Current Biology

Convergence and molecular evolution of floral fragrance after independent transitions to self-fertilization

Graphical abstract
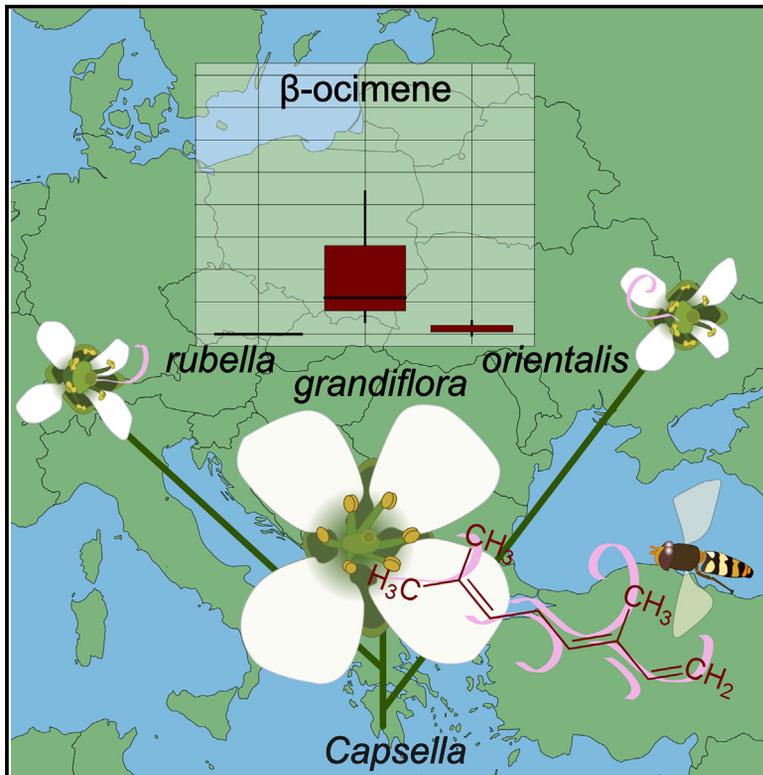
Authors

Natalia Joanna Woźniak, Kevin Sartori, Christian Kappel, ..., Michael Lenhard, Joachim Kopka, Adrien Sicard

Correspondence

adrien.sicard@slu.se

In brief

Flower scent serves as chemical signals to attract pollinators in many plants. Wozniak et al. describe a very similar evolution of floral volatile emission after independent transitions to autogamy in the genus *Capsella*, emphasizing its ecological importance even when animal pollination is not required.

Highlights

- Flower scent evolved similarly in the independent selfing lineages of *Capsella*

- *β*-ocimene emission is highly reduced in both lineages

- This reduction is attributed to distinct mutations in each selfing species

- In *C. rubella*, it is linked to a decrease in chloroplastic levels of TPS02

CellPress

Report

Convergence and molecular evolution of floral fragrance after independent transitions to self-fertilization

Natalia Joanna Woźniak,[1] Kevin Sartori,[2] Christian Kappel,[1] Thi Chi Tran,[1] Lihua Zhao,[2] Alexander Erban,[3] Jannicke Gallinger,[4] Ines Fehrle,[3] Friederike Jantzen,[1] Marion Orsucci,[2] Velemir Ninkovic,[4] Stefanie Rosa,[2] Michael Lenhard,[1] Joachim Kopka,[3] and Adrien Sicard[2,5,*]

[1]Institut für Biochemie und Biologie, Universität Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam-Golm, Germany
[2]Department of Plant Biology, Swedish University of Agricultural Sciences, Uppsala BioCenter, 75007 Uppsala, Sweden
[3]Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, 14476 Potsdam, Germany
[4]Department of Ecology, Swedish University of Agricultural Sciences, 750 07 Uppsala, Sweden
[5]Lead contact
*Correspondence: adrien.sicard@slu.se
https://doi.org/10.1016/j.cub.2024.04.063

SUMMARY

Studying the independent evolution of similar traits provides valuable insights into the ecological and genetic factors driving phenotypic evolution.[1] The transition from outcrossing to self-fertilization is common in plant evolution[2] and is often associated with a reduction in floral attractive features such as display size, chemical signals, and pollinator rewards.[3] These changes are believed to result from the reallocation of the resources used for building attractive flowers, as the need to attract pollinators decreases.[2,3] We investigated the similarities in the evolution of flower fragrance following independent transitions to self-fertilization in *Capsella*.[4–9] We identified several compounds that exhibited similar changes in different selfer lineages, such that the flower scent composition reflects mating systems rather than evolutionary history within this genus. We further demonstrate that the repeated loss of β-ocimene emission, one of the compounds most strongly affected by these transitions, was caused by mutations in different genes. In one of the *Capsella* selfing lineages, the loss of its emission was associated with a mutation altering subcellular localization of the ortholog of TERPENE SYNTHASE 2. This mutation appears to have been fixed early after the transition to selfing through the capture of variants segregating in the ancestral outcrossing population. The large extent of convergence in the independent evolution of flower scent, together with the evolutionary history and molecular consequences of a causal mutation, suggests that the emission of specific volatiles evolved as a response to changes in ecological pressures rather than resource limitation.

RESULTS AND DISCUSSION

Independent transitions to selfing in the genus *Capsella* are accompanied by convergent changes in floral fragrance

The flowers of many outcrossing species emit a complex mixture of volatile chemical compounds that are believed to increase pollinator visitation rates.[3,10,11] Although transitions to selfing were associated with a decreased emission of particular compounds,[12–17] the extent to which independent selfer lineages evolved similar floral fragrances remains unknown. The genus *Capsella* provides a genetically tractable model to study the independent evolution of selfing syndrome traits.[18] In this genus, the two selfer lineages, *C. rubella* (*Cr*)[4–6] and *C. orientalis* (*Co*),[7,8] respectively, evolved from the independent breakdown of the self-incompatibility system in an eastern and a western lineage of the outcrossing *C. grandiflora* (*Cg*)-like ancestors.[8] *Cg* is visited by a wide range of insect species, including wild bees,

honey bees, wasps, and hoverflies, some of which are known to rely on olfactory clues to locate nutritious flowers.[19–21] Modern *Cg* individuals descend from the western lineage, while "eastern" outcrossing populations are believed to be extinct. Consistently, clustering these species based on genetic distance places *Cr* in a sister clade to *Cg*, while *Co* branches as an outgroup (Figure 1A).

In contrast to *Cr* and *Co*, the outcrossing species *Cg* emits high levels of benzaldehyde, indicating that flower fragrances have changed in both selfers.[12,13] To determine the extent of similarity between these changes, we analyzed the floral fragrance of 5 representative accessions for each of the three *Capsella* species. Total emissions did not differ consistently between mating types, indicating that the transition to selfing is not systematically associated with a global reduction of floral scent (Figure S1). We, therefore, characterized flower scent composition by calculating dissimilarity indices between individuals based on volatile compound proportions (Figures 1B and 1C). The first dimension of a non-metric multidimensional scaling
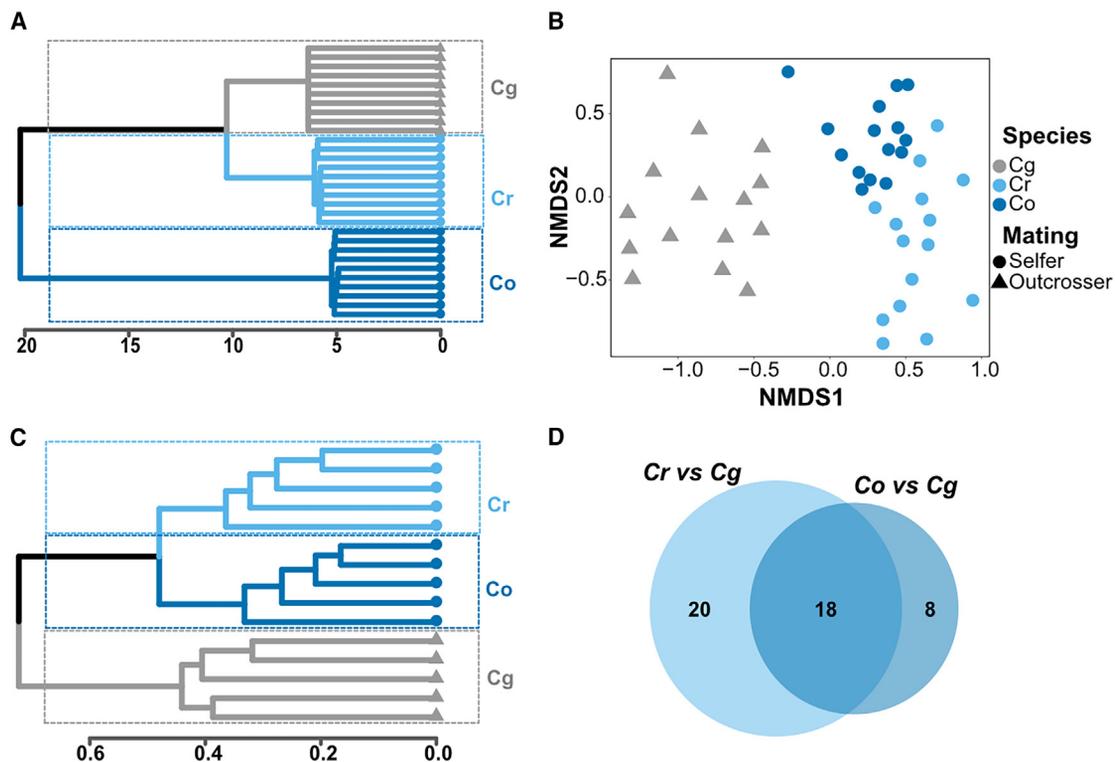
Figure 1. Convergent evolution of flower scent after independent transitions to selfing in the genus *Capsella*

(A) Hierarchical clustering based on genetic distances using the Ward clustering algorithm on hamming distances between 10 *Cg*, *Cr*, and *Co* populations. The scale bars indicate log(hamming distances).

(B) Non-metric multidimensional scaling (NMSD) plot based on Bray-Curtis dissimilarities in flower scent composition between samples (*n* = 15 per species). SI and SC refer to self-incompatible and self-compatible, respectively.

(C) Clustering cladogram based on average Bray-Curtis dissimilarities calculated from volatile compound proportions. The scale bars indicate Bray-Curtis dissimilarities.

(D) Venn diagram illustrating the number of compounds significantly changed only in *Cr*, only in *Co*, or in both *Co* and *Cr* compared with *Cg*.

See also Figures S1 and S2 and Table S1.

(NMDS) of these dissimilarity indices discriminated the samples according to their mating systems, with *Co* and *Cr* displaying higher NMDS1 values than *Cg* (Figures 1 and S1). In contrast to the genetic relationships (Figure 1A), clustering the different samples based on the flower scent composition grouped *Cr* and *Co* as sister clades, while *Cg* branched as an outgroup (Figure 1C). Accordingly, we observed that the mating system explains most (about 40%) of the variation in flower scent composition (PERMANOVA; $F = 43.128$, $p < 0.001$). We next performed a random forest analysis to identify flower volatile emission profiles characteristic of the individuals' mating systems. The resulting classification model correctly assigned all populations to the right category with a probability above 75% (Figures S1F and S1G). To determine the contribution of each volatile compound to this classification, we calculated the mean decrease in model accuracy when the different volatiles were removed from the dataset (Figure S1H). Nineteen compounds were particularly important for the classification, with (E)-β-ocimene, (Z)-β-ocimene, and benzenacetaldehyde causing the largest drop in accuracy. Overall, 38 and 26 compounds varied significantly between *Cg* and *Cr* or *Cg* and *Co*, respectively (Figures 1D and S2; Table S1). Among those, 18 compounds, comprising the most

important ones for discriminating the mating systems, changed in the same direction in both selfers ($p = 0.01$; hypergeometric probability) (Figure 1D).

