (4.5)	138.2 (22.4)	37.9 (9.8)	n.d.
5	β-Phellandrene	987.4 (19.3)	1274.2 (123)	78.1 (5.5)	25.9 (11.4)	655.0 (21.9)	674.7 (109)	56.9 (14.8)	19.0 (3.1)
6	γ-Terpinene	3.2 (0.2)	5.1 (0.3)	273.3 (0.9)	81.5 (39.5)	n.d.	n.d.	195.6 (62.4)	35.8 (5.3)
7	δ-Elemene	41.0 (0.6)	89.5 (1.7)	34.0 (4.1)	n.d.	141.4 (25)	173.2 (16.7)	97.0 (17.6)	n.d.
8	Caryophyllene	26.1 (0.4)	46.9 (0.8)	27.1 (3.1)	n.d.	65.6 (6.8)	52.4 (11.8)	69.2 (13.2)	n.d.
9	α-Humulene	18.7 (2.0)	30.8 (0.3)	20.5 (1.1)	n.d.	51.5 (1.2)	44.7 (4.6)	44.6 (5.7)	n.d.

*Standard deviation.
†n.d., not detected.