In Planta Recapitulation of Isoprene Synthase Evolution from Ocimene Synthases

Mingai Li,[1] Jia Xu,[1,2] Alberto Algarra Alarcon,[3,4] Silvia Carlin,[3] Enrico Barbaro,[1] Luca Cappellin,[3] Violeta Velikova,[1,5] Urska Vrhovsek,[3] Francesco Loreto,[6] and Claudio Varotto*[,1]

[1]Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige (TN), Italy

[2]Dipartimento di Biologia, Università di Padova, Padova, Italy

[3]Department of Food Quality and Nutrition, Research and Innovation Centre, San Michele all'Adige (TN), Italy

[4]Institute of Ecology, University of Innsbruck, Innsbruck, Austria

[5]Institute of Plant Physiology and Genetics, Bulgarian Academy of Sciences, Sofia, Bulgaria

[6]Department of Biology, Agriculture and Food Sciences, The National Research Council of Italy (CNR), Rome, Italy

***Corresponding author:** E-mail: claudio.varotto@fmach.it.

Associate editor: Juliette de Meaux

Abstract

Isoprene is the most abundant biogenic volatile hydrocarbon compound naturally emitted by plants and plays a major role in atmospheric chemistry. It has been proposed that isoprene synthases (IspS) may readily evolve from other terpene synthases, but this hypothesis has not been experimentally investigated. We isolated and functionally validated in *Arabidopsis* the first isoprene synthase gene, *AdoIspS*, from a monocotyledonous species (*Arundo donax* L., Poaceae). Phylogenetic reconstruction indicates that AdoIspS and dicots isoprene synthases most likely originated by parallel evolution from TPS-b monoterpene synthases. Site-directed mutagenesis demonstrated in vivo the functional and evolutionary relevance of the residues considered diagnostic for IspS function. One of these positions was identified by saturating mutagenesis as a major determinant of substrate specificity in AdoIspS able to cause in vivo a dramatic change in total volatile emission from hemi- to monoterpenes and supporting evolution of isoprene synthases from ocimene synthases. The mechanism responsible for IspS neofunctionalization by active site size modulation by a single amino acid mutation demonstrated in this study might be general, as the very same amino acidic position is implicated in the parallel evolution of different short-chain terpene synthases from both angiosperms and gymnosperms. Based on these results, we present a model reconciling in a unified conceptual framework the apparently contrasting patterns previously observed for isoprene synthase evolution in plants. These results indicate that parallel evolution may be driven by relatively simple biophysical constraints, and illustrate the intimate molecular evolutionary links between the structural and functional bases of traits with global relevance.

Key words: **isoprene synthase evolution, short-chain terpene synthases parallel evolution, site-directed mutagenesis, substrate specificity, active site size modulation, ocimene synthase.**

Article

Introduction

Isoprene (2-methyl-1,3-butadiene, C_5H_8) is the most abundant naturally emitted biogenic volatile compound, first discovered in plants 60 years ago (Sanadze 1957). Many plants, especially perennial, fast-growing forest species, constitutively emit large amounts of isoprene into the atmosphere (Dani et al. 2014; Loreto and Fineschi 2015), corresponding to a yearly global emission of ~500 Tg C (Guenther et al. 2012). Moreover, isoprene rapid oxidation, initiated by hydroxyl radicals (OH) is estimated to contribute almost half of secondary organic aerosol (SOA) (Worton et al. 2013), which plays an essential role in regulating atmospheric chemistry (Heald et al. 2009; Archibald et al. 2010; St. Clair et al. 2016).

In plants, isoprene biosynthesis is catalyzed in chloroplasts by isoprene synthase (IspS) from dimethylallyl diphosphate anion (DMADP; supplementary fig. S1, Supplementary Material online) which is formed by the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway (Schwender et al. 1997). Isoprene emission rates depend on the activity of isoprene synthase and the pool size of DMADP (Wiberley et al. 2009; Rasulov et al. 2010), which are in turn influenced by many factors, like the endogenous developmental stage of a leaf (Ahrar et al. 2015; Niinemets et al. 2015) and several environmental stimuli and constraints (Loreto and Schnitzler 2010; Sharkey and Monson 2017).

Many experiments have demonstrated that isoprene plays fundamental roles in protecting plants from environmental stresses (Vickers et al. 2009; Velikova et al. 2011; Tattini et al. 2014). There may be many mechanisms acting concurrently, whose examination is beyond the scope of this report (reviewed in Loreto and Schnitzler 2010; Sharkey and

Monson 2017). Recent studies suggest that isoprene may play a general role as signaling molecule to prime and regulate stress-induced genes that activate defenses against abiotic stresses (Harvey and Sharkey 2016).

Isoprene emission could be detected across the whole plant kingdom, but in the angiosperm lineage, isoprene emitting species are scattered across the phylogenetic tree (Monson et al. 2013; Dani et al. 2014). Even though there are ~20% of plant species emitting isoprene from diverse taxonomic groups (Loreto and Fineschi 2015), only a relatively limited number of isoprene synthase genes has been so far isolated and characterized (Sharkey et al. 2013). Several views explaining such sparse phylogenetic occurrence of isoprene have been proposed. The debate focused initially mainly on whether IspS appeared only once early in the evolution of land plants, or multiple times independently (Hanson et al. 1999; Harley et al. 1999; Lerdau and Gray 2003; Sharkey et al. 2005, 2008). With the elucidation of the general patterns of evolution of terpene synthases (Chen et al. 2011), it became clear that the isoprene emission trait must have evolved independently in the major lineages of land plants by parallel evolution (Sharkey et al. 2013). It is, however, still debated whether this model applies also to lower taxonomic levels, and in particular to the rosids, the clade from which the majority of IspS was obtained in recent years. One view proposes that a single gain of isoprene synthase function took place early during the radiation of rosids, followed by multiple losses (Sharkey 2013; Sharkey et al. 2013). Adaptive tradeoffs between metabolic costs and selective advantages due to paleoclimatic fluctuations of atmospheric CO_2 concetration have been proposed to account for the repeated losses of the trait (Sharkey and Yeh 2001; Sharkey et al. 2013). A single isoprene synthase gain in rosids is also supported by the finding that *IspS* genes form a monophyletic group in this plant order (Sharkey et al. 2013). The other scenario proposes that the actual number of independent gains of isoprene emission trait is much higher, suggesting that *IspS* genes can undergo repeated gain and loss at the family and even genus level thanks to the hypothetically low number of amino acidic mutations which may be sufficient for evolution (Monson et al. 2013). Parsimony-based analyses based on the mapping of isoprene presence/absence on species phylogenies obtained with molecular markers independent from *IspS* genes support this conclusion (Monson et al. 2013). Under this view, the retention of the trait is postulated to require adaptive advantage(s) in a narrow range of environments and phenotypes. Other authors support this multiple gains and multiple losses view, further linking it to species richness at the genus level and the perennial growth habit of the species supporting the trait (Dani et al. 2014). As of today, it is not clear to what degree each of these scenarios explains the overall evolutionary trajectory followed by the *IspS* gene family in angiosperms, as the actual number of mutations required for IspS to evolve from other terpene synthases and the precise identity of such ancestral enzymes are unknown (Monson et al. 2013; Sharkey 2013; Sharkey et al. 2013; Sharkey and Monson 2017).

The crystal structure of *P. x canescens* IspS elucidated the overall similarities of isoprene synthases mechanism of catalysis. Like all other known Class I terpene synthases, IspS enzymes require for catalysis the coordination of a triad of magnesium ions, which in turn bind the pyrophosphate anion moiety of the DMADP substrate (Köksal et al. 2010). The amino acids coordinating the metal ions form two separate and highly conserved clusters on IspS helices D and H, called the DDXXD and DTE/NSE motifs, respectively. The DMADP substrate diphosphate anion moiety is bound by the magnesium triad on top of the substrate binding pocket, while its apolar moiety is positioned in the hydrophobic cleft which extends in the lower part of the central cavity of each IspS monomer (Köksal et al. 2010). Two conserved phenylalanine residues (F338 and F485) make Van der Waals contacts with the substrate at the bottom of the hydrophobic cleft. Like in the other Class I TPS, the catalytic mechanism involves the formation of geranyl and/or linalyl carbocations, followed by elimination of the substrate diphosphate moiety, with the diphosphate-leaving group likely serving as a general base (Köksal et al. 2010). By contrast to the majority of class I terpene synthases characterized to date, no obvious conformational change associated to the transition from the enzyme open to close (substrate-bound) state could be inferred for *P. x canescens* IspS, casting doubts on whether IspS may lack such mechanism of active site "capping" (Köksal et al. 2010).

Based on *Populus x canescens* IspS crystal structure and alignment of isoprene synthase proteins from rosids, a core of four conserved amino acids distinguishing isoprene synthase from phylogenetically closely related terpene synthases in angiosperms was identified (Sharkey et al. 2013). Two of the residues (F338 and F485) were suggested to constrain the size of the binding pocket, but the relevance of this hypothesis for IspS function was not experimentally tested (Gray et al. 2011; Sharkey et al. 2013). Since then, these amino acids have been considered as essential markers while screening for enzymes having exclusive isoprene synthase activity from sequence identity. The other two amino acids of the IspS diagnostic tetrad (S446 and N505) (Sharkey et al. 2013) are preferentially found in isoprene synthases, but recent identification of novel *IspS* genes with amino acid variants indicate that these positions display some degree of variation (Ilmén et al. 2015; Oku et al. 2015). All the validated *IspS* genes so far have been identified in rosids, with the only exception of one gene isolated from asterids (Ilmén et al. 2015). More importantly, to date no report on the identification of isoprene synthase genes from monocotyledons exists.

The Arundinoideae is a small subfamily of the Poaceae. Out of the six taxa from four genera investigated, the majority were found to emit isoprene (Hewitt et al. 1990; Kesselmeier and Staudt 1999; Ahrar et al. 2015). The most highly emitting species among them is *Arundo donax* (giant reed), a perennial fast-growing weed, which can grow up to 6–8 m in height in mediterranean climate regions all over the world. *Arundo donax* is used as an important biomass and biofuel crop as it may produce >35 tons ha^{-1} $year^{-1}$ of dry biomass with low agronomic input, and a yield in bio-ethanol that may reach 50% of the dry matter (Angelini et al. 2009).

In this study, we identified and functionally validated an isoprene synthase gene from *A. donax*, which, to the best of our knowledge, is the first isoprene synthase characterized

from monocot species. Phylogenetic reconstruction indicates that monocots and dicots isoprene synthases are most likely the result of parallel evolution. Evolutionarily informed site-directed mutagenesis and characterization of total volatile emission in *Arabidopsis thaliana* as heterologous host support the pivotal role of a single amino acid of the diagnostic tetrad in IspS evolution and suggest the likely ancestors of IspS enzymes. Based on these results we propose a model reconciling in an unified conceptual framework the apparently contrasting views previously proposed for isoprene synthase evolution in rosids. The mechanism of neofunctionalization of active site size reduction demonstrated in this study for IspS seems to be general, as the very same amino acidic position is implicated in the parallel evolution of different short-chain terpene synthases from both angiosperms and gymnosperms.

Results

Identification of a Putative *IspS* from *A. donax*

Several terpene synthase genes were identified in the transcriptome of *A. donax* (Sablok et al. 2014) based on sequence homology to poplar *IspS*, and the best hit (*AdoIspS*) was selected for further confirmation. The identified *AdoIspS* genomic locus (GeneBank accession No. KX906604) codes for an open reading frame of 581 amino acids with a predicted molecular mass of 69 kD. The first 16 amino acids were identified by the TargetP program as the putative transit peptide for chloroplast import (Emanuelsson and Nielsen 2000). As already reported for other *IspS* genes, attempts to express a functional protein in *E. coli* were not successful despite removal of the transit peptide (Sasaki et al. 2005; Ilmén et al. 2015; Oku et al. 2015). An overexpression construct containing the full-length cDNA of putative *AdoIspS* gene under the control of CaMV 35S promoter (p35S:AdoIspS-WT) was, therefore, transformed into Col-0 wild-type *Arabidopsis* plants. Isoprene emission was detected in the majority of a set of 120 independent T1 transgenic lines overexpressing *AdoIspS* compared with Col-0 wild-type plants (supplementary fig. S2, Supplementary Material online). Expression of the *AdoIspS* transgene was confirmed by semi-quantitative RT-PCR in 12 lines emitting isoprene (supplementary fig. S3, Supplementary Material online), indicating that isoprene emission was due to stable expression of the transgene into the *Arabidopsis* genome.