These results indicated a convergent evolution of volatile production after independent transitions to selfing in the genus *Capsella*. The similarities are such that it is possible to classify individuals into different mating types solely based on the floral fragrance composition. Most of the compounds recurrently lost in selfers are common volatiles in floral bouquets believed to promote pollinator attraction.[10,22–27] In particular, β-ocimene has been proposed to act as a generalist pollinator attractant and has been shown to increase pollinator visit rates.[28–33] The proportion of dimethyl trisulfide and dimethyl disulfide also increased in both selfers. These compounds have an unpleasant sulfurous odor attractive to silphid beetles and some dipterans,[34–37] but they also act as repellents for other insect species[38] and have neurotoxic effects on generalist insects.[39] The nature of these compounds and the lack of consistent decrease in flower volatile emission suggest that the transitions to selfing are associated with a shift in flower blend function toward decreasing pollinator attraction and possibly increasing defenses against pests.
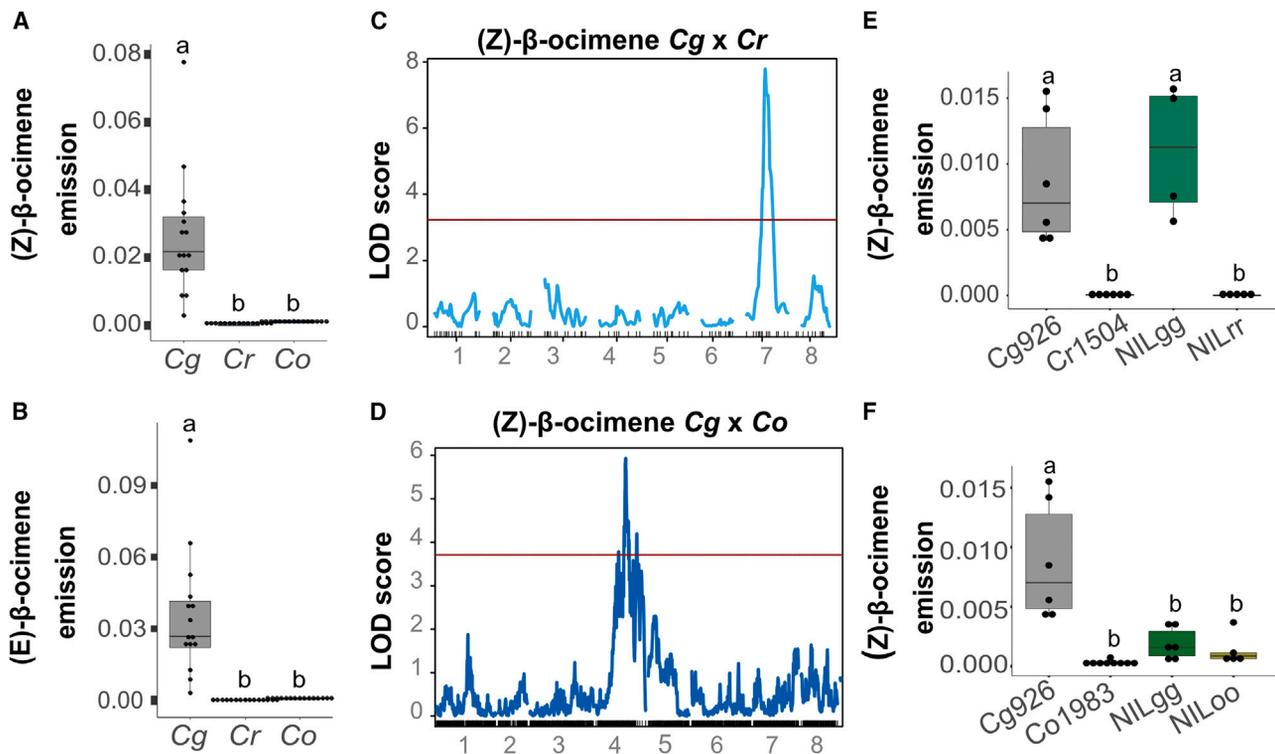
Figure 2. Mutations at different loci led to the loss of β-ocimene in *C. rubella* and *C. orientalis*

(A and B) The abundance of (Z)-β-ocimene and (E)-β-ocimene relative to total volatile emission in each *Capsella* species (n = 15 for each species).

(C and D) QTL mapping for the emission of (Z)-β-ocimene (Tag_mass 93) in the *Cr1504 × Cg926* RIL population (C) and *Co1983 ×* Cg926 F2 population (D). Red dashed lines indicate the genome-wide 5% significance threshold.

(E and F) (Z)-β-ocimene emission normalized to total volatile emission in the headspace of NILs segregating for the *Cg926* and *Cr1504* QTL_1 alleles (E) or the *Cg926* and *Co1983* QTL_1 alleles (F).

Letters indicate significant differences as determined by one-way ANOVA with Tukey's HSD test (α = 0.05). Boxplots' lower and upper hinges correspond to the first and third quartiles, respectively. The upper and lower whiskers add or subtract 1.5 interquartile ranges to/from the 75 and 25 percentile, respectively. The middle lines represent the median.

See also Figure S3 and Table S2.

Mutations at different loci underlie the loss of β-ocimene emission after the transition to selfing in *Capsella*

Globally, *Cr* shows greater changes in volatile emission than *Co* (Figure S2), suggesting that the molecular processes underlying these changes differ between the two selfer lineages. To test this, we mapped quantitative trait loci (QTLs) influencing the emission of β-ocimene, one of the compounds with the largest contribution to the discrimination between mating categories. (E)-β-ocimene and (Z)-β-ocimene were emitted by all *Cg* samples but detected in very low proportion or not at all in selfers (Figures 2A and 2B). Both compounds are known to be produced through the same biosynthetic pathway[30] and showed highly correlated emission levels in selfer × outcrosser populations (Pearson R^2 = 0.9936 in *Cr × Cg* and 0.9807 in *Co × Cg*). This suggests that the genetic change(s) in either selfer lineage decrease the emission of both isomers in the same manner. We, thus, focused further analyses on (Z)-β-ocimene, using it as a proxy for overall changes in β-ocimene emission. We detected one significant QTL affecting (Z)-β-ocimene emission in each population, but while it was located on chromosome 7 in the *Cr × Cg* recombinant inbred lines (RILs) (later referred to

as QTL_1), it mapped to chromosome 4 in the *Co × Cg* F2 population (QTL_2), suggesting that different loci underlie the loss of scent in the two selfing lineages (Figures 2C and 2D; Table S2). Flower size did not significantly affect (Z)-β-ocimene emission in any of these populations (linear regression analysis; $p > 0.08$ with R^2 = 0.04093 in *Cr × Cg* RIL and $p > 0.12$ with R^2 = 0.01571 in *Co × Cg* F2). Additionally, none of the QTLs identified overlap with any known flower size QTLs,[18,40] indicating that flower size is not the main factor influencing volatile emission.

To ascertain that different loci underlie the loss of β-ocimene emission in the two *Capsella* selfing species, we formally tested whether allele variation in the QTL_1 influences (Z)-β-ocimene emission in both selfers. To this end, we generated near-isogenic lines (NILs) segregating for either the *Cr* and *Cg* alleles (*NILrg*) or the *Co* and *Cg* alleles (*NILog*) in a chromosomal region spanning QTL_1 but otherwise homozygous for the selfer alleles throughout most of the genome. We next selected plants homozygous for alternative alleles and compared their scent emission profile (Figures 2E and 2F). The *NILrg* progeny individuals homozygous for the *Cg* allele (*NILgg*) emit (Z)-β-ocimene to a similar level as wild-type *Cg* plants. As for *Cr*, no (Z)-β-ocimene could be detected in *NILrg* progeny homozygous for the *Cr* allele

(*NILrr*). This result confirmed the major influence of QTL_1 on volatile emission and further demonstrated that allelic variation at this locus can fully explain the loss of (Z)-β-ocimene emission in *Cr* (Figure 2E). By contrast, both plants homozygous for the *Cg* allele (*NILgg*) and those homozygous for *Co* (*NILoo*) derived from *NILog* emitted (Z)-β-ocimene at a detectable and similar level, further suggesting that QTL_1 does not influence (Z)-β-ocimene emission in *Co* (Figure 2F). Therefore, at least in the case of (Z)-β-ocimene, the parallel reduction of the same scent compound is not caused by mutations at the same locus in the two selfers.

Allelic variation at TPS02 underlies the loss of β-ocimene in *C. rubella*

To better understand the molecular history of flower fragrance evolution after the transition to selfing, we sought to identify the gene underlying the loss of (Z)-β-ocimene in *Cr*. To this end, we screened about 1,200 *NILrg* progeny plants for ones harboring a recombination event within the QTL_1 region. Comparing the level of (Z)-β-ocimene emission with the position of the recombination breakpoints allowed us to narrow down the causal mutations to an interval of 115 kb between positions 9,330,402 and 9,445,345 on chromosome 7 (Figure 3A). Although all individuals heterozygous within this interval emitted a detectable amount of (Z)-β-ocimene, this compound could not be detected in any individuals homozygous for the *Cr* allele. Within this region, we identified two genes coding for terpene synthases (TPSs), syntenic to *Arabidopsis thaliana TPS02* and *TPS03*, with an average sequence similarity of 85% and 84% between the proteins of the two species. TPS02 was previously shown to contain a chloroplast transit peptide (cTP) that regulates its subcellular localization and allows it to act as β-ocimene synthase in *A. thaliana*.[41] *In silico* predictions of target peptides identified a cTP in both *Cr* TPS02 (TPS02r) and *Cg* TPS02 (TPS02g) in an extra 35 (36 for TPS02g) amino acid N-terminal domain that is absent in any of the *TPS03* alleles (Table S3). Consistently, *Cg* TPS02 but not *Cg* TPS03 could rescue the absence of β-ocimene emission in the *A. thaliana* accession Col-0, which has been shown to be caused by a frameshift mutation in *TPS02*[41] (Figure S3A). We, thus, hypothesized that allelic variation at *TPS02* underlies the loss of (Z)-β-ocimene emission in *Cr*. To test this, we compared the ability of the genomic fragments encoding for *Cg* TPS02 (TPS02g) or *Cr* TPS02 (TPS02r) to restore (Z)-β-ocimene in the *NILrr*. Although (Z)-β-ocimene could not be detected in any of the *TPS02r* lines tested, all the lines transformed with *TPS02g* emitted a detectable amount of this compound (Figure 3B). Together, these results indicated that allelic variation in *TPS02* is responsible for the change in (Z)-β-ocimene emission between *Cr* and *Cg* and that introducing *CgTPS02* into *Cr* fully rescued the emission of this compound.

Inefficient targeting of TPS02 to the chloroplast is associated with the loss of β-ocimene in *C. rubella*

TPS02 is expressed at similar levels in the inflorescences of both *NILgg* and *NILrr* plants and did not show any differences in splicing patterns (Figures S3B and S3C), suggesting that the loss of (Z)-β-ocimene emission in *Cr* was caused by mutations affecting the TPS02 amino acid sequence. The comparison between TPS02 alleles segregating in our mapping population (i.e.,

a cross between *Cr1504* and *Cg926*) identified six non-synonymous mutations, including five amino acid exchanges and one nucleotide triplet deletion (Figure 3C). To pinpoint which of these amino acids abolished (Z)-β-ocimene emission, we correlated the genotype at the candidate non-synonymous polymorphisms with (Z)-β-ocimene emission in five *Cg* and five *Cr* plants from different accessions (Figure 3C). (Z)-β-ocimene was emitted at various levels in all *Cg* plants but was not detected in any *Cr* accessions. Only the polymorphism at position 9,384,490 (8th amino acid) fully correlates with the production of (Z)-β-ocimene in our assay, making it the prime candidate to underlie the loss of (Z)-β-ocimene. The corresponding single-nucleotide polymorphism (SNP) induces a serine-to-proline conversion, which corresponds to a polar-to-non-polar amino acid exchange within the cTP. Both TPS02g and TPS02r demonstrated the ability to use the precursor of β-ocimene, geranyl pyrophosphate (GPP), as substrate in *in vitro* assays without significant kinetic differences (Figure 3D). TPS02r tends to have a lower affinity for this substrate compared with TPS02g (Km = 50.71 and 19.99 μM, respectively; Wilcoxon rank sum test p = 0.1), but the catalytic activities of both enzymes were comparable. To assess whether the cellular localization of TPS02 is affected by the polymorphism at position 9,384,490, we tagged TPS02 with yellow fluorescent protein (YFP) and estimated the efficiency of chloroplast targeting by determining the above-threshold correlation coefficient between the YFP and the chlorophyll A (ChlA) fluorescent signals (Figure 3E). This analysis indicated that a significantly larger proportion of the TPS02g is localized within the chloroplast. We next mutated nucleotide 9,384,490 in TPS02g to introduce a proline while keeping the rest of TPS02g protein unchanged. This point mutation decreased the amount of TPS02g in the chloroplast, suggesting that SNP 9,384,490 impaired the efficiency of TPS02 targeting to the chloroplast (Figure 3E). Considering that the level of chloroplastic TPS02 influences (Z)-β-ocimene production in *A. thaliana*,[41] it is likely that the decrease in chloroplastic TPS02 in *Cr* contributed to reduce β-ocimene biosynthesis, thereby preventing its detection within the floral blend. In the cytoplasm of *A. thaliana*, TPS02 contributes to synthesizing the sesquiterpene (E,E)-α-farnesene,[41] a compound activated upon herbivore attack as an indirect defense mechanism attracting natural pests' enemies.[42–45] TPS02g and TPS02r also demonstrated the ability to use the precursor of sesquiterpenes, farnesyl pyrophosphate (FPP), as a substrate in *in vitro* assays with similar catalytic efficiencies (Figure S3D). This suggests that the increase of cytoplasmic TPS02 could reinforce sesquiterpene production in *Cr*. However, we were not able to detect (E, E)-α-farnesene in any of our samples, suggesting that this compound is not strongly emitted by any *Capsella* flowers. Since the biosynthesis of α-farnesene can be induced by herbivory in other species,[41] it remains possible that change in *TPS02* localization increases this compound emission in response to tissue damage. The NILs segregating for TPS02 *Cg* and *Cr* alleles will provide a unique resource to test this hypothesis in the future.

Evolutionary history of the loss of β-ocimene

To gain insights into the evolutionary origin of the loss of β-ocimene emission, we used publicly available genome sequences
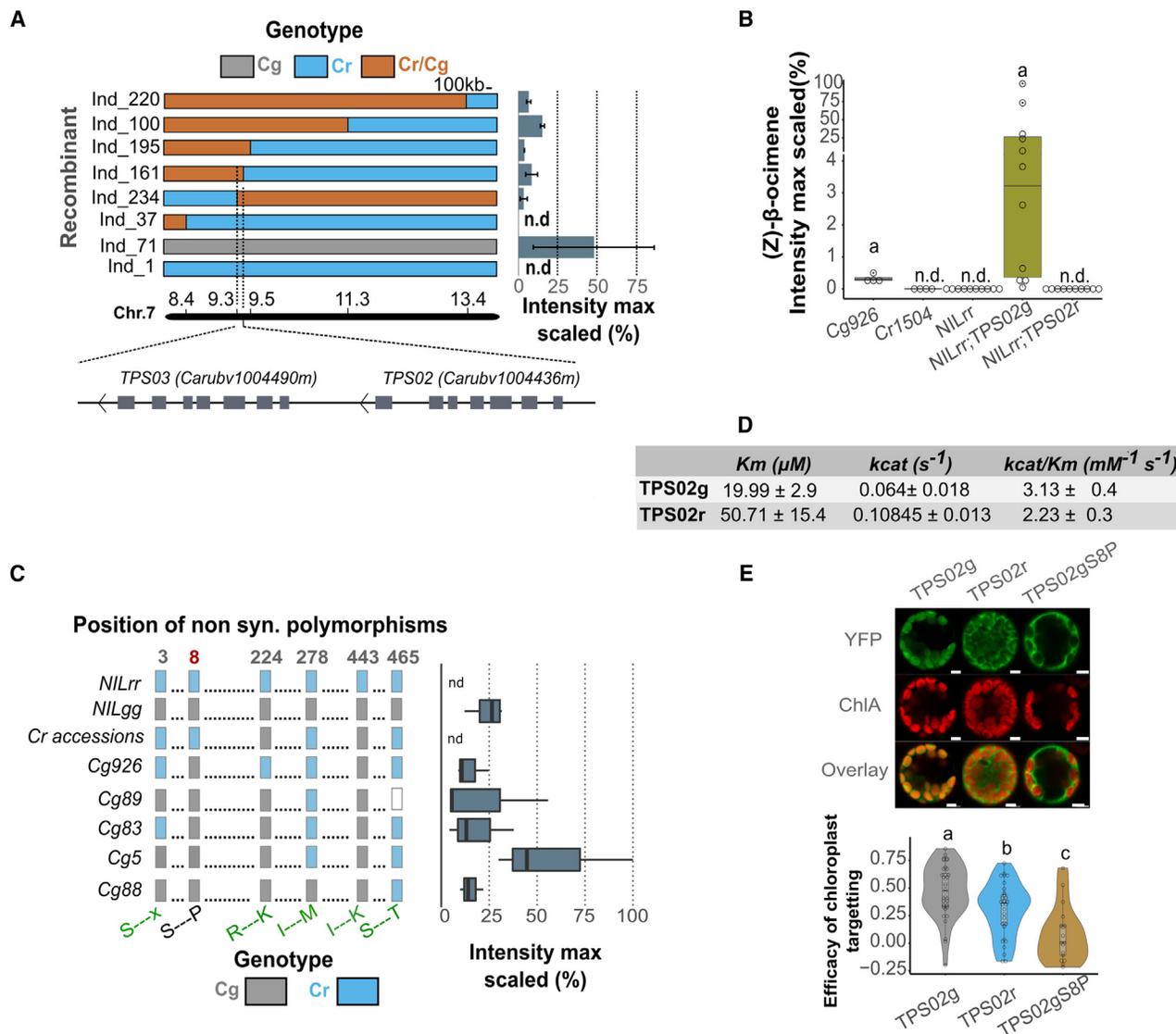
Figure 3. A mutation in the chloroplast transit peptide of TPS02 underlies the loss of β-ocimene emission in *C. rubella*

(A) Fine-mapping of QTL_1. The color-coded bars on the left indicate the informative recombinants' genotype. Positions on chromosome 7 are indicated in Mb at the bottom. The bar chart on the right indicates (Z)-β-ocimene emission intensity scaled to the maximal value in the experiment. Values are mean ± SD from 2 measurements. The schematic representation of genes involved in monoterpene biosynthesis in the mapping interval shown at the bottom.

(B) (Z)-β-ocimene emission in *Cg926*, *Cr1504, NILrr;TPS02g*, and *NILrr;TPS02r*. Each dot represents an individual measurement.

(C) (Z)-β-ocimene (Tag_mass 93) emission in different *Cg* and *Cr* accessions. The genotype at each non-synonymous SNP is represented on the right by color-coded rectangles. Boxplots on the right indicate the intensity of the (Z)-β-ocimene emitted by each individual scaled to the maximal value in the experiment. The SNP showing the best correlation with (Z)-β-ocimene emission is indicated in red.

(D) Kinetic parameters of TPS02r and TPS02g using geranyl pyrophosphate (GPP), as substrate. Each value corresponds to the average ± SE of three replicates.

(E) Cellular localization of TPS02r, TPS02g, or TPS02gS8P in *Arabidopsis* leaf protoplasts (top). YFP corresponds to TPS02, and the chlorophyll A (ChlA) signal indicates the position of the chloroplasts. Scale bars correspond to 5 μm. The bottom panel shows a violin plot representing the efficiency of TPS02r, TPS02g, and TPS02gS8P chloroplast targeting (above-threshold Pearson correlation coefficient between the YFP and ChlA signals). Each dot represents an individual measurement.

Letters indicate significant differences as determined by one-way ANOVA with Tukey's HSD test (α = 0.05). Boxplots' lower and upper hinges correspond to the first and third quartiles, respectively. The upper and lower whiskers add or subtract 1.5 interquartile ranges to/from the 75 and 25 percentile, respectively. The middle lines represent the median.

See also Figure S3 and Table S3.

to reconstruct the TPS02 haplotypes from 182 *Cg*, 16 *Co*, and 51 *Cr* accessions. We first investigated the genetic structure at the *TPS02* locus within *Cr* by inferring individuals' genetic ancestries using ADMIXTURE (Figures 4A and S3). The best fit of individuals'

ancestries was obtained for six clusters, with two of these clusters including 37 out of 51 individuals. Hierarchical clustering of the *Cr* haplotypes uncover two major clades regrouping together 48 *Cr* individuals, later referred to as the *Crub_183* and *Cr1504*
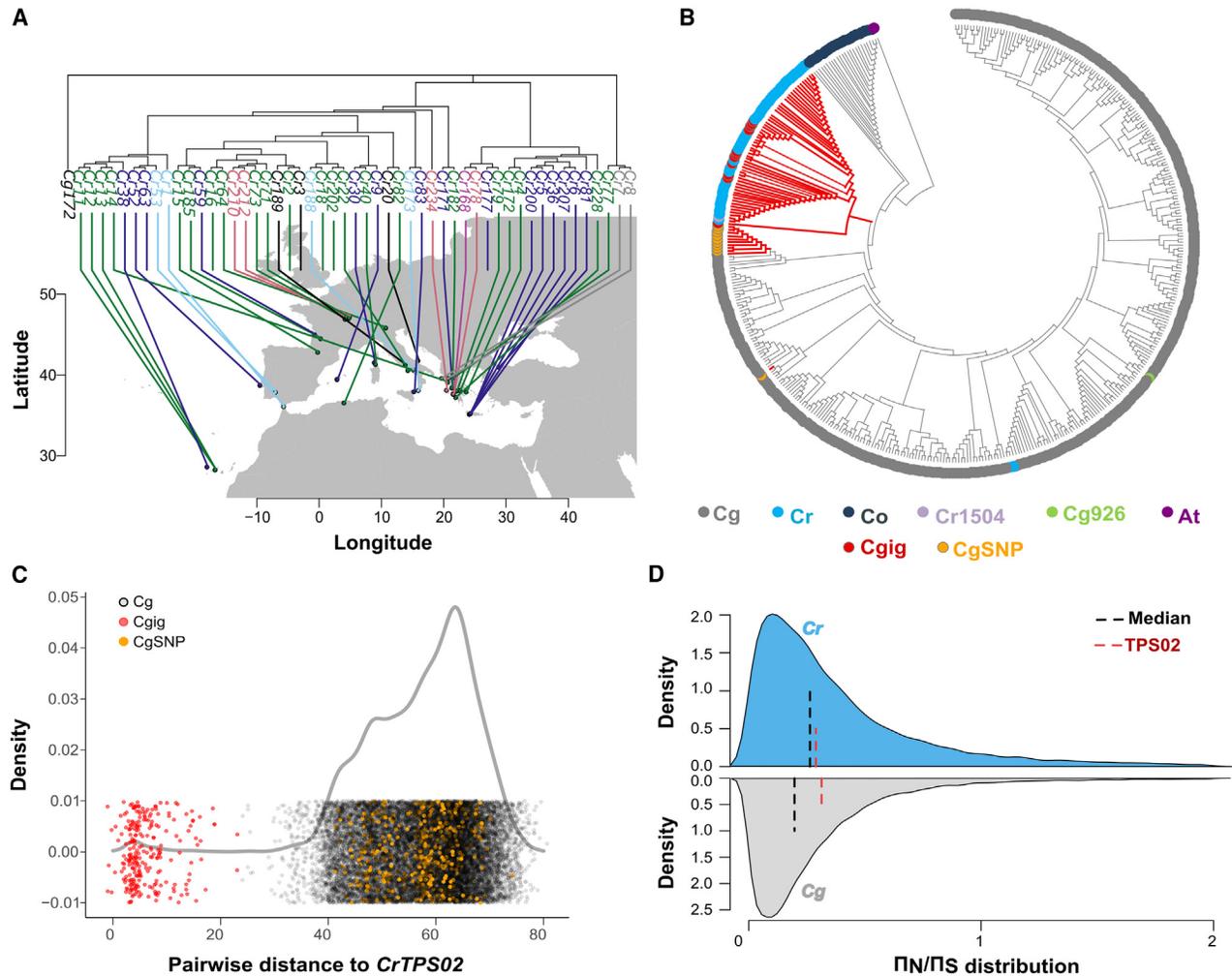
Figure 4. Evolutionary history of TPS02 in *Capsella*