To conclusively demonstrate that isoprene emission was due to the *AdoIspS* transgene, PCR amplification with primers specific to the *AdoIspS* transgene and isoprene emission measurement were conducted in parallel in the T2 segregating population of the highly emitting single copy line AdoIspS-44 (supplementary fig. S4, Supplementary Material online). Isoprene emission above background level was detected only in individuals positive for the *AdoIspS* transgene ($n = 78$) and not in the negative ones ($n = 30$). The null hypothesis of random association between isoprene emission and transgene presence in the analyzed sample was rejected at $P < 0.001$ (two tailed Fisher's exact test; $P = 2.27E\text{-}27$). These results demonstrate that the *AdoIspS* gene is the first isoprene synthase gene identified so far from a monocot species, and that the predicted 16-aa transit peptide is sufficient for chloroplast targeting of the mature protein (Emanuelsson and Nielsen 2000).

The *AdoIspS* gene (3,764 bp in length; see fig. 1*A*) comprises seven exons and six introns alike all monoterpene synthases in angiosperms (Trapp and Croteau 2001). The major differences compared with kudzu and aspen *IspS* genes are a shorter first exon (consequence of the shorter transit peptide) and a greater length of the second intron. The multiple sequence alignment of the C-terminal domain (Class I terpenoid synthase fold) has an overall average of 144 amino acids differences out of 330 aligned positions, while the number of differences of AdoIspS from the other validated IspS ranges from 177 to 195, indicating that this is the most divergent protein in the alignment (fig. 1*B*). Modeling of the tertiary structure based on the crystal structure of *P. x canescens* IspS (Köksal et al. 2010) (PcIspS, 39.3% identity with AdoIspS) confirmed a very similar folding of the two proteins (Cα RMSD = 0.19 Å comparing 522 residues), with typical alpha-helix rich structure (supplementary figs. S5 and S6, Supplementary Material online). In addition, the distribution of electrostatic potential on the surface and active site of the two proteins are largely conserved (supplementary fig. S6, Supplementary Material online). Structural alignment allowed the reliable identification of functionally homologous amino acids in *A. donax* and *P. tremuloides* proteins.

Tissue- and Stress-Specific Expression Patterns of *AdoIspS*

Semi-quantitative rtPCR was performed to elucidate the expression patterns of *AdoIspS* in different organs of *A. donax*. *AdoIspS* was highly expressed in leaf blade, while its expression dramatically decreased in leaf sheath (supplementary fig. S7, Supplementary Material online). Only traces of expression were detected in emerging rhizome bud, and there was no expression in root, node and internode.

To estimate the general stress responses of *AdoIspS* in *A. donax* at molecular level, young seedlings grown in hydroponic solution were treated with either 150 mM NaCl, 15% PEG 6000, 500 μM $CdSO_4$, or a 42 °C heat shock, and the entire shoots were collected to examine the expression pattern of *AdoIspS* at different time points in different stress conditions, (fig. 2*A*–*D*), respectively. The salt treatment did not result in significant variation of the *AdoIspS* expression levels compared with untreated controls at any time points (fig. 2*A*). PEG and heavy metal treatments induced the upregulation of *AdoIspS* at specific time points (30′ for both treatments; 11 h and 3 h for PEG and heavy metal treatments, respectively; fig. 2*B* and *C*). Heat shock caused the largest transcriptional response of *AdoIspS*, resulting in a significant upregulation of the gene at all time points tested, up to a 8-fold with respect to unstressed levels (fig. 2*D*).

Distinctive Features of *A. donax* IspS with Respect to Dicot IspSs

Phylogenetic reconstructions with representatives of all major classes of angiosperm terpene synthases using both

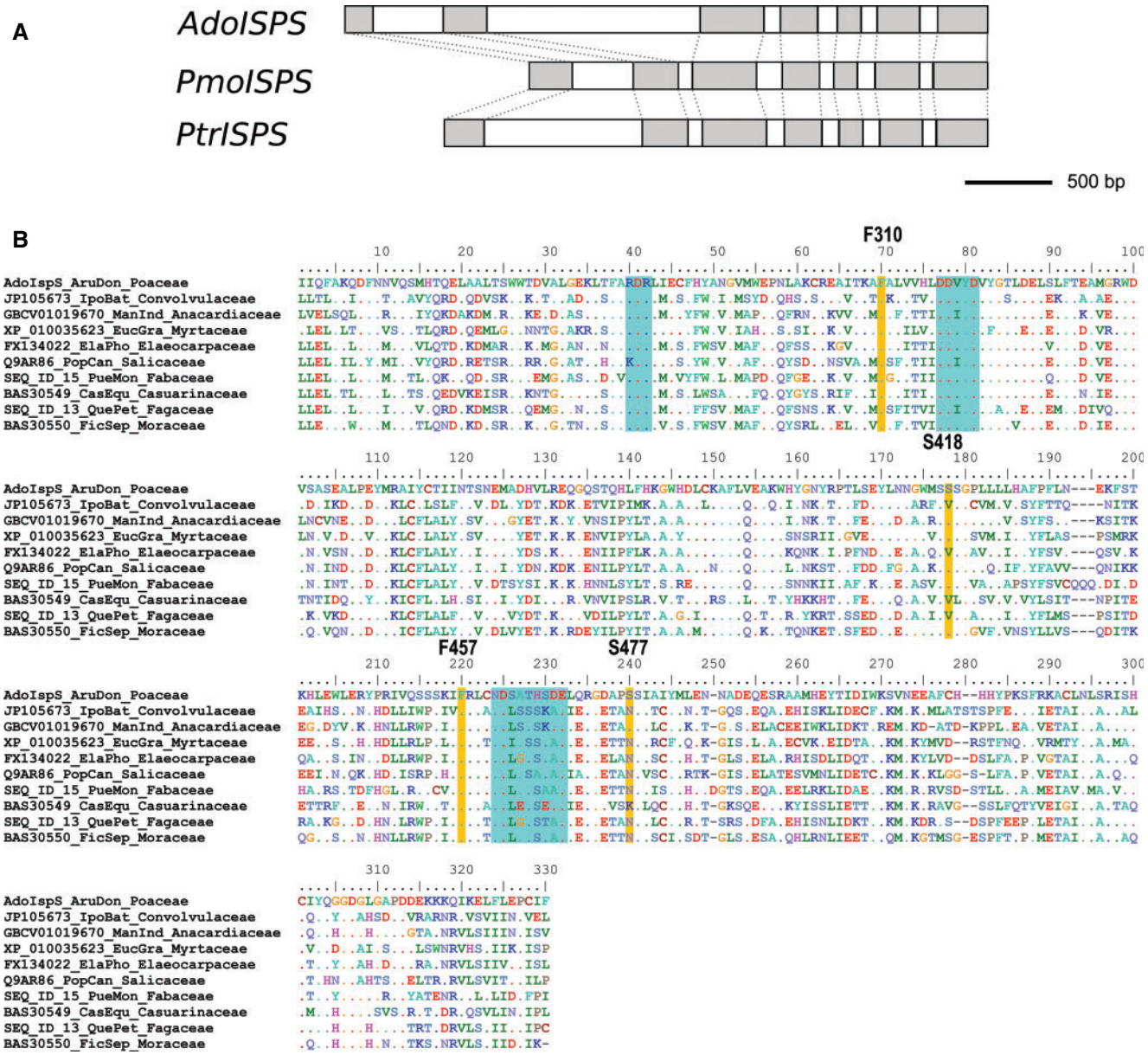

Fig. 1. Comparison of AdoIspS to selected isoprene synthases from dicots. (*A*) Gene structure of Isoprene synthase from *Arundo donax* (*AdoIspS*), kudzu (*Pueraria montana*, *PmoIspS*) and aspen (*Populus tremuloides*, *PtrIspS*). Exons correspond to gray boxes, introns to white ones. (*B*) Multiple sequence alignment of the C-terminal domain (Class I terpenoid synthase fold) for one selected IspS protein per family. Orange shading indicates residues of the diagnostic tetrad (bold numbers refer to aligned positions of the AdoIspS sequence), while light blue shading indicates the RXR, DDXXD and DTE/NSE motifs. Dots represent residues which are identical to those in AdoIspS.

maximum likelihood (ML) and Bayesian inference (BI) confirmed that AdoIspS belongs to class b terpene synthases (supplementary figs. S8 and S9, Supplementary Material online). AdoIspS is basal with respect to the majority of the other known isoprene synthases, and forms a monophyletic clade with two putative terpene synthases from *Sorghum bicolor* with unknown function (Sb06g002820 myrcene/ocimene synthase, putative; and Sb01g039090 lyase/magnesium ion binding). To gain a more precise relationship of AdoIspS with all isoprene synthases validated to date, ML and bayesian reconstructions were carried out on a larger set of class b terpene synthases and some class g terpene synthases as outgroup (fig. 3). The placement of AdoIspS is in both cases basal to the major isoprene synthase group (closely related to poplar IspS), but differs depending on the phylogenetic method employed. In the ML tree (fig. 3), AdoIspS is at the base of TPS-b Clade 2, while in the BI tree it forms a separate clade basal to both Clade 1 and Clade 2 (part of the clade of basal IspS, supplementary fig. S10, Supplementary Material online). The ML topology corresponds to the one published earlier (Sharkey et al. 2013) and to the one observed in the reconstruction with all TPS classes, thus lending support to the positioning of AdoIspS at the base of Clade 2. Analogously to AdoIspS, the isoprene synthase gene from *Casuarina equisetifolia* (Oku et al. 2015) and the myrcene/isoprene synthase gene from *Humulus lupulus* (Sharkey et al. 2013) do not

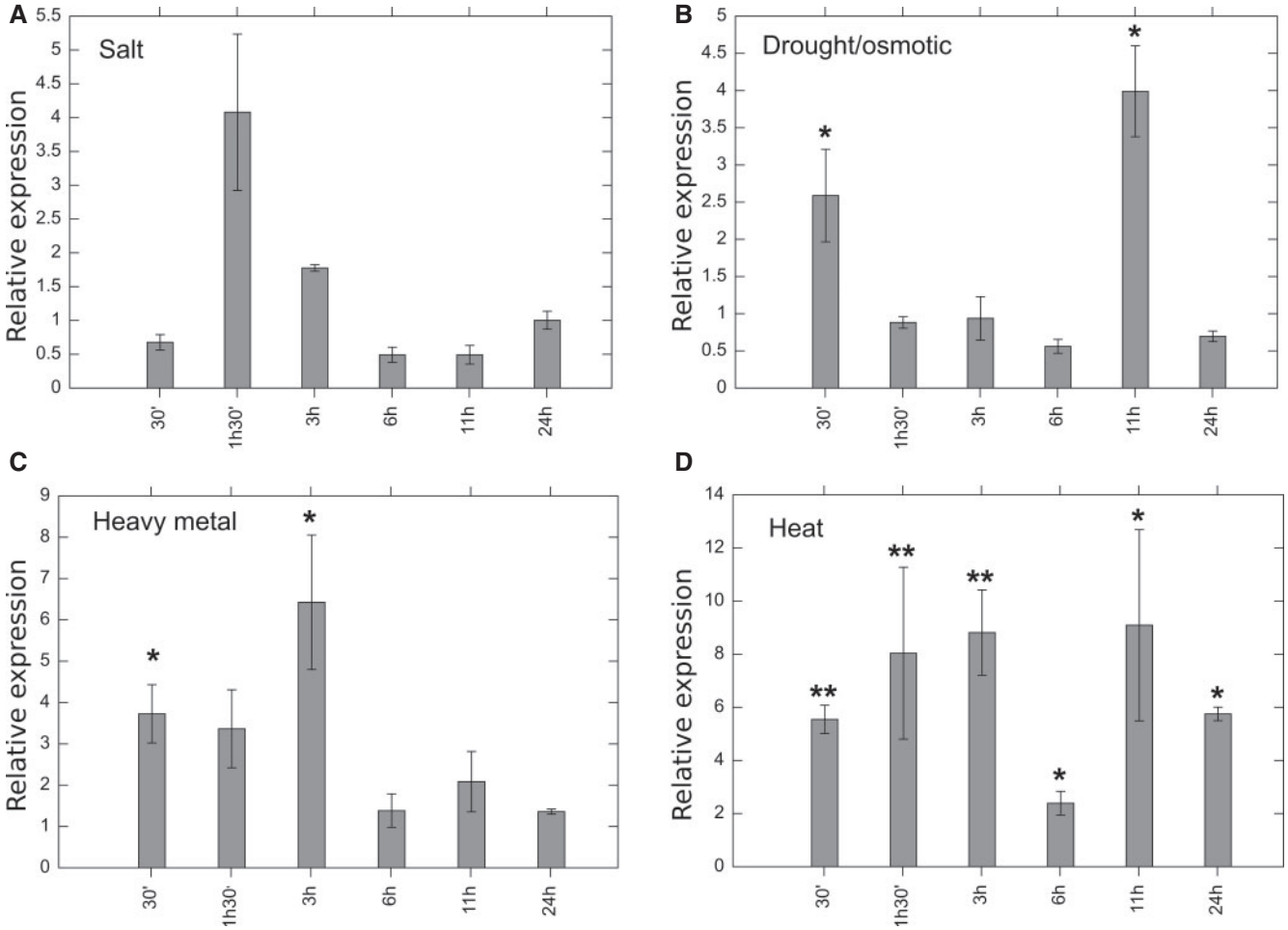

FIG. 2. Time course of *AdoIspS* expression in response to 150 mM NaCl (*A*), 15% PEG 6000 (*B*), 500 μM $CdSO_4$ (*C*), and 42 °C heat shock (*D*) treatments of *Arundo donax* shoots. The relative expression of *AdoIspS* in stressed versus unstressed plants was calculated for different lengths of the stress treatment (0.5, 1.5, 3.0, 6.0, 11.0 and 24.0 h) and normalized to the mean of internal control *AdoActin* at the same time points. Data represent the mean ± SD of three independent experiments. Expression differences were analyzed with Student's *t*-test, correcting for multiple tests using Holm's method. For each time point, the black stars indicate statistically significant differences between *AdoIspS* expression means of treated and untreated samples (*statistically significant at $P < 0.05$; **statistically significant at $P < 0.01$).

belong to the main IspS clade, but to TPS-b Clade 1, confirming that they likely originated from independent events of parallel evolution (fig. 3 and supplementary fig. S10, Supplementary Material online).