(A) Geographical distribution of the different *Cr* TPS02 haplotype groups, as determined by an admixture analysis (Figure S3). The tree on the top illustrates the phylogenetic relationship among corresponding *Cr* accessions.

(B) Neighbor-joining tree representing the genetic relationships between all *Capsella* TPS02 haplotypes identified. The bootstrap consensus tree was inferred from 1,000 replicates. Branches corresponding to partitions reproduced in less than 50% of bootstrap replicates are collapsed. Red branches illustrate haplotypes containing the candidate causal mutation at SNP 9.384.490. *Cg* haplotypes containing the *Cr* allele at this position are marked by a red dot if they show signs of recent introgression (Cgig) or by an orange dot if they do not show any signs of recent introgression (CgSNP).

(C) Distribution of the pairwise genetic distances, as numbers of polymorphisms, between *Cg* and *Cr* TPS02 haplotypes. Values corresponding to the distance between *Cgig* and *Cr* haplotypes are shown in red, while the values comparing *CgSNP* and *Cr* haplotypes are shown in orange.

(D) Π_N/Π_S distribution in *Cg* and *Cr*. Median and TPS02 Π_N/Π_S are indicated with a black and red dashed line, respectively.

See also Figure S4.

haplotype groups. The geographical distribution of these haplotypes does not reflect the genetic relationship between *Cr* individuals, suggesting that they were differentiated at the early stage of *Cr* history before its geographical spread (Figure 4A). To understand how these different clusters relate to the *TPS02g* haplotypes, we reconstituted a neighbor-joining tree using all *TPS02* assembled haplotypes (Figure 4B). The haplotypes corresponding to the 48 *Cr* individuals mentioned above cluster within a lineage also containing 11 haplotypes found in 8 *Cg* individuals. The latter show, nevertheless, excessively long *Cr*-like haplotypes and abnormally low genetic differentiation from *Cr* haplotypes, indicating that they reflect recent hybridization events with the

Cr haplotypes not yet broken up by recombination (Figures 4C and S4). The 3 *Cr* sequences not included in the main haplotypes cluster within *Cg* haplotypes. Their unusual genetic similarities to *Cg* sequences, particularly low levels of long *Crub_183* or *Cr1504* haplotypes (Figure S4), and the fact that the individuals harboring these haplotypes originate from Greece, where the two *Capsella* species remain in sympatry, also suggests that these haplotypes correspond to post-divergence gene flow from *Cg* to *Cr*.

After excluding haplotypes with genomic signatures of recent introgression, we observed that the *Cr* allele at the candidate causal SNP (position 9,384,490) was found in all *Cr* from the Crub_183 and Cr1504 haplotype groups but also in 11 *Cg*

individuals (Figures 4 and S4). None of these *Cg* individuals shows signs of recent introgression. Their pairwise distances to *Cr* haplotype are dispersed through the center of the distribution of all *Cg* vs. *Cr* haplotypes comparison, and they show average levels of long *Cr*-like haplotypes. All other non-synonymous mutations within *TPS02* fixed within *Cr* were found to similarly segregate in *Cg* (Figure S4). Together, these observations support a scenario in which the *Cr* allele leading to the loss of β-ocimene has been captured from standing variation in the ancestral outcrossing population.

Selfing syndrome traits such as scent emission are often believed to evolve as a result of the relaxation of selective pressure acting on genes controlling pollinator attraction. To determine if the functional constraints acting on *TPS02* have indeed been relaxed in *Cr*, we computed the ratio between non-synonymous sites and synonymous sites' nucleotide polymorphism (Π_N/Π_S). Consistent with the expectation of the decreased efficiency of natural selection in selfers,[9] the genome-wide Π_N/Π_S distribution is shifted toward higher values in *Cr* (Figure 4D). However, TPS02 Π_N/Π_S did not change between the two species, suggesting that TPS02 is under similar levels of functional constraints in *Cr* and *Cg*. The loss of β-ocimene emission in *Cr* does not, therefore, seem to have been caused by the relaxation of purifying selection acting on the gene controlling its production.

In this study, we uncovered a number of volatiles whose emission is repeatedly affected by the transition toward self-fertilization, suggesting their potential importance in optimizing different reproduction modes. The mutation underlying the loss of emission of one of these compounds, (*Z*)-β-ocimene, in *Cr*, was found to affect the subcellular localization of an enzyme involved in the terminal step of monoterpene biosynthesis. Therefore, it seems unlikely that this mutation enables significant reallocation of resources to other fitness functions, as previously proposed for selfing syndrome traits.[46] Instead, *TPS02* remains under similar functional constraints in *Cr* as in the outbreeding ancestor, possibly reflecting its rechanneling toward other metabolic processes. Such mutations at the onset of self-fertilization could favor the establishment of selfer lineages by limiting conspecific gene flow through pollinator movements while potentially reinforcing other metabolic functions. Consistently, β-ocimene emission in *Capsella* appears to increase the visitation rate of several pollinator groups in nature.[21] This example demonstrates how deviations in phenotypic value associated with shifts in ecological constraints do not always associate with changes in the extent of selective pressure at underlying loci. Dissecting the biological functions of different convergent selfing syndrome traits will help in understanding the evolutionary and ecological pressures that shape flower phenotypes in self-fertilizing species.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability

- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Floral Volatile sampling in *Capsella*
  - Floral volatile sampling in *Arabidopsis thaliana*
  - VOCs Identification and quantification
  - Flower fragrance analysis
  - QTL mapping
  - Fine mapping
  - Molecular cloning and plant transformation
  - Gene expression and splicing analysis
  - TPS02 activity assay
  - Confocal imaging
  - Sequence and population-genetic analyses
- QUANTIFICATION AND STATISTICAL ANALYSIS

REFERENCES

1. Blount, Z.D., Lenski, R.E., and Losos, J.B. (2018). Contingency and determinism in evolution: replaying life's tape. Science *362*, eaam5979. https://doi.org/10.1126/science.aam5979.

2. Barrett, S.C.H. (2002). The evolution of plant sexual diversity. Nat. Rev. Genet. *3*, 274–284. https://doi.org/10.1038/nrg776.

3. Sicard, A., and Lenhard, M. (2011). The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. Ann. Bot. *107*, 1433–1443. https://doi.org/10.1093/aob/mcr023.

4. Guo, Y.-L., Bechsgaard, J.S., Slotte, T., Neuffer, B., Lascoux, M., Weigel, D., and Schierup, M.H. (2009). Recent speciation of Capsella rubella from Capsella grandiflora, associated with loss of self-incompatibility and an extreme bottleneck. Proc. Natl. Acad. Sci. USA *106*, 5246–5251. https://doi.org/10.1073/pnas.0808012106.

5. Foxe, J.P., Slotte, T., Stahl, E.A., Neuffer, B., Hurka, H., and Wright, S.I. (2009). Recent speciation associated with the evolution of selfing in

Capsella. Proc. Natl. Acad. Sci. USA *106*, 5241–5245. https://doi.org/10.1073/pnas.0807679106.

6. Koenig, D., Hagmann, J., Li, R., Bemm, F., Slotte, T., Neuffer, B., Wright, S.I., and Weigel, D. (2019). Long-term balancing selection drives evolution of immunity genes in Capsella. eLife *8*, 1–27. https://doi.org/10.7554/eLife.43606.

7. Bachmann, J.A., Tedder, A., Laenen, B., Fracassetti, M., Désamoré, A., Lafon-Placette, C., Steige, K.A., Callot, C., Marande, W., Neuffer, B., et al. (2018). Genetic basis and timing of a major mating system shift in Capsella. Preprint at bioRxiv. https://doi.org/10.1101/425389.

8. Hurka, H., Friesen, N., German, D.A., Franzke, A., and Neuffer, B. (2012). "Missing link" species Capsella orientalis and Capsella thracica elucidate evolution of model plant genus Capsella (Brassicaceae). Mol. Ecol. *21*, 1223–1238. https://doi.org/10.1111/j.1365-294X.2012.05460.x.

9. Slotte, T., Hazzouri, K.M., Ågren, J.A., Koenig, D., Maumus, F., Guo, Y.-L., Steige, K., Platts, A.E., Escobar, J.S., Newman, L.K., et al. (2013). The Capsella rubella genome and the genomic consequences of rapid mating system evolution. Nat. Genet. *45*, 831–835. https://doi.org/10.1038/ng.2669.

10. Knudsen, J.T., and Gershenzon, J. (2006). The chemical diversity of floral scent. In Biology of Floral Scent (Springer). https://doi.org/10.1201/9781420004007.ch2.

11. Schiestl, F.P., and Johnson, S.D. (2013). Pollinator-mediated evolution of floral signals. Trends Ecol. Evol. *28*, 307–315. https://doi.org/10.1016/j.tree.2013.01.019.

12. Sas, C., Müller, F., Kappel, C., Kent, T.V., Wright, S.I., Hilker, M., and Lenhard, M. (2016). Repeated inactivation of the first committed enzyme underlies the loss of benzaldehyde emission after the selfing transition in Capsella. Curr. Biol. *26*, 3313–3319. https://doi.org/10.1016/j.cub.2016.10.026.

13. Jantzen, F., Lynch, J.H., Kappel, C., Höfflin, J., Skaliter, O., Wozniak, N., Sicard, A., Sas, C., Adebesin, F., Ravid, J., et al. (2019). Retracing the molecular basis and evolutionary history of the loss of benzaldehyde emission in the genus Capsella. New Phytol. *224*, 1349–1360. https://doi.org/10.1111/nph.16103.

14. Raguso, R.A., Kelber, A., Pfaff, M., Levin, R.A., and McDade, L.A. (2007). Floral biology of North American Oenothera sect. Lavauxia (Onagraceae): advertisements, rewards, and extreme variation in floral depth. Ann. Mo. Bot. Gard. *94*, 236–257. https://doi.org/10.3417/0026-6493(2007)94[236:FBONAO]2.0.CO;2.

15. Majetic, C.J., Castilla, A.R., and Levin, D.A. (2019). Losing a Scent of one's Self: is there a Reduction in Floral Scent Emission in Self-Pollinating Phlox cuspidata versus Outcrossing Phlox drummondii? Int. J. Plant Sci. *180*, 86–92. https://doi.org/10.1086/701102.

16. Doubleday, L.A.D., Raguso, R.A., and Eckert, C.G. (2013). Dramatic vestigialization of floral fragrance across a transition from outcrossing to selfing in Abronia umbellata (Nyctaginaceae). Am. J. Bot. *100*, 2280–2292. https://doi.org/10.3732/ajb.1300159.

17. Petrén, H., Toräng, P., Ågren, J., and Friberg, M. (2021). Evolution of floral scent in relation to self-incompatibility and capacity for autonomous self-pollination in the perennial herb Arabis alpina. Ann. Bot. *127*, 737–747. https://doi.org/10.1093/aob/mcab007.

18. Woźniak, N.J., Kappel, C., Marona, C., Altschmied, L., Neuffer, B., and Sicard, A. (2020). A similar genetic architecture underlies the convergent evolution of the selfing syndrome in Capsella. Plant Cell *32*, 935–949. https://doi.org/10.1105/tpc.19.00551.

19. Primante, C., and Dötterl, S. (2010). A syrphid fly uses olfactory cues to find a non-yellow flower. J. Chem. Ecol. *36*, 1207–1210. https://doi.org/10.1007/s10886-010-9871-6.

20. Bisrat, D., and Jung, C. (2022). Roles of flower scent in bee?flower mediations: a review. J. Ecol. Environ. *46*, 3. https://doi.org/10.5141/jee.21.00075.

21. Xiong, Y.Z., Kappel, C., Hagemann, L., Jantzen, F., Wozniak, N., Sicard, A., Huang, S.Q., and Lenhard, M. (2023). Testing the effect of individual scent compounds on pollinator attraction in nature using quasi-isogenic Capsella lines. Am. J. Bot. *110*, e16237. https://doi.org/10.1002/ajb2.16237.

22. Knudsen, J.T., Eriksson, R., Gershenzon, J., and Ståhl, B. (2006). Diversity and distribution of floral scent. Bot. Rev. *72*, 1–120. https://doi.org/10.1663/0006-8101(2006)72[1:DADOFS]2.0.CO;2.

23. Knudsen, J.T., Tollsten, L., and Bergström, L.G. (1993). Floral scents-a checklist of volatile compounds isolated by head-space techniques. Phytochemistry *33*, 253–280. https://doi.org/10.1016/0031-9422(93)85502-I.

24. Knudsen, J.T., and Tollsten, L. (1993). Trends in floral scent chemnistry in pollination syndromnes: floral scent comnposition in moth- pollinated taxa. Bot. J. Linn. Society *113*, 263–284. https://doi.org/10.1111/j.1095-8339.1993.tb00340.x.

25. Knudsen, J.T., Tollsten, L., Groth, I., Bergström, G., and Raguso, R.A. (2004). Trends in floral scent chemistry in pollination syndromes: floral scent composition in hummingbird-pollinated taxa. Bot. J. Linn. Soc. *146*, 191–199. https://doi.org/10.1111/j.1095-8339.2004.00329.x.

26. Qian, C.Y., Quan, W.X., Xiang, Z.M., and Li, C.C. (2019). Characterization of volatile compounds in four different rhododendron flowers by GC_GC-QTOFMS. Molecules *24*, 3327. https://doi.org/10.3390/molecules24183327.

27. Weiss, J., Mühlemann, J.K., Ruiz-Hernández, V., Dudareva, N., and Egea-Cortines, M. (2016). Phenotypic space and variation of floral scent profiles during late flower development in Antirrhinum. Front. Plant Sci. *7*, 1903. https://doi.org/10.3389/fpls.2016.01903.

28. Farré-Armengol, G., Fernández-Martínez, M., Filella, I., Junker, R.R., and Peñuelas, J. (2020). Deciphering the biotic and climatic factors that influence floral scents: A systematic review of floral volatile emissions. Front. Plant Sci. *11*, 1154. https://doi.org/10.3389/fpls.2020.01154.

29. Filella, I., Primante, C., Llusià, J., Martín González, A.M., Seco, R., Farré-Armengol, G., Rodrigo, A., Bosch, J., and Peñuelas, J. (2013). Floral advertisement scent in a changing plant-pollinators market. Sci. Rep. *3*, 3434. https://doi.org/10.1038/srep03434.

30. Farré-Armengol, G., Filella, I., Llusià, J., and Peñuelas, J. (2017). β-ocimene, a Key Floral and Foliar Volatile Involved in Multiple Interactions between Plants and Other Organisms. Molecules *22*, 1148. https://doi.org/10.3390/molecules22071148.

31. Andersson, S., Nilsson, L.A., Groth, I., and Bergström, G. (2002). Floral scents in butterfly-pollinated plants: possible convergence in chemical composition. Bot. J. Linn. Soc. *140*, 129–153. https://doi.org/10.1046/j.1095-8339.2002.00068.x.

32. Peng, F., Byers, K.J.R.P., and Bradshaw, H.D. (2017). Less is more: independent loss-of-function OCIMENE synthase alleles parallel pollination syndrome diversification in monkeyflowers (Mimulus). Am. J. Bot. *104*, 1055–1059. https://doi.org/10.3732/ajb.1700104.

33. Byers, K.J.R.P., Vela, J.P., Peng, F., Riffell, J.A., and Bradshaw, H.D. (2014). Floral volatile alleles can contribute to pollinator-mediated reproductive isolation in monkeyflowers (Mimulus). Plant J. *80*, 1031–1042. https://doi.org/10.1111/tpj.12702.

34. Yan, G., Liu, S., Schlink, A.C., Flematti, G.R., Brodie, B.S., Bohman, B., Greeff, J.C., Vercoe, P.E., Hu, J., and Martin, G.B. (2018). Behavior and electrophysiological response of gravid and non-gravid Lucilia cuprina (Diptera: Calliphoridae) to carrion-associated compounds. J. Econ. Entomol. *111*, 1958–1965. https://doi.org/10.1093/jee/toy115.

35. Zito, P., Sajeva, M., Raspi, A., and Dötterl, S. (2014). Dimethyl disulfide and dimethyl trisulfide: so similar yet so different in evoking biological responses in saprophilous flies. Chemoecology *24*, 261–267. https://doi.org/10.1007/s00049-014-0169-y.

36. von Hoermann, C., Ruther, J., and Ayasse, M. (2016). Volatile organic compounds of decaying piglet cadavers perceived by Nicrophorus vespilloides. J. Chem. Ecol. *42*, 756–767. https://doi.org/10.1007/s10886-016-0719-6.

37. Podskalská, H., Růžička, J., Hoskovec, M., and Šálek, M. (2009). Use of infochemicals to attract carrion beetles into pitfall traps. Entomol. Exp. Appl. *132*, 59–64. https://doi.org/10.1111/j.1570-7458.2009.00871.x.

38. Ling, S., Qiu, H., Xu, J., Gu, Y., Yu, J., Wang, W., Liu, J., and Zeng, X. (2022). Volatile dimethyl disulfide from guava plants regulate developmental performance of Asian citrus psyllid through activation of defense responses in neighboring orange plants. Int. J. Mol. Sci. *23*, 10271. https://doi.org/10.3390/ijms231810271.

39. Dugravot, S., Grolleau, F., Macherel, D., Rochetaing, A., Hue, B., Stankiewicz, M., Huignard, J., and Lapied, B. (2003). Dimethyl disulfide exerts insecticidal neurotoxicity through mitochondrial dysfunction and activation of insect KATP channels. J. Neurophysiol. *90*, 259–270. https://doi.org/10.1152/jn.01096.2002.

40. Sicard, A., Stacey, N., Hermann, K., Dessoly, J., Neuffer, B., Bäurle, I., and Lenhard, M. (2011). Genetics, evolution, and adaptive significance of the selfing syndrome in the genus Capsella. Plant Cell *23*, 3156–3171. https://doi.org/10.1105/tpc.111.088237.

41. Huang, M., Abel, C., Sohrabi, R., Petri, J., Haupt, I., Cosimano, J., Gershenzon, J., and Tholl, D. (2010). Variation of herbivore-induced volatile terpenes among Arabidopsis ecotypes depends on allelic differences and subcellular targeting of two terpene synthases, TPS02 and TPS03. Plant Physiol. *153*, 1293–1310. https://doi.org/10.1104/pp.110.154864.

42. Boevé, J.L., Lengwiler, U., Tollsten, L., Dorn, S., and Turlings, T.C.J. (1996). Volatiles emitted by apple fruitlets infested by larvae of the European apple sawfly. Phytochemistry *42*, 373–381. https://doi.org/10.1016/0031-9422(95)00948-5.

43. Scutareanu, P., Bruin, J., Posthumus, M.A., and Drukker, B. (2003). Constitutive and herbivore-induced volatiles in pear, alder and hawthorn trees. Chemoecology *13*, 63–74. https://doi.org/10.1007/s00049-002-0228-7.

44. Lin, J., Wang, D., Chen, X., Köllner, T.G., Mazarei, M., Guo, H., Pantalone, V.R., Arelli, P., Stewart, C.N., Wang, N., et al. (2017). An (E,E)-α-farnesene synthase gene of soybean has a role in defence against nematodes and is involved in synthesizing insect-induced volatiles. Plant Biotechnol. J. *15*, 510–519. https://doi.org/10.1111/pbi.12649.

45. Danner, H., Boeckler, G.A., Irmisch, S., Yuan, J.S., Chen, F., Gershenzon, J., Unsicker, S.B., and Köllner, T.G. (2011). Four terpene synthases produce major compounds of the gypsy moth feeding-induced volatile blend of Populus trichocarpa. Phytochemistry *72*, 897–908. https://doi.org/10.1016/j.phytochem.2011.03.014.

46. Brunet, J. (1992). Sex allocation in hermaphroditic plants. Trends Ecol. Evol. *7*, 79–84. https://doi.org/10.1016/0169-5347(92)90245-7.

47. Rohloff, J., and Bones, A.M. (2005). Volatile profiling of Arabidopsis thaliana – putative olfactory compounds in plant communication. Phytochemistry *66*, 1941–1955. https://doi.org/10.1016/j.phytochem.2005.06.021.

48. Zorrilla-Fontanesi, Y., Rambla, J.-L., Cabeza, A., Medina, J.J., Sánchez-Sevilla, J.F., Valpuesta, V., Botella, M.A., Granell, A., and Amaya, I. (2012). Genetic analysis of strawberry fruit aroma and identification of O-Methyltransferase FaOMT as the locus controlling natural variation in Mesifurane content. Plant Physiol. *159*, 851–870. https://doi.org/10.1104/pp.111.188318.

49. Rambla, J.L., López-Gresa, M.P., Bellés, J.M., and Granell, A. (2015). Metabolomic profiling of plant tissues. Methods Mol. Biol. *1284*, 221–235. https://doi.org/10.1007/978-1-4939-2444-8_11.

50. Josephs, E.B., Lee, Y.W., Stinchcombe, J.R., and Wright, S.I. (2015). Association mapping reveals the role of purifying selection in the maintenance of genomic variation in gene expression. Proc. Natl. Acad. Sci. USA *112*, 15390–15395. https://doi.org/10.1073/pnas.1503027112.