AdoIspS shared with isoprene synthases and other class I terpene synthases the functionally relevant motifs previously identified (Degenhardt et al. 2009): the two conserved metal-binding motifs (DDXXD and DTE/NSE) and the RxR motif, involved in stabilization of the released pyrophosphate (fig. 1*B*). As confirmed by structural alignments with *Populus x canescens* IspS (PcIspS; Köksal et al. 2010), only three out of the four amino acids that together are considered diagnostic of IspS activity (Sharkey et al. 2013) (hereafter indicated as "IspS diagnostic tetrad") were conserved in AdoIspS. In the AdoIspS model, F310, S418 and F457 cluster close to each other and contribute to the definition of the substrate binding site of the enzyme, while S477, the last amino acid of the H-α1 loop, is separated from the others at the top of the active site (supplementary fig. S11A and B, Supplementary Material online). As the IspSs identified so far from species/genera belonging to the same family do not differ in their diagnostic tetrad (supplementary table S1, Supplementary Material online), we derived a family-specific diagnostic tetrad profile, which equally weighs the validated variations among families and mapped it onto the phylogram representing the known taxonomy of the considered families (fig. 4). AdoIspS diagnostic tetrad is identical to that of Anacardiaceae (Ilmén et al. 2015) and differs from the standard one of Salicaceae by the presence in the fourth position of a serine (S477) instead of an asparagine (N505 in *Populus* species; see protein structural alignment in fig. 1*B*). Excluding from computation the *H. lupulus* gene, which is primarily a myrcene synthase with secondary IspS activity (Sharkey et al. 2013), both first and third positions of the diagnostic tetrad were invariantly F residues, while the other 2 positions allowed a limited number of variants (supplementary table S1, Supplementary Material online). The ranking in decreasing order of conservation among positions was: Pos1 = Pos3 > Pos2 > Pos4. Only four amino acidic combinations (FSFN, FVFN, FSFS, FVFK) were found to be present in obligate isoprene synthases across

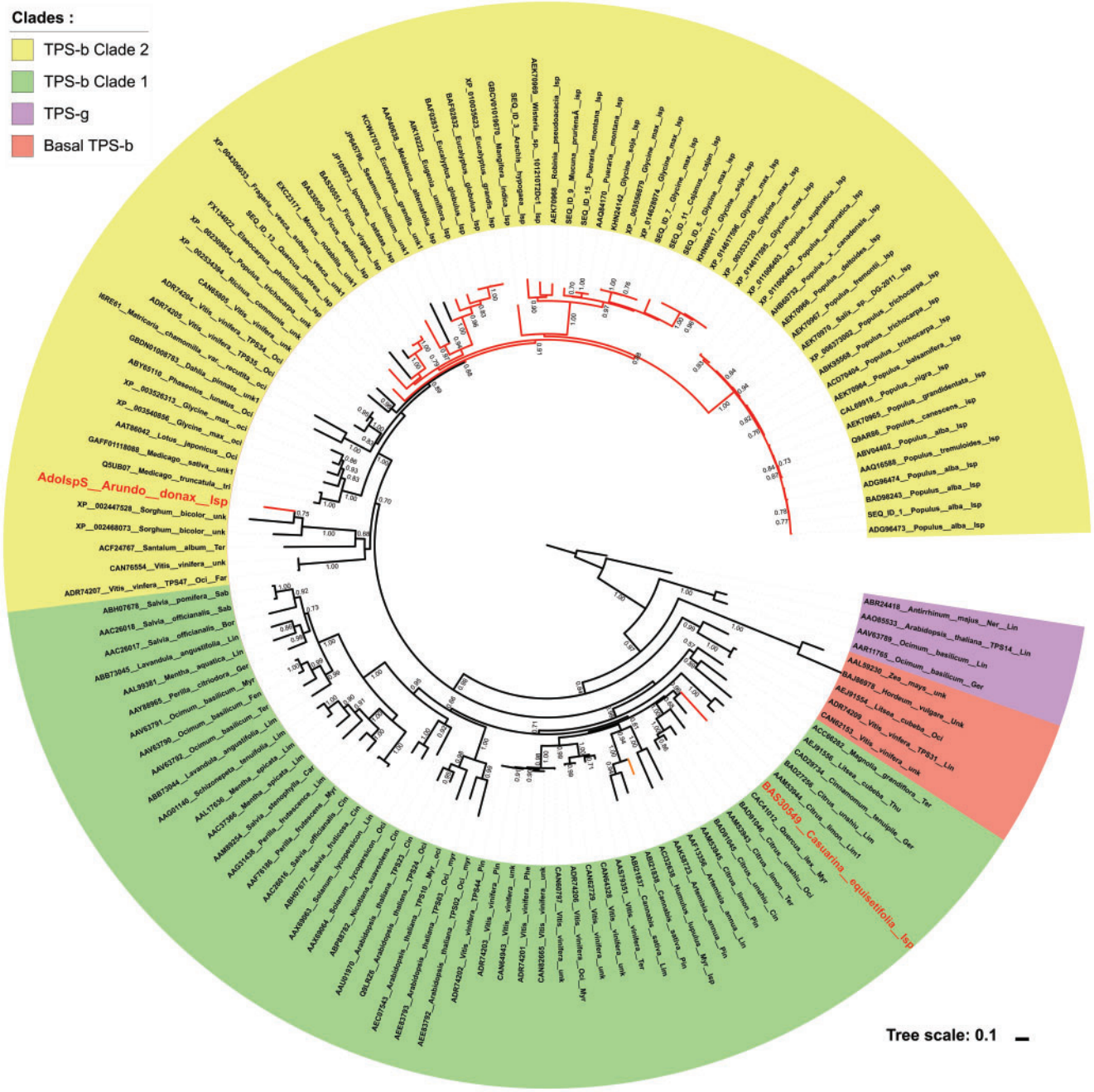

Fig. 3. Phylogenetic reconstruction of class b terpene synthases and some representatives from class g with maximum likelihood. The branches of validated proteins with exclusive IspS activity are marked in red, while the Myrcene synthase with minor IspS activity from *Humulus lupulus* is marked in orange. The two names in red indicate IspS enzymes that most likely are the result of parallel evolution with respect to the major clade of "standard" IspS (all from the rosids but for the single fabid protein from *Ipomoea batatas*). For clarity, only SH-like aLRT support values >0.5 are shown.

the 10 families for which sequences are available. With the exception of the very divergent tetrad from *C. equisetifolia* (FVFK), the first three tetrads were repeatedly found in more than one (sometimes distantly related) families, indicating parallel and/or homoplasious evolution. IspS sequence information was used to search by Blast for additional IspS proteins from monocots. Two bona fide IspS proteins from Poaceae species *Phragmites australis* (Genbank accession number GEKX01061072) and *Phyllostachis edulis* (moso bamboo; accession PH01005131G0060; http://www.bamboogdb.org/; last accessed April 18, 2017) were identified which have the same diagnostic tetrad found in *A. donax*. Curated searches in the two fully sequenced palm genomes available to date (*Phoenix dactylifera* and *Elaeis guineensis*) failed to identify homologues of IspS synthase proteins with any of the diagnostic tetrads identified so far.

In Vivo Analysis of Functional Relevance of AdoIspS Diagnostic Tetrad Residues

In order to dissect the contribution of each amino acid of the IspS diagnostic tetrad to isoprene emission, we created by

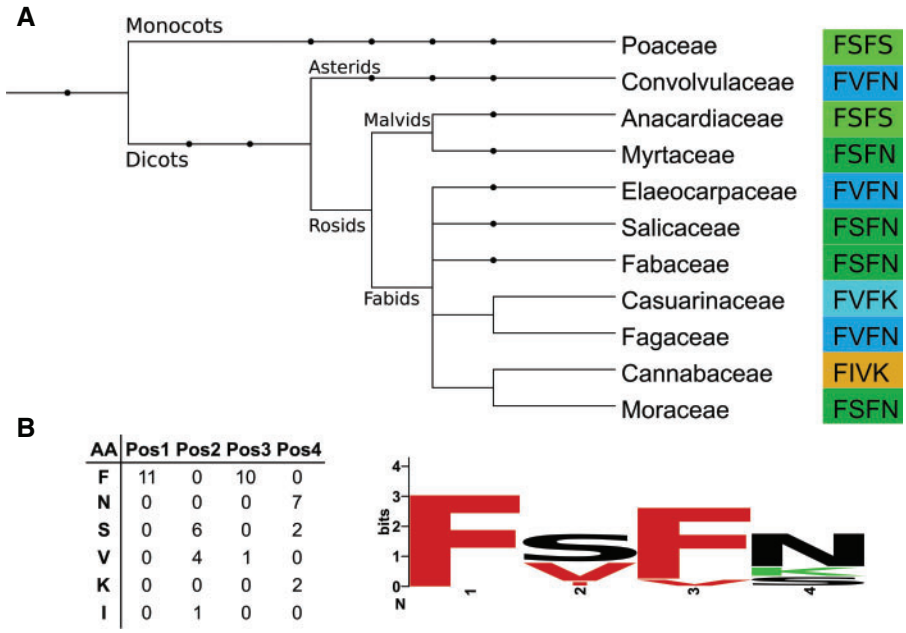

FIG. 4. The extended version of the IspS diagnostic tetrad based on families. (*A*) Phylogram representing the known taxonomy of the angiosperm families from which IspS genes have been validated. The four diagnostic tetrads are color-coded based on their sequence. (*B*) Family-based sequence logo of all IspS validated so far.

site-directed mutagenesis (1) substitutions of each of the four residues to alanine (F310A, S418A, F457A and S477A), (2) the double mutation F310A/F457A, (3) the S477N mutation (which replaces the *A. donax* residue with the most common one present in Salicaceae) and (4) a total of additional 36 mutations independently substituting F310 and F457 with each of the remaining 18 amino acids (saturating mutagenesis). Transgenic lines with any of the single alanine mutants dramatically decreased total isoprene emission by 69–98% with respect to WT IspS levels (fig. 5*A*). In addition, isoprene emission of the S477N mutant was almost totally abolished, indicating that S477 cannot be functionally replaced by its most common amino acid counterpart from dicots. Strikingly, replacement of the first phenylalanine residue with alanine (F310A) caused a dramatic change in total volatile emission, with a reduction of isoprene emission to $<10\%$ of WT AdoIspS and the preponderant production of monoterpenes, demonstrating neofunctionalization of the mutant enzyme to a monoterpene synthase. By contrast, the F457A mutation had a negative effect on both monoterpene and isoprene emission, as indicated by the strong reduction observed in the F310AF457A double mutant (fig. 5*A* and *B*).

The three-dimensional models of AdoIspS WT and F310A protein structure show that the size of the active site of the F310 mutant enzyme is increased by comparison to that of the WT protein, where the protrusion of F310 aromatic ring causes a significant reduction of the space to accommodate the substrate (fig. 6). To systematically evaluate in vivo the role of active site reduction in determining enzyme activity, we performed the saturating site-directed mutagenesis of each phenylalanine residue (F310 and F457).