51. Lommen, A. (2009). Metalign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. Anal. Chem. *81*, 3079–3086. https://doi.org/10.1021/ac900036d.

52. Luedemann, A., Strassburg, K., Erban, A., and Kopka, J. (2008). TagFinder for the quantitative analysis of gas chromatography - mass spectrometry

53. Allwood, J.W., Erban, A., de Koning, S., Dunn, W.B., Luedemann, A., Lommen, A., Kay, L., Löscher, R., Kopka, J., and Goodacre, R. (2009). Inter-laboratory reproducibility of fast gas chromatography-electron impact-time of flight mass spectrometry (GC-EI-TOF/MS) based plant metabolomics. Metabolomics *5*, 479–496. https://doi.org/10.1007/s11306-009-0169-z.

54. Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. Trends Genet. *16*, 276–277. https://doi.org/10.1016/s0168-9525(00)02024-2.

55. Brown, N.P., Leroy, C., and Sander, C. (1998). MView: A web-compatible database search or multiple alignment viewer. Bioinformatics *14*, 380–381. https://doi.org/10.1093/bioinformatics/14.4.380.

56. Sperschneider, J., Catanzariti, A.-M., DeBoer, K., Petre, B., Gardiner, D.M., Singh, K.B., Dodds, P.N., and Taylor, J.M. (2017). LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. Sci. Rep. *7*, 44598. https://doi.org/10.1038/srep44598.

57. Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis, version 11. Mol. Biol. Evol. *38*, 3022–3027. https://doi.org/10.1093/molbev/msab120.

58. Marçais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics *27*, 764–770. https://doi.org/10.1093/bioinformatics/btr011.

59. Lunter, G., and Goodson, M. (2011). Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res. *21*, 936–939. https://doi.org/10.1101/gr.111120.110.

60. Van der Auwera, G.A., and O'Connor, D. (2020). Genomics in the Cloud: Using Docker, GATK, and WDL in Terra, First Edition (O'Reilly Media).

61. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience *4*, 7. https://doi.org/10.1186/s13742-015-0047-8.

62. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. *32*, 268–274. https://doi.org/10.1093/molbev/msu300.

63. Flamini, G., Cioni, P.L., and Morelli, I. (2002). Differences in the fragrances of pollen and different floral parts of male and female flowers of Laurus nobilis. J. Agric. Food Chem. *50*, 4647–4652. https://doi.org/10.1021/jf020269x.

64. Vallarino, J.G., Erban, A., Fehrle, I., Fernie, A.R., Kopka, J., and Osorio, S. (2018). Acquisition of volatile compounds by gas chromatography–mass spectrometry (GC-MS). In Methods in Molecular Biology (Humana Press), pp. 225–239. https://doi.org/10.1007/978-1-4939-7819-9_16.

65. Gouinguené, S.P., and Turlings, T.C.J. (2002). The effects of abiotic factors on induced volatile emissions in corn plants. Plant Physiol. *129*, 1296–1307. https://doi.org/10.1104/pp.001941.

66. Boachon, B., Junker, R.R., Miesch, L., Bassard, J.-E., Höfer, R., Caillieaudeaux, R., Seidel, D.E., Lesot, A., Heinrich, C., Ginglinger, J.-F., et al. (2015). CYP76C1 (cytochrome P450)-mediated linalool metabolism and the formation of volatile and soluble linalool oxides in Arabidopsis Flowers: A strategy for defense against floral antagonists. Plant Cell *27*, 2972–2990. https://doi.org/10.1105/tpc.15.00399.

67. Gross, J., Gallinger, J., and Rid, M. (2019). Collection, identification, and statistical analysis of volatile organic compound patterns emitted by Phytoplasma infected plants. Methods Mol. Biol. *1875*, 333–343. https://doi.org/10.1007/978-1-4939-8837-2_25.

68. Luedemann, A., von Malotky, L., Erban, A., and Kopka, J. (2011). TagFinder: preprocessing software for the fingerprinting and the profiling of gas chromatography–mass spectrometry based metabolome analyses. Methods Mol. Biol. *860*, 255–286. https://doi.org/10.1007/978-1-61779-594-7_16.

(GC-MS)-based metabolite profiling experiments. Bioinformatics *24*, 732–737. https://doi.org/10.1093/bioinformatics/btn023.

69. Vandendool, H., and Kratz, P.D. (1963). A generalization of the retention index system including linear temperature programmed gas—liquid partition chromatography. J. Chromatogr. *11*, 463–471. https://doi.org/10.1016/S0021-9673(01)80947-X.

70. Ranganathan, Y., and Borges, R.M. (2010). Reducing the babel in plant volatile communication: using the forest to see the trees. Plant Biol. (Stuttg) *12*, 735–742. https://doi.org/10.1111/j.1438-8677.2009.00278.x.

71. Broman, K.W., Wu, H., Sen, S., and Churchill, G.A. (2003). R/qtl: QTL mapping in experimental crosses. Bioinformatics *19*, 889–890. https://doi.org/10.1093/bioinformatics/btg112.

72. Williams, D.C., McGarvey, D.J., Katahira, E.J., and Croteau, R. (1998). Truncation of limonene synthase preprotein provides a fully active 'pseudomature'' form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. Biochemistry *37*, 12213–12220. https://doi.org/10.1021/bi980854k.

73. Becker, D., Kemper, E., Schell, J., and Masterson, R. (1992). New plant binary vectors with selectable markers located proximal to the left T-DNA border. Plant Mol. Biol. *20*, 1195–1197. https://doi.org/10.1007/BF00028908.

74. Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. *16*, 735–743. https://doi.org/10.1046/j.1365-313x.1998.00343.x.

75. Wu, F.H., Shen, S.C., Lee, L.Y., Lee, S.H., Chan, M.T., and Lin, C.S. (2009). Tape-Arabidopsis sandwich - A simpler Arabidopsis protoplast isolation method. Plant Methods *5*, 16. https://doi.org/10.1186/1746-4811-5-16.

76. Vardakou, M., Salmon, M., Faraldos, J.A., and O'Maille, P.E. (2014). Comparative analysis and validation of the malachite green assay for the high throughput biochemical characterization of terpene synthases. MethodsX *1*, 187–196. https://doi.org/10.1016/j.mex.2014.08.007.

77. R Core Team (2018). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

78. Ågren, J.Å., Wang, W., Koenig, D., Neuffer, B., Weigel, D., and Wright, S.I. (2014). Mating system shifts and transposable element evolution in the plant genus Capsella. BMC Genomics *15*, 602. https://doi.org/10.1186/1471-2164-15-602.

79. Sicard, A., Kappel, C., Lee, Y.W., Woźniak, N.J., Marona, C., Stinchcombe, J.R., Wright, S.I., and Lenhard, M. (2016). Standing genetic variation in a tissue-specific enhancer underlies selfing-syndrome evolution in Capsella. Proc. Natl. Acad. Sci. USA *113*, 13911–13916. https://doi.org/10.1073/pnas.1613394113.

80. Luu, K., Bazin, E., and Blum, M.G.B. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. Mol. Ecol. Resour. *17*, 67–77. https://doi.org/10.1111/1755-0998.12592.

81. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109.

82. Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics *26*, 419–420. https://doi.org/10.1093/bioinformatics/btp696.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli BL21(DE3)pLysS	Promega Biotech	Cat# L1195
Biological samples		
Capsella rubella	List given in Table S1	N/A
Capsella grandiflora	List given in Table S1	N/A
Capsella orientalis	List given in Table S1	N/A
Cg926 x *Cr1504* RILs	Sicard et al.[40]	N/A
Cg926 x *Cr1983* F2	Woźniak et al.[18]	N/A
NILog	This paper	N/A
NILrg	This paper	N/A
NILrr; TPS02g	This paper	N/A
NILrr; TPS02r	This paper	N/A
Col-0,TPS02g	This paper	N/A
Col-0,TPS03g	This paper	N/A
Chemicals, peptides, and recombinant proteins		
Murashige-Skoog medium B5	Duchefa Biochemie	Cat# M0231.0050
Agar-Agar	Carl Roth	Cat# 4807.4
Gibberellic acid	Duchefa Biochemie	Cat# G0938.5000
Porapak Q, 80-100 mesh,	Sigma-Aldrich	Cat# 20331
Dichloromethane	ThermoFisher	Cat# 1006682500
Tridecane	Sigma-Aldrich	Cat# T57401-25G
β-ocimene	Sigma-Aldrich	Cat# W353901
TRIzol ® Reagent	ThermoFisher	Cat# 15596018
LB medium	Duchefa Biochemie	Cat# L1704.2500
IPTG	ThermoFisher	Cat # R0392
TRIS HCL	Sigma-Aldrich	Cat # T3253
Imidazole	VWR international	Cat # 1.04716.0250
NaCl	Duchefa Biochemie	Cat # S0520
Glycerol	Sigma-Aldrich	Cat # 49781
$MgCL_2$	VWR international	Cat # 1,05833,0250
DTT	ThermoFisher	Cat # R0862
Lysozyme	Sigma-Aldrich	Cat # 62971-10G-F
DNAse I	ThermoFisher	Cat # EN0521
Protease inhibitors	Sigma-Aldrich	Cat # 11836170001
Geranyl pyrophosphate	Sigma-Aldrich	Cat # G6772-1VL
Farnesyl pyrophosphate	Sigma-Aldrich	Cat # 700300P-1MG
Critical commercial assays		
SensiMix SYBR Low-ROX kit	Meridian Bioscience	Cat# QT625-05
Superscript III Reverse Transcriptase	ThermoFisher	Cat# 18080085
TurboDNAse	Thermofisher	Cat# AM2238
Deposited data		
Capsella rubella and *orientalis* whole genome resequencing	Koenig et al.[6]	ENA: PRJEB6689 (https://www.ebi.ac.uk/ena/browser/home)
Capsella grandiflora whole genome resequencing	Josephs et al.[50]	ENA: PRJNA275635 (https://www.ebi.ac.uk/ena/browser/home)

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
Oligonucleotides	List given in Methods S1B	N/A
Recombinant DNA		
Plasmid *TPS02r*	This paper	N/A
Plasmid *TPS02g*	This paper	N/A
Plasmid *TPS03g*	This paper	N/A
Plasmid 35S::TPS02gYFP	This paper	N/A
Plasmid 35S::TPS02rYFP	This paper	N/A
Plasmid 35S::TPS02g^{S8P}YFP	This paper	N/A
Plasmid TPS02g-(His)$_6$-tag	This paper	N/A
Plasmid TPS02r-(His)$_6$-tag	This paper	N/A
Software and algorithms		
AMDIS, V. 2.71	National Institute of Standards and Technology NIST	https://chemdata.nist.gov/
R(version 4.0.3)	R Core Team	https://www.r-project.org/
Agilent ChemStation (G1701 DA D.30.00 MSD)	Agilent	https://www.agilent.com/en/product/software-informatics/analytical-software-suite/chromatography-data-systems/openlab-chemstation
MetAlign	Lommen[51]	https://www.wur.nl/en/show/metalign-1.htm
TagFinder v4.1	Luedemann et al.[52] and Allwood et al.[53]	http://www-en.mpimp-golm.mpg.de/03-research/researchGroups/01-dept1/Root_Metabolism/smp/TagFinder/index.html
ImageJ 1.54f	N/A	http://rsbweb.nih.gov/ij/
EMBOSS Needle	Rice et al.[54]	https://www.ebi.ac.uk/jdispatcher/psa/emboss_needle
Clustal Omega	Rice et al.[54]	https://www.ebi.ac.uk/jdispatcher/msa/clustalo
MView	Brown et al.[55]	https://www.ebi.ac.uk/jdispatcher/msa/mview
LOCALIZER v1.0.4	Sperschneider et al.[56]	http://localizer.csiro.au/
MEGA11	Tamura et al.[57]	https://www.megasoftware.net/
JELLYFISH	Marçais and Kingsford[58]	http://www.cbcb.umd.edu/software/jellyfish.
Stampy v1.0.19	Lunter and Goodson.[59]	https://www.chg.ox.ac.uk/research/scientific-cores/bioinformatics-statistical-genetics/software
GATK v4	Van der Auwera and O'Connor[60]	https://gatk.broadinstitute.org/hc/en-us/sections/360007407851-4-0-0-0
PLINK v1.9	Chang et al.[61]	https://github.com/chrchang/plink-ng
IQTREE	Nguyen et al.[62]	http://www.iqtree.org/
Other		
ZebaTM Spin Desalting columns (7kMWCO)	Thermo scientific	Cat # 89882
Scripts and code for analyses	Figshare	https://doi.org/10.6084/m9.figshare.21379581
Metabolomic files generated in this study	Figshare	https://doi.org/10.6084/m9.figshare.21379581

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Adrien.Sicard (adrien.sicard@slu.se).

Materials availability

Plasmids and plant materials generated in this study will be available from the lead contact upon request.

Data and code availability

- Metabolomics' files generated in this study and average responses of all mass features observed after non-targeted profiling of the volatile organic compounds are available on Figshare. DOIs are listed in the key resources table.
- Original code has been deposited at Figshare and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

In this study, we examined plants or sequence data from the following Species: *Capsella grandiflora*, *Capsella rubella*, *Capsella orientalis* and *Arabidopsis thaliana* (accession *Col-0*).

The *Capsella grandiflora* (*Cg*), *C. rubella* (*Cr*) and *C. orientalis* (*Co*) accessions used in this study were collected from diverse locations in Europe, Africa, and Asia. Detailed information on the locations is given in Methods S1A. The *Cr* x *Cg* RILs and *Co* x *Cr* F2 population used for the QTL mapping experiments were previously described and generated from crosses between *Cr1504* and *Cg926* and *Co1983* and *Cg926*, respectively.[18,40] The near-isogenic lines (NILs) segregating at the (*Z*)-β-ocimene emission QTL (QTL_1) locus were generated by introgressing the QTL_1 *Cg926* alleles present in the *Cr* x *Cg* RILs and *Co* x *Cr* F2 into *Cr1504* (NILrg) and *Co1983* (NILog) through six and four rounds of backcrossing, respectively. These lines were further stabilized through four rounds of selfing while maintaining the QTL_1 heterozygous. The resulting lines were termed *NILrg* and *NILog*, respectively

Seeds were surface-sterilized, plated on half-strength Murashige-Skoog medium supplemented with 0.8 % (w/v) Agar-Agar and 3.3 mg L^{-1} gibberellic acid (Duchefa Biochemie) and placed for two days at 4°C. Germination took place at 22°C for ten days. To promote the transition to flowering seedlings were vernalized for ten days at 4°C. Plants were then grown in growth chambers under long-day conditions (16h light/8h dark) with a temperature cycle of 22°C during the day and 16°C during the night and at 70% humidity with a light level of 200 μmol m-2s^{-1}.

METHOD DETAILS

Floral Volatile sampling in *Capsella*

For the purpose of previous studies,[12,13] volatile organic compounds (VOC) emitted by *Cr* x *Cg* RILs and *Co* x *Cg* F3 population have been collected using a dynamic headspace sampling method coupled with gas chromatography-mass spectrometry (GC-MS). The extraction, assigning, and calculation of most compound intensities was done, as described below, through a semi-automated approach. The quantification of (*Z*)-β-ocimene emission in the *Cr* x *Cg* RILs was, however, conducted manually as described in[12] because the close proximity of additional compounds in the available chromatograms impaired the automatic quantification of this compound.

In this study, volatile compounds emitted by NILs, *Capsella* accessions and transgenic *Capsella* plants were analyzed by the headspace solid phase microextraction (HS-SPME) and gas chromatography coupled with electron impact ionization/quadrupole mass spectrometry (GC-EI-MS). HS-SPME coupled with GC-MS facilitates sample processing, improves throughput and was shown to highly efficiently capture VOCs in several species,[47–49,63] including those previously detected in *Capsella* flower blend.[12]

Three inflorescences for each plant individual analyzed were cut and placed inside a 20 ml airtight glass vial following.[64] The number of open flowers per each inflorescence was noted and later used for normalization. Up to 19 individuals were measured in one measurement sequence batch. A blank sample with no inflorescence was included in each sequence batch, and unless stated otherwise, each individual was measured twice. The number of individuals analyzed are indicated in the respective figure legends. Samples were collected between 10:00 am and 1:00 pm and placed in a dark container to minimize post-sampling volatile emission[65] for a period of approximately 30 minutes before proceeding with quantification. Time until measurement was kept approximatively constant for all samples. Sample collection was also randomized to minimize the influence of uncontrolled environmental variables. A sanity check revealed that differences in sample processing did not affect the total amount of volatiles emitted (Kruskal-Wallis, *p* = 0.837).

Headspace samples were analyzed by solid phase microextraction (SPME) and gas chromatography coupled with electron impact ionization/quadrupole mass spectrometry (GC-EI-MS) by using Agilent 6890N24 gas chromatograph (Agilent Technologies, Böblingen, Germany; http://www.agilent.com) and a StableFlex^{TM} SPME-fiber with 65μm polydimethylsiloxane/divinylbenzene (PDMS-DVB) coating (Supelco, Bellefonte, USA). The headspace of samples for SPME was put under the following procedure: incubation at 45°C for 10 minutes, adsorption at 45°C for 5 minutes and desorption at 250°C for 1 minute onto a DB-624 capillary column with 60m x 0.25mm x 1.40μm film thickness (Agilent Technologies, Böblingen, Germany). The GC temperature was programmed as follows: 40°C isothermal for 2 minutes, then 10°C/min ramping to 260°C final temperature, which was then held at a constant level for 10 min. The Agilent 5975B VL GC-MSD system was run using a constant flow of helium at a level of 1.0 mL/min. Desorption from the SPME fibre took place at 16.6 psi with an initial 0.1 min pulsed-pressure at 25 psi. The purge of the SPME fibre was run for 1 min at a purge flow of 12.4 mL/min.[64]

Floral volatile sampling in *Arabidopsis thaliana*

Volatile Organic compounds emitted from *Arabidopsis* flowers were collected as described in Boachon et al.[66] Forty inflorescences for each replicate were detached and placed in a 2 ml glass vials filled with water. The vials were then placed in a 250 ml glass vial and

enclosed in an oven bag made of PET. Charcoal-filtered air was pumped into the bag in a close loop system in which emitted volatile were traped into a glass tube filled with 30 mg of Porapak Q (80–100 mesh, Sigma). VOCs were sampled for 16 hours starting from 7 am before been eluted in 200 μl of dichloromethane spiked with tridecane (10ng/ μl). One μl of solvent extracts of the headspace flower samples were analyzed by gas chromatography coupled with mass spectrometry (GC-MS) using an Agilent 7890N GC system coupled to an Agilent 5975C mass spectrometer (Agilent Technologies, Inc., Santa Clara, CA, USA). A HP-1MS capillary column (30 × 0.25 mm id × 0.25 μm film thickness, 100% Dimethylpolysiloxane) was used for the GC separation of compounds. Therefore, 1 μl of each extract was injected to the system. Helium was used as carrier gas (Helium 6.0) with a pressure of 6.128 psi. The GC temperature program was as follows: Initial oven temperature of 30 °C was held for 1 min, increased at a rate of 5 C min-1 to 150 °C, followed by a rate of 10 C min-1 to the final temperature of 260 °C and held for 10 min. The mass spectrometer was operated in the electron impact (EI) mode at 70 eV, with a gain set to 10. All data were obtained by collecting the full-scan mass spectra within the range of 40–500 m/z. The occurrence of β-ocimene in the chromatograms was determined by comparison to the characteristic ion fragmentation pattern and retention index from the reference standard, using the "Automated Mass Spectral Deconvolution and Identification System" (AMDIS, V. 2.71; National Institute of Standards and Technology NIST, Boulder, CO).[67] The settings for de-convolution were set to low sensitivity, medium resolution, and shape requirement with a component width of 32 and one adjacent peak subtraction. Identification criteria were applied as follows: match factor ≥ 80% with relative retention index deviation ≤ 5% + 0.01 from the reference value. Match factor penalty was set to very strong with a maximum penalty of 25.

VOCs Identification and quantification

The chromatograms were recorded with the mass range set to 30-300 m/z and at 20 Hz scan rate. The obtained chromatograms were controlled visually, exported in NetCDF file format using Agilent ChemStation-Software and baseline-corrected with MetAlign soft-ware.[51] Data processing into a standardized numerical data matrix and compound identification were done using the TagFinder soft-ware.[52,53] Identification of compounds was performed based on comparison to reference standards using mass spectral and reten-tion time matching to the reference collection of the Golm Metabolomics Database (GMD) for volatile compounds. Note that in the case of β-ocimene, a mix of isomers was used (Sigma - W353901). The isomers were annotated based on their retention index order in the NIST library. To assign compound as present in the analyzed sample a the criteria of at least 3 specific mass fragments per compound and a retention time deviation < 3.0% was used. Mass to charge ratios (m/z) for selective and specific quantification of annotated compounds are reported in the format Tag_mass followed by m/z. To define non-annotated compounds, we correlated the mass fragments for each retention time. When mass fragments were not well correlated (r^2 < 0.4), we performed a hierarchical clustering based on dissimilarity indices using an optimal number of clusters determined based on the silhouette method (factoextra package in R). Non-annotated compounds are named Time_###_**, with ### corresponding to the Time_group (i.e., running number of co-eluting masses features) and ** to the cluster number. The numbers of Time_groups and clusters are provided by the TagFinder software.[68] These yet unknown volatile compounds are characterized by assigning m/z and retention time (RT) to each Time_###_** feature. Retention indices were calculated following.[69]

Flower fragrance analysis

For quantification purposes, all mass features were evaluated for the best specific, selective and quantitative representation of observed analytes. The blank values of each volatile compound were subtracted from the corresponding individual samples for each sequence batch. Each sample was then standardized to the number of open flowers. For each compound at a given retention time, we selected the three best correlated (Pearson rank correlation coefficient) mass fragments and averaged blank-corrected and standardized intensities of technical replicates of each individual. Analytes present in at least two biological replicates of at least one population were kept for further analysis. The intensity of the volatile compound detected for each individual was then summed to estimate the total emission. The relative proportion of each detected compound present in the flower scent was determined by dividing each volatile compound's intensity by the sum of the intensity of all detected compounds in the corresponding samples to obtain relative abundances of each annotated and non-annotated volatile. The comparison of β-ocimene emission between ge-notypes was performed by scaling the standardized response of the corresponding mass fragments to the maximal intensity de-tected in the experiment.

To assess the difference in volatile composition between *Capsella* species, we first performed a multivariate analysis on the relative proportion of each compound. We calculated Bray-Curtis dissimilarities that we illustrated with a non-metric multidimensional scaling (NMDS) plot (vegan package in R). We next performed hierarchical clustering based on the average Bray-Curtis dissimilarity scores to visualize the relationship between the volatile profiles emitted by the different populations analyzed (Dendextend in R). We formally tested whether the mating systems affected the floral scent composition by conducting a permutational multivariate analysis of variance (PERMANOVA, adonis function from the vegan package in R using 1000 permutations) using the mating system, species and populations as factors. Additionally, we investigated the importance of different compounds for classifying populations into different mating types using the Random Forest algorithm (randomForest package in R).[70] We use the algorithm to calculate the prob-ability of being assigned to the correct mating systems categories based on the relative abundance of different scent compounds (n_{tree} = 1000, m_{try} = 7). We next extracted the mean decrease accuracy for each compound which describes the extent to which they contribute to the accurate classification of the samples.