In general, isoprene emission was more tolerant at both positions to substitution of residues with apolar and large side chains, with the exception of F457N, which emitted amounts of isoprene similar to those of WT AdoIspS (fig. 7*A*). In addition, the concomitant or exclusive in vivo neofunctionalization to monoterpene synthase was significantly detected only in F310 mutants (especially by alanine replacement and to a minor extent also glycine, serine and cysteine; fig. 7*B*). The majority of mutations abolished both isoprene and monoterpene emission (fig. 7*A* and *B*).

To determine the specific monoterpenes emitted by transgenic plants overexpressing p35S:AdoIspS neofunctionalized mutants, GCxGC-ToF-MS analysis of the spectrum of terpenes using the fragments of mass 93, 121 and 136 Da identified several peaks with large quali-quantitative differences with AdoIspS-WT lines (supplementary figs. S12–S15, Supplementary Material online). The large majority of the F310A volatile compounds were acyclic monoterpenes and none of them resulted from carbocation hydroxylation with the exception of negligible amounts of linalool produced also in the control line (Degenhardt et al. 2009). The monoterpene emitted in largest quantity in the neofunctionalized AdoIspS lines was ocimene, which is below detection level in transgenic lines expressing WT AdoIspS (fig. 8 and supplementary fig. S15, Supplementary Material online). While the presence of neo-allo-ocimene isomers in both the commercial standard and the samples from mutant lines are almost certainly due to conversion from beta-ocimene due to thermal desorption of volatiles from the SPME fiber (Chamorro et al. 2014) (supplementary figs. S12 and S13, Supplementary Material online), alpha-ocimene as well as the other unidentified ocimene isomers are bona fide minor products of the mutant enzyme. The only other acyclic monoterpene, beta myrcene, was found in minor amount, while the cyclic monoterpenes gamma terpinene and terpinolene amounted together to

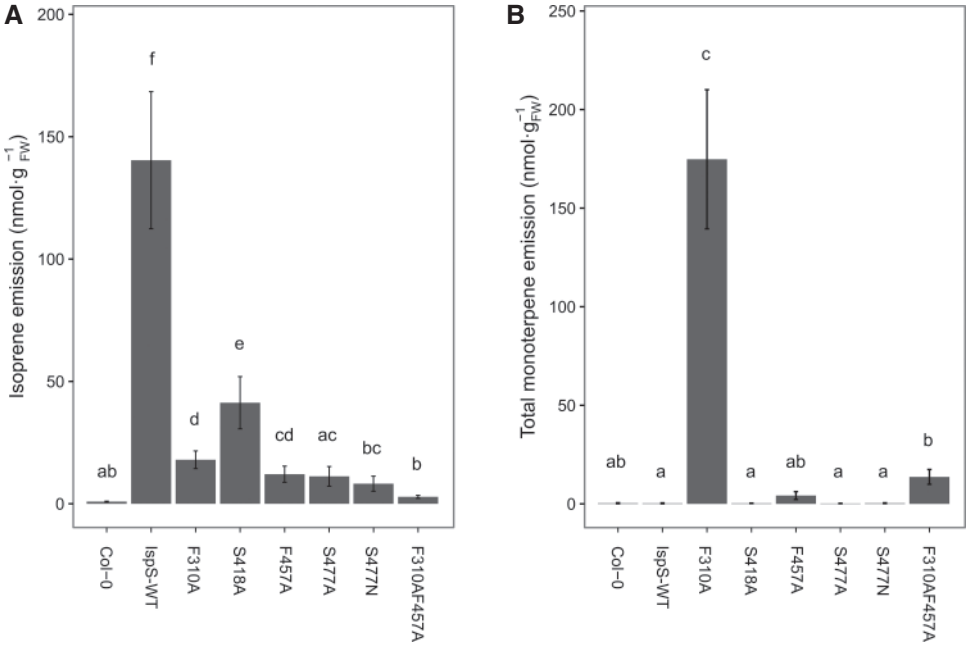

Fig. 5. Volatile emissions from transgenic *Arabidopsis* lines transformed with WT *AdoIspS* construct and the six independent site-directed mutageneses F310A, S418A, F457A, S477A, S477N, F310AF457A. (*A*) Isoprene emission. (*B*) Monoterpene emission. Letters on top of the bars are the result of one-way analysis of variance (ANOVA) tests. Different letters indicate statistically different means based on Tukey's test with $P < 0.05$.

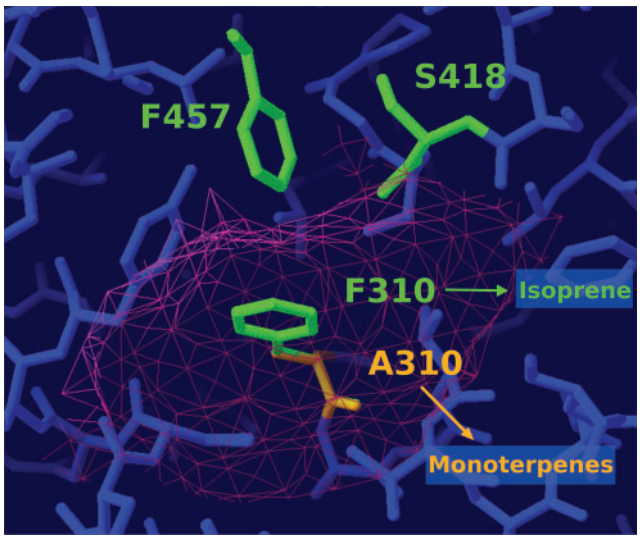

Fig. 6. Superposition of structural models of AdoIspS WT and F310A mutant. Modeling was carried out using the SWISS-MODEL server based on the crystal structure of *Populus x canescens*. The resulting models were superposed using the DeepView program. Three of the amino acids of the WT diagnostic tetrad (F310, S418 and F457) are shown in green, while the F310A mutation is shown in orange. Part of the surface of the F310 A mutant active site is shown as a pink network. The aromatic ring of F310 (present in the WT enzyme) clearly protrudes in the middle of the active site cavity, thus causing its size reduction.

<10% of the total monoterpenes (fig. 8). The major monoterpenes emitted were the same in all neofunctionalized lines (supplementary fig. S15, Supplementary Material online). Thus, the F310A, F310G, F310C, and F310S mutations cause the same in vivo neofunctionalization of AdoIspS, converting it into a functional ocimene synthase.

Discussion

AdoIspS Is the First Isoprene Synthase Gene Identified from Monocots

Several isoprene synthase genes have been characterized in dicots, especially from the rosids (Sharkey et al. 2013). The *AdoIspS* gene is the first *IspS* characterized from a monocot species. From a functional point of view, several features of *AdoIspS* are similar to *IspS*s from dicots. The *AdoIspS* transgene is both necessary and sufficient to confer the ability to synthesize isoprene to the nonemitter *Arabidopsis* (Sharkey et al. 2005; Sasaki et al. 2007; Vickers et al. 2009). Isoprene synthases catalyze the formation of isoprene only when localized in the chloroplast (Vickers et al. 2011). Thus, the chloroplast transit peptide of the monocot *A. donax* is sufficient for the correct targeting of the mature peptide to the chloroplast also in *Arabidopsis*. No emission of other volatile compounds besides isoprene was observed in *Arabidopsis* plants transformed with WT *AdoIspS*, indicating that in the physiological conditions tested the encoded protein is a highly specific isoprene synthase like poplar and kudzu *IspS* (Sharkey et al. 2005; Loivamäki et al. 2007; Schnitzler et al. 2010). The genomic structure and the number of exon–introns are highly conserved to those of dicots, and also expression of the gene in *A. donax* is detected nearly exclusively in photosynthetic tissues, and especially the leaf blade (Sasaki et al. 2005; Sharkey et al. 2005). Like dicot *IspS* (Sasaki et al. 2005), also *AdoIspS* expression is mainly responsive to heat stress, but, to a lower extent, also to heavy metal and drought/osmotic stresses, which matches responses of isoprene emission to stress observed in nature (Loreto and Schnitzler 2010). Given the difficulties we encountered in expression of a functional protein in bacteria, the in-depth comparison of AdoIspS

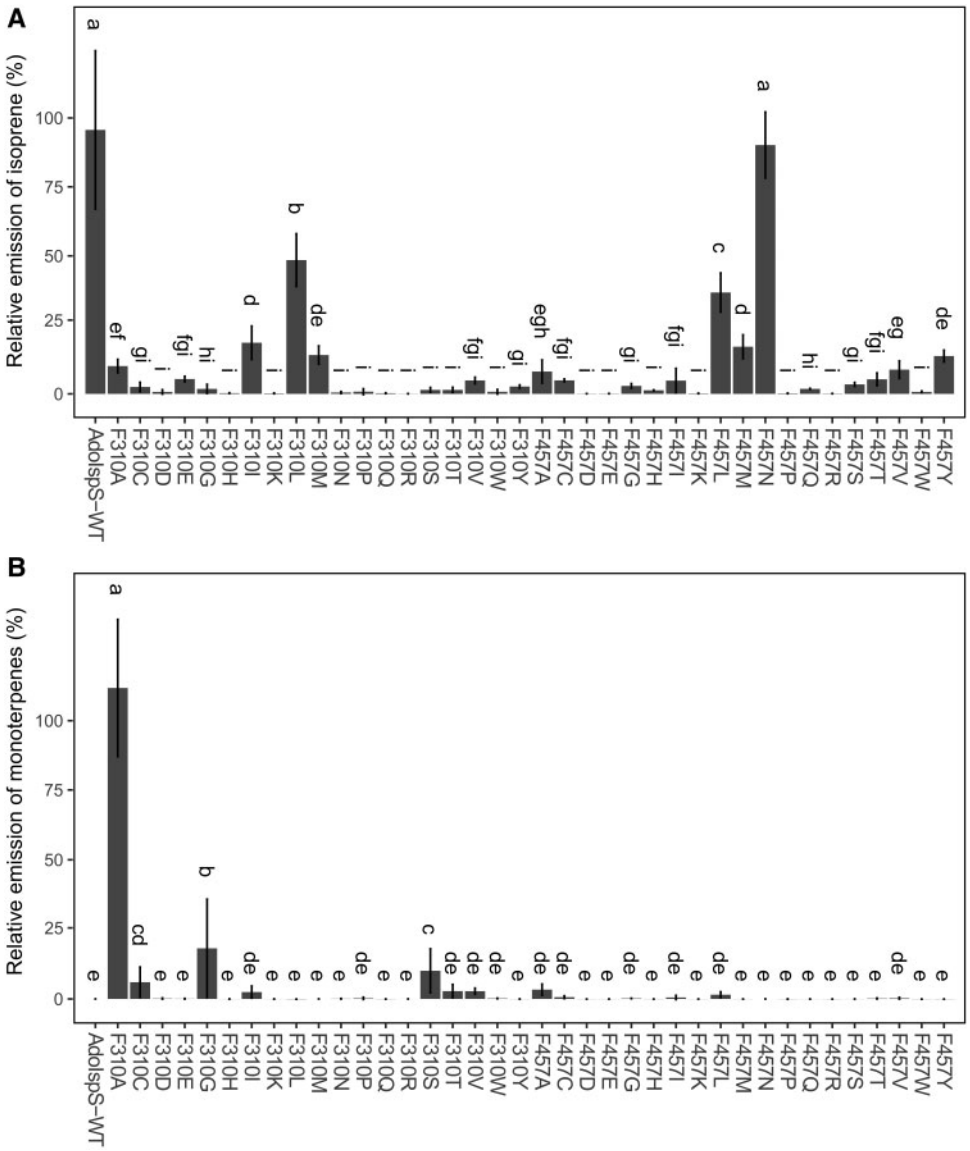

Fig. 7. Relative volatile emissions from transgenic *Arabidopsis* lines transformed with WT *AdoIspS* and 38 saturating mutagenesis mutants for diagnostic residue F310 and F457. (*A*) Relative isoprene emission. (*B*) Relative monoterpene emission. Relative emission level was calculated as the ratio of net emission (obtained by subtracting background: Emission level of Col-0 for isoprene emission and WT AdoIspS_79 for monoterpene emission, respectively) divided by the isoprene emission of the internal reference (WT AdoIspS_79 in the same PTR-ToF-MS run on the same day). Average of the six top ranking emitters either for isoprene or monoterpene emission out of 30 T2 transgenic lines measured were used for this analysis. Letters on top of the bars are the result of one-way analysis of variance (ANOVA) tests. Different letters indicate statistically different means based on Tukey's test with $P < 0.05$. The absolute values corresponding to 100% average ($\pm$ SD) emissions of isoprene and monoterpene were, respectively, 355.2 $\pm$ 106.5 and 428.6 $\pm$ 94.5.

enzymatic activity with other IspS awaits validation of a suitable expression system of the tagged enzyme. However, the results presented demonstrate that *AdoIspS* shares all major functional features with dicot *IspS*s.

The *IspS* gene from *A. donax* fills a large gap in the taxonomic sampling of validated *IspS*, as it is to date the single-product *IspS* gene from the most taxonomically divergent lineage with respect to the main IspS rosids clade. It has been previously hypothesized that five to six independent origins of IspS may have occurred in angiosperms, one of them being in commelinids (encompassing both Poales and Arecales orders; Sharkey et al. 2013). The separation of AdoIspS from the main rosids IspS clade supports parallel IspS evolution in Poaceae and dicotyledons. Despite divergent evolution cannot be completely ruled out, a parallel evolution scenario seems to be more consistent with the small number of phylogenetically related species emitting isoprene in Poales (Sharkey et al. 2013), as well as with their highly derived position among monocots (Chase et al. 2016). The recent isolation of *Casuarina equisetifolia* IspS, a species from the rosid order fagales, constitutes an independent evidence of parallel evolution of single-product IspS based on the clearly polyphyletic origin of fagales IspS (Monson et al. 2013; Oku et al. 2015). Taken together, these results suggest that isoprene emission independently evolved at least three times in angiosperms, twice in rosids and once in Poaceae.

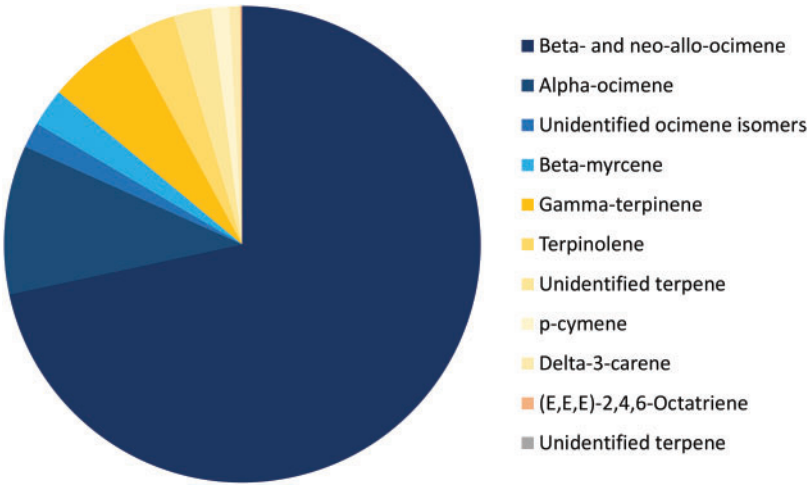

Fig. 8. Relative amounts of the major monoterpenes emitted by F310A AdoIspS. Acyclic monoterpenes are indicated in blue, while cyclic monoterpenes in different hues of orange and red. The unidentified monoterpene (gray) is not visible because present in extremely low amounts.

A Possible Role for S418 and S477 in Carbocation Stabilization and Active Site Closure

Our revision of the IspS diagnostic tetrad using a family-based normalization captures the relevant variation of amino acid combinations among the angiosperm single-product IspS known so far, indicating that positions 2 and 4 are subject to a more relaxed selection with respect to specification of IspS activity. Lack of full conservation has been recently interpreted as indication of accessory function for these residues (Ilmén et al. 2015). Relative emissions of AdoIspS S418A and S477A mutants in vivo demonstrate the essential function of S477, suggesting a possibly less fundamental, although relevant, role of S418 in IspS activity.

The region corresponding to the serine triplet centered around AdoIspS position 418 has been recently identified (Schrepfer et al. 2016) as part of a triad of amino acids responsible for a mechanism of induced fit substrate ionization in taxadiene synthase (TXS). Only minor movements of S418 and neighbouring residues can be inferred by the structural model obtained from PcIspS and this mechanisms of substrate ionization is in contrast to the proposed role of DMADP diphosphate as the general base for carbocation formation (Köksal et al. 2010). The carbonyl oxygen atom of S418 is, however, in closest proximity to the $C3^1$ and C4 methyl groups of the substrate analogue DMASPP, suggesting that it may contribute to the stabilization of the geranyl and/or linalyl carbocation (Köksal et al. 2010). Consistent with the possible involvement of the S triplet in carbocation stabilization, the region homologous to AdoIspS S417-L422 has been shown to contribute to product specificity in monoterpene synthases (Kampranis et al. 2007), as suggested by the deformation of the G1/G2 helix-break motif observed in several TPS upon transition to the closed conformation (Baer et al. 2014). This model is compatible with the Ser-Val variability observed in isoprene synthases, indicating that side chain size may be more relevant than polarity for optimal positioning/function of position 418 alpha carbonyl oxygen.

S477 is the last amino-acid of the H-α1 loop, which is involved in capping of the active site of several TPS (Whittington et al. 2002). Despite mutation of AdoIspS S477 structural homologue in MdAFS1, an apple α-farnesene synthase, abolishes K^+ dependence and phenocopies the corresponding AdoIspS mutation (Green et al. 2009), a primary role of S477 in K^+ coordination seems unlikely because this ion enhances, but is not essential for IspS activity (Köksal et al. 2010). Bornyl synthase K512 residue, the structural homologue of AdoIspS S477, donates a hydrogen bond to the diphosphate group of the substrate (Whittington et al. 2002), indicating that S477 may instead contribute to diphosphate positioning/sensing during transition to the closed conformation.

These observations suggest that S477 may be part of a mechanism of substrate sensing and conformational remodeling, associated to transition to closed conformation, ancestrally shared by several TPS enzymes (Whittington et al. 2002). Crystallographic data indicate that in isoprene synthases the transition from open to close conformation may not happen. However, it cannot be excluded that structural changes may have been prevented in vitro by molecule packing in the crystal lattice (Köksal et al. 2010), leaving open the possibility of a so far undetected close state transition. Additional studies will, thus, be required to elucidate whether this hypothesis is supported by in vivo evidence, and the eventual significance of this mechanism in IspS function.

A Pivotal Role of F310 in the Evolution of Short-Chain TPS

Saturating mutagenesis of positions 310 and 457 uncovered several amino acids compatible with detectable levels of isoprene emission, but only one of the possible combinations among them (F310/F457) has been found in naturally occurring single-product IspS enzymes to date. This result is consistent with the fact that only a minority of the possible mutational paths in the fitness landscape are accessible to Darwinian selection of protein function due to pleiotropy

(Weinreich et al. 2006). The high proportion of substitutions abolishing or drastically reducing both isoprene and ocimene synthase emission (indicating that they affect the overall structure/stability of the enzyme) supports this view.

Previous studies predicted that both F310 and F457 residues could decrease the size of the substrate binding pocket and make van der Waals contacts with the substrate DMADP, thus determining isoprene production instead of monoterpene synthesis from GDP (Sharkey et al. 2005; Köksal et al. 2010). Our results demonstrate in vivo that only position 310 plays a fundamental role in AdoIspS substrate specificity determination by active site size modulation. In the lack of data from purified enzymes, the relative entity of the ocimene and isoprene enzymatic activities underlying the dramatic changes in volatile emission observed in transgenic plants harboring the AdoIspS F310A mutant is currently unknown. Due to the limited pool of isoprene and ocimene precursors in *Arabidopsis*, neofunctionalization and the consequent activation of a novel sink for GDP may alter in nonlinear ways the amounts of AdoIspS F310A substrates, which in addition could compete with each other for active site binding. Strikingly, also single mutations of AdoIspS F310 homologue in widely divergent cineole synthases from plants (*Salvia officinalis*) and fungi (*Hypoxylon* sp.) shifted in vitro substrate specificity from GDP to FDP, causing neofunctionalization from monoterpene to sesquiterpene synthase activity (Kampranis et al. 2007; Shaw et al. 2015). *Salvia officinalis* cineole synthase N338A mutant, the structural homologue of AdoIspS F310A, displayed a nearly 200-fold decrease of the catalytic efficiency of the enzyme when using FDP instead of GDP as only substrate (Kampranis et al. 2007). Unfortunately, neither enzymatic activities were tested with different ratios of both substrates, nor was the impact of this change assessed in vivo in terms of total volatile emission. The actual enzymatic underpinnings of the preponderant ocimene emission from AdoIspS F310A, thus, remain an open question awaiting the establishment of a suitable expression system. Irrespective of the enzyme kinetics details, these results strongly support that the very same amino acidic position is involved in substrate specificity determination of different short-chain TPS-b (hemiterpene, monoterpene and sesquiterpene synthases) and that in the case of isoprene synthases this single mutation has a large effect on the total emission of volatiles in vivo. This plasticity determinant is in common between the monocotyledonous and dicotyledonous lineages. Additional evidences, however, indicate that the pivotal role of this position in plants likely extends even further, to the gymnosperm lineage. F310, in fact, is structurally homologous to the phenylalanine hypothesized to determine substrate specificity of a conifer methylbuthenol synthase (Gray et al. 2011), strongly supporting that the independent evolution of hemiterpene synthases in gymnosperms and angiosperms relied on the parallel acquisition of a phenylalanine residue in the same position of the class I terpenoid synthase fold helix D. Analogously, substrate and product specificity has been demonstrated to rely on this very same residue in the parallel evolution of cineole synthases in plants and fungi (Kampranis et al. 2007; Shaw et al. 2015). By contrast to the double function in substrate preference and product specificity of cineole synthases, the saturating mutagensis of F310A indicate that in AdoIspS the F310 residue is involved in determination of product specificity as a consequence of substrate preference, as the same products are present in all neofunctionalized mutants.

The substrate actually used by some TPS enzymes can be up to one isoprenoid unit shorter than the one potentially fitting into the active site cleft (Köksal et al. 2011). This suggests that current substrate preference is the relatively recent result of one/few key mutations preventing accommodation of a larger, ancestral substrate for which the overall active site cavity architecture/size was optimized in the course of evolution (Pazouki and Niinemets 2016). As the architecture and maximal size of the active site depends from many residues, these features are expected to evolve slowly and to constrain the effect of mutations of plasticity residues, which, by definition, are major switches for evolutionarily fast neofunctionalization (Yoshikuni and Keasling 2007). Analogously to the "molecular ruler" mechanism of prenyl transferases (Wang et al. 2016), thus, we propose that, in class I TPS, plasticity towards different substrates relies mainly on the geometric constraints imposed by α-helical modularity and by the distance from the site of coordination of the magnesium triad. The side chain of any amino acid like F310, laying exactly two helical turns below (i.e., seven positions before) the first aspartic residue of the DDXXD motif, is exposed to the active site hydrophobic cleft, and defines, according to its size, the possibility to accommodate in a conformation compatible with catalysis DMADP versus GDP or GDP versus FDP as substrate. This structural modularity confers to terpene synthases a surprising capacity to revert to original functions, which is not always shared by other enzymes (Kaltenbach et al. 2015).