Because the above analysis suggested a convergent evolution of flower scent in the two selfers, we further characterize differ-ences in flower fragrance composition between the different species and mating systems. To this end, we constructed linear models

with zero-inflated beta distribution (glmmadmb package in R) using mating systems and/or species as fixed factors and population nested within mating systems or species as random variables. The model terms' significance was determined using the likelihood test (lrtest function from the lmtest package in R). In parallel, we also conducted pairwise comparisons of the abundance of the different compounds between *Capsella* species using Tukey's honest significant differences test to identify significantly reduced or increased compounds in the selfers compared to *C. grandiflora* (agricolae package in R). The resulting number of compounds emitted at significantly different levels by the *Capsella* species were used to visualize the overlap between the evolution of flower scent that occurred in the two selfers (VennDiagram package in R).

QTL mapping
The genotypes of *Cr x Cg* RILs and *Co x Cg* F2 mapping populations have been obtained as previously described.[18,40] Scent compounds emitted by both populations were collected as described above. QTLs for the selected volatile compounds were mapped using R/QTL.[71] The single-QTL genome scan was conducted using a non-parametric model. QTLs were tested in 1 cM intervals and assessed using LOD scores (the $\log_{10}$ likelihood ratio). Genome-wide permutations (1000) under the non-parametric model were used to determine significance thresholds (5%). A 1.5-LOD support interval was used to estimate the position of each QTL. The results of the QTL analyses are presented in Table S2.

Fine mapping
We first identified a *NILrg* segregating in the region containing the QTL_1 interval by genotyping plants with markers G06, G07 and G09. To refine the position of QTL_1, around 1,200 *NILrg* progenies were screened for plants harbouring a meiotic recombination between the markers m1621 and m1629, each located at the opposed borders of the QTL confidence interval (Table S2). The position of the recombination breakpoints was determined by genotyping informative recombinants with additional markers located in the focal region. Floral scent was collected from each of the selected recombinant individuals two times using the static headspace sampling method as described above. The primers used for the genotyping are shown in Methods S1B.

Molecular cloning and plant transformation
Cr (TPS02r) and *Cg* (TPS02g and TPS03g) genomic constructs were used to complement the *NIL* plants homozygous for the *Cr* allele (*NILrr*) or *Arabidopsis thaliana* Col-0 plants. All constructs were PCR-amplified from either *NILrr* or *NILgg* (NIL homozygous for the *Cg* allele) genomic DNA. A 6.1 kb TPS02r and 5.9 kb TPS02g genomic fragments starting ~1.3 kb before the transcription start site of TPS02 and ending ~1.3kb after the stop codon were amplified in two PCR reactions (sequences of primers used are presented in Methods S1B). TPS03g genomic fragment was 5.2 kb in length including 1.3 kb upstream the transcription start site and 0.9 kb of 3' sequences. The genomic fragments were assembled and subcloned into a modified version of pBluescript II KS (StrataGen, pBlueM-LAPUCAP) using the InFusion® HD Cloning Plus kit (Clontech), yielding the construct pBS-TPS02r and pBS-TPS02g. Following the same cloning approach, a sequence coding for the yellow fluorescent protein was inserted before the codon stop position of TPS02 in the resulting plasmids. The sequences corresponding to the TPS02YFP recombinant proteins were then transferred to pAS95, modified pBluescript II KS containing a CaMV 35S promoter and terminator, yielding the 35S::TPS02gYFP and the 35S::TPS02rYFP constructs. The serine at the 8th amino acid position was then converted into a proline by PCR-based site-directed mutagenesis, giving rise to the 35S::TPS02g^{S8P}YFP construct. To express recombinant TPS02g and TPS02r protein in bacteria, we cloned a truncated version of TPS02 ORFs fused to a 6 x histidine (His)$_6$-tag in pet11a. These protein versions started at amino acid 38, the predicted cleavage site of the chloroplast peptide signal. In line with previous observations, this truncation was necessary to obtain a soluble protein version and is believed to naturally occur in plants upon chloroplast entry before enzymatic reaction occurred.[41,72] The primers used to generate the constructs are shown in Methods S1B.

After verification by Sanger sequencing, the fragments were transferred into the AscI site of the plant transformation vector pBar-MAP, a derivative of pGPTVBAR (Becker et al.[73]). These genomic constructs were then used to transform *Capsella NILrr* plants or *Arabidopsis thaliana* Col-0 accession by floral dip[74] or Arabidopsis thaliana leaf protoplasts using PEG-mediated transfection.[75]

Gene expression and splicing analysis
Total RNA was extracted from inflorescences using Trizol (Life Technologies). RNAs were treated with TurboDNAse (Ambion) and reverse transcribed using oligo(dT) and the Superscript III Reverse Transcriptase (Invitrogen). Quantitative RT-PCR was performed using the SensiMix SYBR Low-ROX kit (Bioline) and a LightCycler® 480 (Roche). The primers used are described in Methods S1B. For each genotype, three biological replicates were used, and for each of them, three technical replicates were included. To detect splicing difference, primers onw540 and onw541 were used to amplify the full-length *TPS02* transcripts.

TPS02 activity assay
TPS02-(HIS)$_6$ were transformed into *Escherichia coli* strain BL21(DE3)pLysS (Life Technologies). *E.coli* cultures (400 ml) were grown in LB medium at 37°C until OD_{600} =0.6 before inducing the expression of recombinant proteins with 1 mM IPTG and transferring them at 20°C for 16 hours. Cells were then collected by centrifugation at 4000 g for 20 min and resuspended into 15 ml of chilled extraction buffer (50 mM tris HCL [pH 7.5], 20 mM imidazole, 300 mM NaCl, 10 % glycerol (v/v), 5mM MgCL$_2$, 1mM DTT, 0,3 mg/ml Lysozyme (Sigma), 1 U/ml of DNAse I and protease inhibitors(Roche)). Cells were disrupted by sonication (3 x 30s – 50% amplitude of ultrasonic liquid processsors, MISONIX) and the soluble protein fraction was recovered by centrifugation for 30 min at 4°C. Recombinant

proteins were purified using NEBExpress® Ni-NTA Magnetic Beads (New England Biolabs) and transferred into the enzymatic assay buffer (25mM MES, 25mM CAPS, 50mM Tris, 5mM MgCl2 at [pH 7.5]) using ZebaTM Spin Desalting columns (7kMWCO – Thermo scientific). Protein concentration was then determined by SDS-page. Steady-state analyses of TPS02 activity were conducted using the malachite green phosphate assay kit (Sigma-Aldrich, No. MAK307) following.[76] Assays were conducted using 0.2 μM enzyme in 50 μl of enzymatic assay buffer supplemented with 5 mU of the coupling enzyme inorganic phosphatase (Sigma) with varying concentrations of substrate (from 1.56 to 100 μM), Geranyl pyrophosphate (GPP) or Farnesyl pyrophosphate (FPP), for 5 min at 30°C. The reactions were stopped by adding 12 ul of Malachite green development solution and incubated 15 min at room temperature before reading at 620 nm using a FLUOstar Omega plate reader (BMG Labtech). The data was fitted to a non-linear regression of the Michaelis-Menten equation using the drc package implemented in R.[77]

Confocal imaging

To analyze the cellular localization of TPS, we imaged transformed *A. thaliana* protoplast with a confocal laser scanning microscope Zeiss LSM710 (Zeiss) using an excitation wavelength of 514 nm, with emission collected above 630 nm for Chlorophyll A and between 520 - 570 for YFP. Above-threshold Pearson's correlation coefficients between the Chloroplhyll A and YFP signal intensities were determined using ImageJ (http://rsbweb.nih.gov/ij/) to assess the efficiency of TPS02 targeting within chloroplasts.

Sequence and population-genetic analyses

Pairwise sequence alignments of nucleotide sequences of *NILrr* and *NILgg* TPS02 alleles were performed using EMBOSS Needle.[54] Protein sequence alignments were performed using the Clustal Omega and visualized by MView.[54,55] Chloroplast transit peptides were predicted using the LOCALIZER 1.0.4.[56]

To correlate the genotype at candidate polymorphisms with scent emission, five *Cg* individuals and five *Cr* accessions (three individuals per each accession) were sequenced in the regions containing five non-synonymous variants within *TPS02* and phenotyped for (Z)-β-ocimene emission as mentioned above.

The dataset used for the population genetic analysis included 180 Cg individuals from a single population and species-wide samples of 13 *Cg* and, 51 *Cr* and 16 *Co* individuals.[6,50,78] TPS02 haplotypes were reconstituted by a combination of local assembly and multiple pair-end phasing (similarly to Sicard et al.[79]). The resulting sequences were used to reconstruct neighbour-joining trees with MEGA11 using the Tamura 3-parameter method and gamma distribution.[57] MEGA11 was also used to compute genetic distances as the number of nucleotide differences per sequence comparison. The hierarchical clustering of the samples based on the genotypes at *Cr1504/Cg926* TPS02 non-synonymous variants was done using Elucidean distances. The counts of *Cr1504* or *Crub183* k-mers of different sizes in different samples were estimated using JELLYFISH.[58]

To visualize the genetic distances between the three Capsella species, reads were mapped to the *Capsella* reference genome[9] using Stampy v1.0.19.[59] After bioinformatics processing using Picard tools (https://picard.sourceforge.net/), Raw SNP calls were generated by the joint calling of all individuals in GATKv4.[60] To determine the genetic relationship between *Capsella* species, we first identify the ten most unrelated individuals from each species. For each species separately, we pruned the dataset with a target of ten individuals by varying the relatedness cutoff using plink 1.9.[61] We then computed the distance matrix with the same software using the "hamming distance" procedure. The dendrogram was created by performing a hierarchical clustering based on the resulting distance matrices using the Ward method in the R package *hclust*.

C. rubella phylogeny was reconstituted using the substitution model in the IQTREE software.[62] To root the tree, we used a *C. grandiflora* individual representative of the species genetic identity, i.e. the individual with the lowest contribution on the principal components of a genetic principal component analysis performed with the R package *pcadapt*.[80] We looked at the *TPS02* haplotype diversity across *C. rubella* genotypes by extracting the variants located in the *TPS02* locus and performing an admixture analysis.[81] To further assess the regime of selection under which TPS02 evolves, we performed a whole genome screen of synonymous versus non-synonymous polymorphism (π_S and π_N respectively) in coding regions. Based on the reference genome sequences, we defined the 4-fold degenerate nucleotide positions as synonymous sites (S sites) and the 0-fold degenerate nucleotide positions as non-synonymous sites (N sites). We compute the nucleotide diversity in S and N sites per gene and per species using customized r script and the *nuc.div* function from *pegas* R package.[82]

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were conducted in R statistical programming language.[77] Tukey-HSD tests were performed using agricolae package add-ons implemented in R. Data were presented as mean and standard deviation and comparisons with p-values below 0.05 were considered significant for multiple comparisons of phenotypic means. Data were illustrated as box plots using the ggplot2 package implemented in R.[77] Outliers were defined as values deviating from the mean by more than 3 Standard deviations. All boxplots' lower and upper hinges correspond to the first and third quartiles, respectively. The upper and lower whiskers add or subtract 1.5 interquartile ranges to/from the 75 and 25 percentile, respectively. The middle lines represent the median.