Based on this general mechanism, the preponderant production of ocimene in all AdoIspS F310 neofunctionalized mutants compared with the very low amounts of myrcene and other monoterpenes confirms the analogy between isoprene and ocimene reaction chemistries (Sharkey et al. 2013). This in turn points to ocimene synthases as the most likely ancestral TPS from which IspS evolved by active site size reduction following acquisition of F at position 310 (Gray et al. 2011) (fig. 9). Given the intrinsically high capacity of TPS to evolve new functions (Christianson 2008), discerning past evolutionary relevance from evovability potential of plasticity residues identified exclusively by structurally informed choices is generally difficult. The choice of the residues to mutagenize in this study, on the contrary, was evolutionarily informed, as it relied on the conservation of the four residues that are preferentially (diagnostic tetrad positions 2 and 4) or exclusively (diagnostic tetrad positions 1 and 3) present together in isoprene synthases, but in no other known terpene synthase (Sharkey et al. 2013). The saturating mutagenesis and functional characterization in vivo of the two F residues, thus, amounts to the closest experimental recapitulation of the mutational landscape faced during IspS gene evolution by plant species lacking the trait, in line with Jensen's hypothesis and many directed mutagenesis experiments (Khersonsky and Tawfik 2010). The strong in vivo reduction of the original

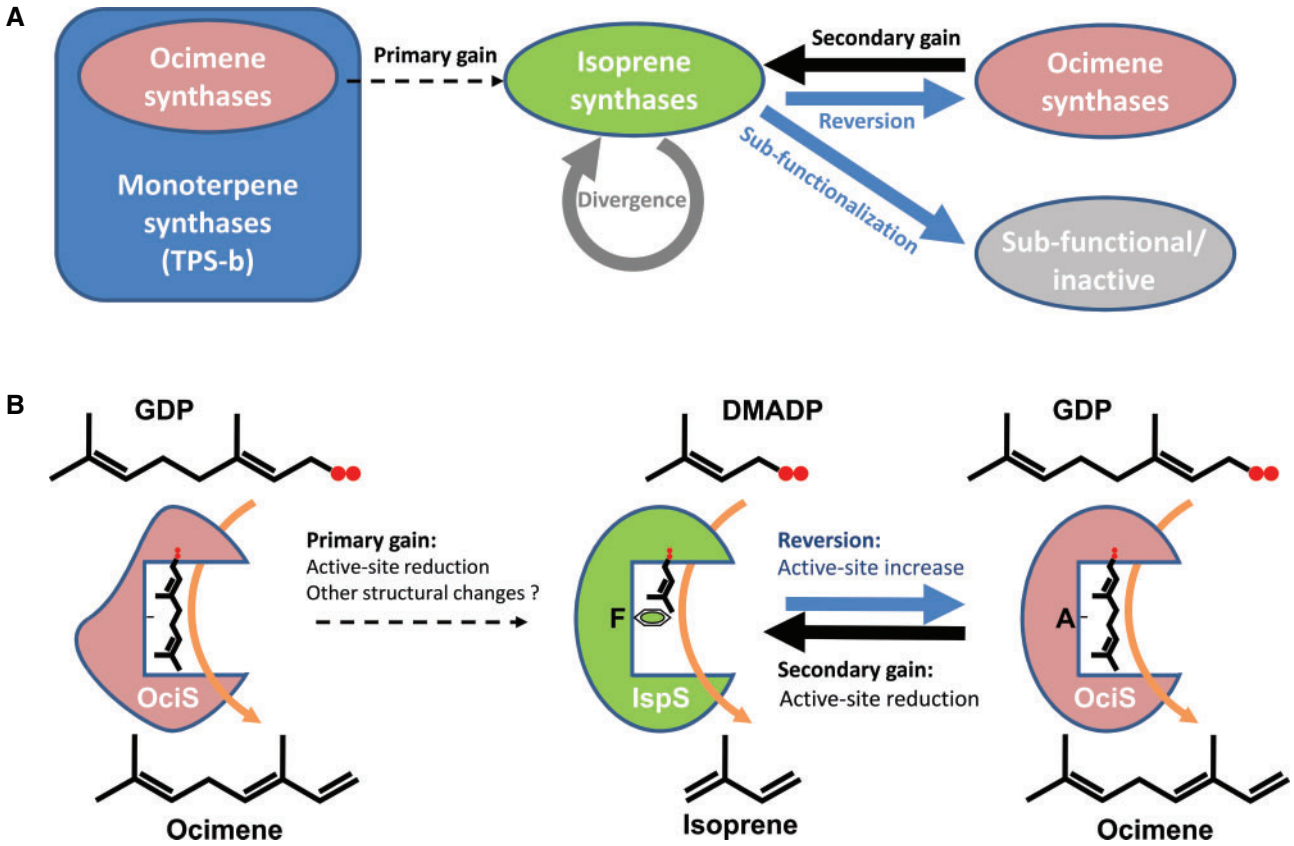

Fig. 9. Model of gain/loss of isoprene synthase function in angiosperms. Solid arrows indicate experimentally supported processes, dashed arrows postulated ones. (*A*) *IspS* genes are proposed to have evolved multiple times from ocimene synthases (TPS-b clade), but the rate of first recruitment from the pool of pre-existing ocimene synthases is currently unknown. It is postulated that primary gains are relatively rare (dashed black line), but once an *IspS* gene has been acquired in a clade, single mutations can switch the enzyme substrate specificity back and forth between isoprene and ocimene synthase by maintaining active site overall structure, potentially leading to repeated losses and secondary gains of the trait in the same clade (blue arrow "Reversion" and black arrow "Secondary gain"). Additional mutations can lead to either divergence or can hit functionally fundamental amino acids leading to sub-functionalization (or complete protein inactivation). (*B*) Molecular mechanism for the evolution of isoprene synthases supported by the results of this work. Both primary and secondary gains of isoprene emission are associated with a mechanism of active site reduction (depicted as a shallower cavity in the enzyme) by replacement of the seventh position before the first D of the DDXXD motif (AdoIspS position 310) with the bulky phenylalanine residue (F). The reduction of the size of active site prevents the efficient use of geranyl pyrophosphate (GDP) as substrate for the production of ocimene, causing instead the conversion of dimethylallyl pyrophosphate (DMADP) to isoprene. Additional mutations leading to putative structural changes may be required for the first recruitment (dashed arrow) of an Ocimene synthase (OciS) from the pool of TPS-b genes present in each species. Enzymatic activity can potentially revert to ocimene synthesis by a second substitution to the same position with a smaller residue, which is preferentially alanine (A) in AdoIspS.

isoprene emission, maintained in the F310A lines as promiscuous function, indicates strong negative trade-offs in the evolution of IspS and rapid fixation of the trait, suggesting that the adaptive driver(s) of trait evolution may have concomitantly selected in favor of the new function and against the original one (Khersonsky et al. 2006; Khersonsky and Tawfik 2010). As the mutagenesis has been carried out on an extant IspS gene, a limitation of the approach used in this study is that it cannot take into account mutations with functional relevance that may have happened after the gain of IspS function. Several other evidences, however, independently corroborate derivation of isoprene synthases from ancestral ocimene synthases. IspS active site volume is as large as that of other monoterpene synthases and it was demonstrated that GDP competes with DMADP for binding to IspS active site (Köksal et al. 2010, 2011), indicating, for the reasons provided above, that the ancestors of IspS probably were monoterpene synthases (Dani et al. 2014) using GDP as substrate. In addition, the catalysis of IspS is inefficient compared with monoterpene synthases, further implying that IspS specificity for DMADP may be a derived trait (Lerdau and Gray 2003). Finally, evolution of IspS from ocimene synthases ancestors is also supported by the derived position of all known IspS genes with respect to other terpene synthases in the TPS-b clade (Sharkey 2013; Sharkey et al. 2013) (fig. 3) and by the monophyly of isoprene and ocimene synthases in the TPS-b2 clade (Sharkey et al. 2013).

With the notable exception of F457N, the tolerance of isoprene activity to amino acid replacements at position 310 and 457 is similar. F457, however, is basically unable to promote significant in vivo neofunctionalization. The different plasticity of the two residues (Yoshikuni and Keasling 2007), indicates that acquisition of an F at position 310 most likely constituted a more relevant constraint in the

evolution of IspS activity compared with appearance of F457. This conclusion is supported by two evidences. First, the secondary neofunctionalization to isoprene synthase activity detected in a myrcene synthase from hops is associated to the exclusive presence of the structural homologue of F310, but not of F457 (Sharkey et al. 2013). Second, relative occurrence in TPS-b monoterpene synthases of F in correspondence of position 457 is higher than for position 310, indicating a lower selectivity of position 457 towards the amino acid side chain.

A Unified Model for *IspS* Evolution

Taken together, the results presented above reconcile the apparently contrasting evolutionary patterns previously observed (Monson et al. 2013; Sharkey 2013; Sharkey et al. 2013) in the unified conceptual framework summarized in figure 9. On one hand, as suggested by the relatively low number of independent gains estimated in this study, primary acquisition of IspS activity seems to be a relatively complex and rare event. All amino acid replacements at position 310 produced low amounts of monoterpenes derived from hydroxylation of geranyl or linalyl cation intermediates (Degenhardt et al. 2009). Even mutation F310N, replacing Phe with the Asn responsible of coordination of the water molecule which hydroxylates the substrates in both plant and fungal cineole synthases, does not produce hydroxylated products, being instead basically inactive. This indicates that, by contrast to cineole and BPP synthases, reduced hydroxylation ability due to lack of the cavity necessary for water coordination at the bottom of the active site is ancestral for isoprene as well as ocimene synthases (Kampranis et al. 2007; Köksal et al. 2010). The existence of clear evolutionary constraints related to the overall active site cavity volume and the ability to catalyze hydroxylation and efficient cyclization of carbocation intermediates and the phylogenetic evidences, thus, support that all rosids and possibly fabids IspSs are the likely result of a single, primary gain of IspS function. On the other hand, the potential ease of in vivo interconversion between ocimene and isoprene synthases by single mutations demonstrated in this study provides a rationale contributing to explain the repeated evolutionary loss and especially re-gain of IspS activity at the family and genus level (Monson et al. 2013) after any primary acquisition of the trait (fig. 9).

As discussed earlier, the insights provided in this study through an evolutionarily informed functional analysis of the residues preferentially or exclusively present in isoprene synthases are a good starting point, but for sure not an exhaustive elucidation of the factors contributing to IspS evolution. As recently pointed out, the selective drivers underlying evolution of the trait as of today are still unknown and their identification will be definitely required to dissect the details of lineage-specific conservation or loss of isoprene emission (Sharkey and Monson 2017). The findings of this work, by providing a mechanistic basis for the structural constraints controlling IspS and more generally TPS evolution by neofunctionalization, and identifying the likely ancestors of IspS genes, will hopefully help to stimulate novel research lines focusing on the possible adaptive trade-offs that may exist between emission of evolutionarily sister volatile compounds, like isoprene and ocimene.

Isoprene synthase is the single enzyme responsible for the yearly emission worldwide of ~600 Tg of isoprene, the most abundant biogenic volatile organic compound and possibly the one with the largest impact on global atmospheric chemistry (Sindelarova et al. 2014). Stunningly, the molecular mechanism at the root of the evolution of such hugely important ecological phenomenon has been traced back to the mutation of a single amino acid responsible of active site size and substrate specificity modulation. The elucidation of the likely evolutionary path of isoprene synthase provides a relevant example of how parallel evolution may be driven by relatively simple biophysical constraints and underlines the power of the emerging paradigm of evolutionary biochemistry to define the intimate molecular evolutionary links between the structural and functional bases of traits with global relevance (Harms and Thornton 2013).

Materials and Methods

Plant Materials, Growing Condition, and Stress Treatments

Rhizomes and cohorts of *A. donax* cuttings (collected in Sesto Fiorentino, Florence, Italy 43°49′01.8″N 11°11′57.0″E) were used in this study. The plants generated from rhizomes were grown under the same condition as described previously (Sablok et al. 2014). *Arundo donax* cuttings were used for different stress treatments under the same growing condition used in earlier studies (Fu et al. 2016). When the new shoots reached the five-leaf stage, they were subjected to different stress treatments. The stress treatments performed were: water limitation (addition of 15% PEG 6000 to the basal hydroponic solution), salt (addition of 150 mM NaCl), heavy metal (addition of 500 μM $CdSO_4$), heat (42 °C) for 24 h. The entire shoots (treated and nontreated) were collected separately at different time points (0.5, 1.5, 3, 6, 11 and 24 h after inducing the stress), snap-frozen in liquid N_2 and stored at −80 °C till use. At least three plants per biological replica and three biological replicas per treatment and respective controls were sampled. For volatile compound analyses, transgenic and wild-type *Arabidopsis* plants (Col-0) growing at 23 °C in the growth chamber under standard long-day condition (16 h light/8 h dark) were used.

RNA Extraction and RT PCR Analyses

Total RNA was extracted from different organs of *A. donax* under different stress conditions with Trizol (Invitrogen) according to the manufacturer's instruction, and treated with Dnase I (Rnase-free) (Sigma–Aldrich) to eliminate genomic DNA contamination. One μg of total RNA was reverse transcribed with Superscript III reverse transcriptase (Invitrogen). The synthesized cDNA was used as a template for semi-quantitative and real-time PCR analyses. Semi-quantitative rtPCR was performed for different organs of *A. donax* using the glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene as reference. For stress treatments, real-time PCR was carried out with Platinum SYBR Green qPCR

SuperMix-UDG (Invitrogen) in a Bio-Rad C1000 Thermal Cycler detection system using a putative *A. donax* actin (Fu et al. 2016) and *Arabidopsis AtActin2* (Lim et al. 2014) as a reference gene for *A. donax* and transgenic *Arabidopsis* lines, respectively. All reactions were performed in triplicates and the $2^{-\Delta\Delta CT}$ method was used to calculate fold changes. Significance of expression difference was calculated on Δct averages of three biological replicates based on the Student's *t*-test using Holm's correction for multiple tests ($n = 6$, $\alpha = 0.05$). Primers used for PCR are listed in supplementary table S2 in Supplementary Material online.

Cloning and Plant Transformation

The full-length cDNA of *AdoIspS* was amplified with primers *AdoIspS_For* and *AdoIspS_Rev* (supplementary table S2, Supplementary Material online) using Phusion High-Fidelity DNA Polymerase (NEB). The purified amplicons were cloned into Gateway vector pENTR/D-TOPO (Invitrogen) to yield pENTR_AdoIspS. A total of 42 mutations were introduced into *AdoIspS* WT CDS using the Quikchange Site-Directed Mutagenesis Kit (Stratagene) in the *A. donax* pENTR_AdoIspS plasmid: The four amino acids corresponding to the "IspS diagnostic tetrad" previously identified (Sharkey et al. 2013) (F338, S446 [identified based on its being in "the middle of a triple serine" and being "most commonly Val or Ile in other proteins," but erroneously indicated as S445 in the original publication], F485 and N505 of PcIspS) are F310, S418, F457 and S477 in AdoIspS. Saturating mutagenesis was applied for both F310 and F457, (respectively, mutations F310A, F310C, F310D, F310E, F310G, F310H, F310I, F310K, F310L, F310M, F310M, F310N, F310P, F310Q, F310R, F310S, F310T, F310V, F310W, F310Y; and F457A, F457C, F457D, F457E, F457G, F457H, F457I, F457K, F457L, F457M, F457M, F457N, F457P, F457Q, F457R, F457S, F457T, F457V, F457W, F457Y). Furthermore, S418 and S477 were independently mutagenized to alanine (S418A and S477A), and the monocot-specific S477 residue was replaced also with asparagine (S477N); finally, the double mutant F310A/F457A was constructed. All these 42 pENTR clones with IspS diagnostic tetrad mutations plus wild-type pENTR_AdoIspS (AdoIspS-WT) were further recombined in the destination vector pK7WG2 (Karimi et al. 2002) to generate the final constructs through LR reaction with LR clonase II (Invitrogen). These constructs were transformed independently into *Agrobacterium tumefaciens* strain GV3101-pMP90RK by electroporation and further transformed into *Arabidopsis thaliana* Col-0 ecotype by the floral dip method (Clough and Bent 1998). Transgenic plants were obtained by screening sterilized seeds on solid MS (Murashige and Skoog) medium supplemented with 50 mg l^{-1} of kanamycin. The primer sequences used for construct preparation are listed in supplementary table S2 in Supplementary Material online.

Genotyping

Segregation analysis was carried out on T2 segregating plants transformed with p35S:AdoIspS (AdoIspS-44 line). T2 seeds stratified at 4 °C for three days were sown directly in pots filled with commercial soil GS90L. Four-week-old plants were used for both isoprene emission measurement (see below) and genomic DNA extraction with the CTAB method (Doyle 1987). DNA quality and presence of the *AdoIspS* transgene were assayed by PCR using primer pairs *AtActin2_For* + *AtActin2_Rev* and *AdoIspS_Rev* + *pK7WG2_2S*, respectively. Thermal cycling was carried out as follows: 95 °C for 2 min, 35 cycles at 94 °C for 40 s, 60 °C for 30 s and 72 °C for 2 min. Primer sequences are listed in supplementary table S2 in Supplementary Material online.

PTR-MS Measurements

One young and mature rosette leaf for each genotype (Col-0 and transgenic lines) was cut and transferred into a 20-ml glass vial containing 300 µl of distilled water. The vials were kept open for at least 30 min and immediately closed with metal lids sealed with silicon septa (Agilent), and then incubated for 3 h at 30 °C with light intensity of 130–150 µmol m^{-2} s^{-1}. The measurement was done with a commercial PTR-ToF 8000 apparatus from Ionicon Analytik GmbH, Innsbruck (Austria). The data of each measurement were extracted and converted to parts per billion (ppb) following the methodology previously described (Cappellin et al. 2011), and the concentration was converted to flux on a mass basis (nmol g_{FW}^{-1}) by using the corresponding individual leaf fresh weight.

Volatile Extraction Conditions and Analysis

The headspace-solid phase microextraction (SPME) method was applied to extract volatile compounds with slight modification from the former description (Matarese et al. 2014). One gram of leaves pretreated as for PTR-MS measurement was collected and ground with liquid nitrogen, then mixed with 0.3 g of NaCl and 2 ml of freshly prepared citrate-phosphate buffer (0.1 M Na_2HPO_4, 50 mM citric acid, pH 5.0) in 20 ml headspace glass vials. Each sample was spiked with 50 µl of 2-octanol at 1 mg l^{-1} as an internal standard. Three independent biological replicates for each genotype were analyzed to generate each data point.

The headspace was sampled using 2 cm DVB/CAR/PDMS 50/30 µm fiber from Supelco (Bellefonte, PA) equipped with a Gerstel MultiPurpose Autosampler (GERSTEL GmbH & Co. KG, Mülheim an der Ruhr Germany). Samples were kept at 60 °C for 20 min and then extracted for 35 min at 60 °C. The volatile and semi-volatile compounds were desorbed in the GC inlet at 250 °C for 3 min in splitless mode and the fiber was reconditioned for 7 min at 270 °C. The GC × GC system consisted of an Agilent 7890 A (Agilent Technologies, Santa Clara, CA) gas chromatography equipped with a Pegasus IV time-of-flight mass spectrometer (Leco Corporation, St. Joseph, MI). A VF-Wax column (100% polyethylene glycol; 30 m × 0.25 mm × 0.25 µm, Agilent J&W Scientific Inc., Folsom, CA) was used as a first-dimension (1D) column and a RXI-17Sil MS-column 1.50 m × 0.15 mm × 0.15 µm (Restek Bellefonte) as a second-dimension (2D) column.

The GC system was equipped with a secondary oven and nonmoving quadjet dual-stage thermal modulator. The injector/transfer line was maintained at 250 °C. Oven temperature program conditions were as follows: initial temperature of 40 °C for 4 min, ramped at 6 °C min^{-1} to 250 °C, where it

remained for 5 min. The secondary oven was kept 5 °C above the primary oven throughout the chromatographic run. The modulator was offset by +15 °C in relation to the secondary oven; the modulation time was 7 and 1.4 s of hot pulse duration. Helium (99.9995% purity) was used as carrier gas at a constant flow of 1.2 ml min^{-1}. The MS parameters included electron ionization at 70 eV with ion source temperature at 230 °C, detector voltage of 1,317 V, mass range of m/z 35–450 and acquisition rate of 200 spectra per second.

GCxGC-ToF-MS data alignment and processing

ChromaTOF software version 4.32 was used to perform baseline correction, chromatogram deconvolution and peak alignment. The baseline offset was set to 0.8 and signal to noise (S/N) ratio was set at 100. Required match (similarity) to combine peaks was set to 650. Library (NIST and Wiley) selection was conducted for molecular weights between 35 and 350 with five library hits to return. Mass threshold was set at 50 and minimum similarity to assign compound name was set at 700. Identification of some volatile compounds was performed by injecting pure reference standards (beta ocimene). The semiquantitative analyses were based on peak area ratios relative to the internal standard 2-octanol.

Sequence Alignment and Phylogenetic Reconstruction

To elucidate the relationship of *A. donax* IspS with respect to known angiosperm terpene synthases and dicotyledonous isoprene synthases, we aligned the AdoIspS protein and reference IspS sequences from Salicaceae, Fabaceae and Myrtaceae to TPS proteins from three previously published datasets (Sharkey et al. 2013; Ilmén et al. 2015; Oku et al. 2015). A second dataset was created by aligning AdoIspS and all other IspS proteins functionally validated to date to a selection of TPS-b and TPS-g representatives used for the identification and classification of *Eucalyptus* terpene synthases (Külheim et al. 2015). Multiple sequence alignments were carried out with Mafft v.7 using the highly accurate L-INS-i method (Katoh and Standley 2013), removing short sequences to avoid their potentially deleterious effect on the accuracy of phylogenetic reconstruction. To remove potentially disordered/low-homology regions, the multiple sequence alignments encompassing the remaining proteins were processed with GBLOCKS v. 0.91b (Talavera and Castresana 2007), using the options: maximum number of contiguous nonconserved positions = 8, minimum length of a block = 5, and retaining positions with <50% of gaps. We used the SMS (Smart Model Selection; http://www.atgc-montpellier.fr/sms/; last accessed April 18, 2017) web server to identify the optimal model of amino acid substitution for each alignment. Maximum likelihood (ML) phylogenetic reconstruction was carried out with PhyML v.3.0 (Guindon et al. 2010). ML tree heuristic search was carried out with the nearest-neighbor interchanges (NNI) strategy, using both branch length and topology optimization. Branch support was calculated using the approximate Likelihood-Ratio Test (aLRT). Terpene synthase subfamilies were identified according to (Külheim et al. 2015). Bayesian inference (BI) phylogenetic reconstruction was performed with MrBayes v3.2 (Ronquist et al. 2012), using two independent runs, each with one cold and three heated chains over 5,000,000 generations. Trees were sampled every 5,000 generations and posterior probabilities of splits were obtained from the 50% majority rule consensus of the sampled trees, discarding 25% as burn-in. Tree editing was carried out with the iTOL v.3.0 program (Letunic and Bork 2011). Both alignments and the corresponding ML and BI trees are available from TreeBASE (accession numbers S19973 and S19974).

Structural modeling was carried out with the SwissModel server (Arnold et al. 2006).

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We thank Carsten Külheim for providing the sequences of terpene synthases of *Eucalyptus* and other species used as a basis for supplementary figures S8 and S9, Supplementary Material online, and Michele Poli for help with rtPCR analyses. This work was supported by the Autonomous Province of Trento (Italy) through the MAN-VIP project (Team 2011 Call approved with provincial government resolution no. 2902 on 14 December 2010; V.V.) and core funding of the Ecogenomics group (M.L., J.X., E.B. and C.V.), the Volatile Compounds platform (A.A.A., L.C.), the Metabolomics platform (S.C., U.V.). L.C. acknowledges funding from European Union's Horizon 2020 research and innovation programme under grant agreement No 659315. J.X. gratefully acknowledges the financial support from China Scholarship Council (201306300083).

References

Ahrar M, Doneva D, Koleva D, Romano A, Rodeghiero M, Tsonev T, Biasioli F, Stefanova M, Peeva V, Wohlfahrt G, et al. 2015. Isoprene emission in the monocot Arundineae tribe in relation to functional and structural organization of the photosynthetic apparatus. *Environ Exp Bot*. 119:87–95.

Angelini LG, Ceccarini L, Nassi o Di Nasso N, Bonari E. 2009. Comparison of *Arundo donax* L. and *Miscanthus x giganteus* in a long-term field experiment in Central Italy: analysis of productive characteristics and energy balance. *Biomass Bioenergy* 33:635–643.

Archibald AT, Cooke MC, Utembe SR, Shallcross DE, Derwent RG, Jenkin ME. 2010. Impacts of mechanistic changes on HOx formation and recycling in the oxidation of isoprene. *Atmos Chem Phys*. 10:8097–8118.

Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22:195–201.

Baer P, Rabe P, Fischer K, Citron CA, Klapschinski TA, Groll M, Dickschat JS. 2014. Induced-fit mechanism in class I terpene cyclases. *Angew Chem Int Ed*. 53:7652–7656.

Cappellin L, Biasioli F, Granitto PM, Schuhfried E, Soukoulis C, Costa F, Märk TD, Gasperi F. 2011. On data analysis in PTR-TOF-MS: from raw spectra to data mining. *Sensors Actuators B Chem*. 155:183–190.

Chamorro E, Ruiz P, Quijano J, Luna D, Restrepo L, Zuluaga S, Duque-Noreña M. 2014. Understanding the thermal [1s,5s] hydrogen shift isomerization of ocimene. *J Mol Model*. 20:2390.

Chase MW, Christenhusz MJM, Fay MF, Byng JW, Judd WS, Soltis DE, Mabberley DJ, Sennikov AN, Soltis PS, Stevens PF, et al. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc*. 181:1–20.

Chen F, Tholl D, Bohlmann J, Pichersky E. 2011. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J*. 66:212–229.

Christianson DW. 2008. Unearthing the roots of the terpenome. *Curr Opin Chem Biol*. 12:141–150.

Clough SJ, Bent A. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J*. 16:735–743.

Dani KGS, Jamie IM, Prentice IC, Atwell BJ. 2014. Evolution of isoprene emission capacity in plants. *Trends Plant Sci*. 19:439–446.

Degenhardt J, Köllner TG, Gershenzon J. 2009. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* 70:1621–1637.

Doyle JJ. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 19:11–15.

Emanuelsson O, Nielsen H. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol*. 300:1005–1016.

Fu Y, Poli M, Sablok G, Wang B, Liang Y, La Porta N, Velikova V, Loreto F, Li M, Varotto C. 2016. Dissection of early transcriptional responses to water stress in *Arundo donax* L. by unigene-based RNA-seq. *Biotechnol Biofuels* 9:54.

Gray DW, Breneman SR, Topper LA, Sharkey TD. 2011. Biochemical characterization and homology modeling of methylbutenol synthase and implications for understanding hemiterpene synthase evolution in plants. *J Biol Chem*. 286:20582–20590.

Green S, Squire CJ, Nieuwenhuizen NJ, Baker EN, Laing W. 2009. Defining the potassium binding region in an apple terpene synthase. *J Biol Chem*. 284:8661–8669.

Guenther AB, Jiang X, Heald CL, Sakulyanontvittaya T, Duhl T, Emmons LK, Wang X. 2012. The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions. *Geosci Model Dev*. 5:1471–1492.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 59:307–321.

Hanson DT, Swanson S, Graham LE, Sharkey TD. 1999. Evolutionary significance of isoprene emission from mosses. *Am J Bot*. 86:634–639.

Harley PC, Monson RK, Lerdau MT. 1999. Ecological and evolutionary aspects of isoprene emission from plants. *Oecologia* 118:109–123.

Harms MJ, Thornton JW. 2013. Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nat Rev Genet*. 14:559–571.

Harvey CM, Sharkey TD. 2016. Exogenous isoprene modulates gene expression in unstressed *Arabidopsis thaliana* plants. *Plant Cell Environ*. 39:1251–1263.

Heald CL, Wilkinson MJ, Monson RK, Alo CA, Wang G, Guenther A. 2009. Response of isoprene emission to ambient CO_2 changes and implications for global budgets. *Glob Chang Biol*. 15:1127–1140.

Hewitt CN, Monson RK, Fall R. 1990. Isoprene emissions from the grass *Arundo donax* L. are not linked to photorespiration. *Plant Sci*. 66:139–144.

Ilmén M, Oja M, Huuskonen A, Lee S, Ruohonen L, Jung S. 2015. Identification of novel isoprene synthases through genome mining and expression in *Escherichia coli*. *Metab Eng*. 31:153–162.

Kaltenbach M, Jackson CJ, Campbell EC, Hollfelder F, Tokuriki N. 2015. Reverse evolution leads to genotypic incompatibility despite functional and active site convergence. *Elife* 4:1–20.

Kampranis SC, Ioannidis D, Purvis A, Mahrez W, Ninga E, Katerelos NA, Anssour S, Dunwell JM, Degenhardt J, Makris AM, et al. 2007. Rational conversion of substrate and product specificity in a *Salvia* monoterpene synthase: structural insights into the evolution of terpene synthase function. *Plant Cell* 19:1994–2005.

Karimi M, Inzé D, Depicker A. 2002. GATEWAY^{TM} vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci*. 7:193–195.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30:772–780.

Kesselmeier J, Staudt M. 1999. Biogenic volatile organic compunds (VOC): an overview on emission, physiology and ecology. *J Atmos Chem*. 33:23–88.

Khersonsky O, Roodveldt C, Tawfik DS. 2006. Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr Opin Chem Biol*. 10:498–508.

Khersonsky O, Tawfik DS. 2010. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem*. 79:471–505.

Köksal M, Jin Y, Coates RM, Croteau R, Christianson DW. 2011. Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature* 469:116–120.

Köksal M, Zimmer I, Schnitzler JP, Christianson DW. 2010. Structure of isoprene synthase illuminates the chemical mechanism of teragram atmospheric carbon emission. *J Mol Biol*. 402:363–373.

Külheim C, Padovan A, Hefer C, Krause ST, Köllner TG, Myburg AA, Degenhardt J, Foley WJ. 2015. The *Eucalyptus* terpene synthase gene family. *BMC Genomics* 16:450.

Lerdau M, Gray D. 2003. Ecology and evolution of light-independant phytogenic volatile organic carbon. *New Phytol*. 157:199–211.

Letunic I, Bork P. 2011. Interactive Tree of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res*. 39:475–478.

Lim CW, Han SW, Hwang IS, Kim DS, Hwang BK, Lee SC. 2014. The pepper lipoxygenase CaLOX1 plays a role in osmotic, drought and high salinity stress response. *Plant Cell Physiol*. 56:930–942.

Loivamäki M, Gilmer F, Fischbach RJ, Sörgel C, Bachl A, Walter A, Schnitzler J-P. 2007. *Arabidopsis*, a model to study biological functions of isoprene emission?. *Plant Physiol*. 144:1066–1078.

Loreto F, Fineschi S. 2015. Reconciling functions and evolution of isoprene emission in higher plants. *New Phytol*. 206:578–582.

Loreto F, Schnitzler JP. 2010. Abiotic stresses and induced BVOCs. *Trends Plant Sci*. 15:154–166.

Matarese F, Cuzzola A, Scalabrelli G, D'Onofrio C. 2014. Expression of terpene synthase genes associated with the formation of volatiles in different organs of *Vitis vinifera*. *Phytochemistry* 105:12–24.

Monson RK, Jones RT, Rosenstiel TN, Schnitzler JP. 2013. Why only some plants emit isoprene. *Plant Cell Environ*. 36:503–516.

Niinemets Ü, Sun Z, Talts E. 2015. Controls of the quantum yield and saturation light of isoprene emission in different-aged aspen leaves. *Plant Cell Environ*. 38:2707–2720.

Oku H, Inafuku M, Ishikawa T, Takamine T, Ishmael M, Fukuta M. 2015. Molecular cloning and biochemical characterization of isoprene synthases from the tropical trees *Ficus virgata*, *Ficus septica*, and *Casuarina equisetifolia*. *J Plant Res*. 128:849–861.

Pazouki L, Niinemets Ü. 2016. Multi-substrate terpene synthases: their occurrence and physiological significance. *Front Plant Sci*. 7:1019.

Rasulov B, Hüve K, Bichele I, Laisk A, Niinemets Ü. 2010. Temperature response of isoprene emission in vivo reflects a combined effect of substrate limitations and isoprene synthase activity: a kinetic analysis. *Plant Physiol*. 154:1558–1570.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. Mrbayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 61:539–542.

Sablok G, Fu Y, Bobbio V, Laura M, Rotino GL, Bagnaresi P, Allavena A, Velikova V, Viola R, Loreto F, et al. 2014. Fuelling genetic and metabolic exploration of C3 bioenergy crops through the first reference transcriptome of *Arundo donax* L. *Plant Biotechnol J*. 12:554–567.

Sanadze GA. 1957. Nature of gaseous substances from the *Robinia pseudoacacia* leaves. *Rep Akad Nauk GruzSSR* 19:83–86.

Sasaki K, Ohara K, Yazaki K. 2005. Gene expression and characterization of isoprene synthase from *Populus alba*. *FEBS Lett*. 579:2514–2518.

Sasaki K, Saito T, Lämsä M, Oksman-Caldentey KM, Suzuki M, Ohyama K, Muranaka T, Ohara K, Yazaki K. 2007. Plants utilize isoprene emission as a thermotolerance mechanism. *Plant Cell Physiol*. 48:1254–1262.

Schnitzler JP, Louis S, Behnke K, Loivamäki M. 2010. Poplar volatiles – biosynthesis, regulation and (eco)physiology of isoprene and stress-induced isoprenoids. *Plant Biol*. 12:302–316.

Schrepfer P, Buettner A, Goerner C, Hertel M, van Rijn J, Wallrapp F, Eisenreich W, Sieber V, Kourist R, Bruck T. 2016. Identification of amino acid networks governing catalysis in the closed complex of class I terpene synthases. *Proc Natl Acad Sci U S A*. 113:E958–E967.

Schwender J, Zeidler J, Gröner R, Müller C, Focke M, Braun S, Lichtenthaler FW, Lichtenthaler HK. 1997. Incorporation of 1-deoxy-D-xylulose into isoprene and phytol by higher plants and algae. *FEBS Lett*. 414:129–134.

Sharkey TD, Gray DW, Pell HK, Breneman SR, Topper L. 2013. Isoprene synthase genes form a monophyletic clade of acyclic terpene synthases in the Tps-b terpene synthase family. *Evolution* 67:1026–1040.

Sharkey TD, Monson RK. 2017. Isoprene research – 60 years later, the biology is still enigmatic. *Plant Cell Environ*. Forthcoming. Available from: http://doi.wiley.com/10.1111/pce.12930.

Sharkey TD, Wiberley AE, Donohue AR. 2008. Isoprene emission from plants: why and how. *Ann Bot*. 101:5–18.

Sharkey TD, Yeh S, Wiberley AE, Falbel TG, Gong D, Fernandez DE. 2005. Evolution of the isoprene biosynthetic pathway in kudzu. *Plant Physiol*. 137:700–712.

Sharkey TD, Yeh S. 2001. Isoprene emission from plants. *Annu Rev Plant Physiol Plant Mol Biol*. 52:407–436.

Sharkey TD. 2013. Is it useful to ask why plants emit isoprene?. *Plant Cell Environ*. 36:517–520.

Shaw JJ, Berbasova T, Sasaki T, Jefferson-George K, Spakowicz DJ, Dunican BF, Portero CE, Narváez-Trujillo A, Strobel SA. 2015. Identification of a fungal 1,8-cineole synthase from *Hypoxylon* sp. with specificity determinants in common with the plant synthases. *J Biol Chem*. 290:8511–8526.

Sindelarova K, Granier C, Bouarar I, Guenther A, Tilmes S, Stavrakou T, Müller JF, Kuhn U, Stefani P, Knorr W. 2014. Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. *Atmos Chem Phys*. 14:9317–9341.

St. Clair JM, Rivera-Rios JC, Crounse JD, Knap HC, Bates KH, Teng AP, Jorgensen S, Kjaergaard HG, Keutsch FN, Wennberg PO. 2016. Kinetics and products of the reaction of the first-generation isoprene hydroxy hydroperoxide (ISOPOOH) with OH. *J Phys Chem A* 120:1441–1451.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. 56:564–577.

Tattini M, Velikova V, Vickers C, Brunetti C, Di Ferdinando M, Trivellini A, Fineschi S, Agati G, Ferrini F, Loreto F. 2014. Isoprene production in transgenic tobacco alters isoprenoid, non-structural carbohydrate and phenylpropanoid metabolism, and protects photosynthesis from drought stress. *Plant Cell Environ*. 37:1950–1964.

Trapp SC, Croteau RB. 2001. Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics* 158:811–832.

Velikova V, Varkonyi Z, Szabo M, Maslenkova L, Nogues I, Kovacs L, Peeva V, Busheva M, Garab G, Sharkey TD, et al. 2011. Increased thermostability of thylakoid membranes in isoprene-emitting leaves probed with three biophysical techniques. *Plant Physiol*. 157:905–916.

Vickers CE, Possell M, Cojocariu CI, Velikova VB, Laothawornkitkul J, Ryan A, Mullineaux PM, Nicholas Hewitt C. 2009. Isoprene synthesis protects transgenic tobacco plants from oxidative stress. *Plant Cell Environ*. 32:520–531.

Vickers CE, Possell M, Laothawornkitkul J, Ryan AC, Hewitt CN, Mullineaux PM. 2011. Isoprene synthesis in plants: lessons from a transgenic tobacco model. *Plant Cell Environ*. 34:1043–1053.

Wang C, Chen Q, Fan D, Li J, Wang G, Zhang P. 2016. Structural analyses of short-chain prenyltransferases identify an evolutionarily conserved GFDPS clade in Brassicaceae plants. *Mol Plant* 9:195–204.

Weinreich DM, Delaney NF, DePristo MA, Hart DL. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312:111–114.

Whittington DA, Wise ML, Urbansky M, Coates RM, Croteau RB, Christianson DW. 2002. Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc Natl Acad Sci U S A*. 99:15375–15380.

Wiberley AE, Donohue AR, Westphal MM, Sharkey TD. 2009. Regulation of isoprene emission from poplar leaves throughout a day. *Plant Cell Environ*. 32:939–947.

Worton DR, Surratt JD, Lafranchi BW, Chan AWH, Zhao Y, Weber RJ, Park JH, Gilman JB, De Gouw J, Park C, et al. 2013. Observational insights into aerosol formation from isoprene. *Environ Sci Technol*. 47:11403–11413.

Yoshikuni Y, Keasling JD. 2007. Pathway engineering by designed divergent evolution. *Curr Opin Chem Biol*. 11:233–239.