Evolution of the Isoprene Biosynthetic Pathway in Kudzu¹[w]

Thomas D. Sharkey*, Sansun Yeh, Amy E. Wiberley, Tanya G. Falbel, Deming Gong, and Donna E. Fernandez

Department of Botany, University of Wisconsin, Madison, Wisconsin 53706 (T.D.S., S.Y., A.E.W., T.G.F., D.E.F.); and Protemix Corporation, University of Auckland, Auckland City, New Zealand (D.G.)

Isoprene synthase converts dimethylallyl diphosphate, derived from the methylerthrythritol 4-phosphate (MEP) pathway, to isoprene. Isoprene is made by some plants in substantial amounts, which affects atmospheric chemistry, while other plants make no isoprene. As part of our long-term study of isoprene synthesis, the genetics of the isoprene biosynthetic pathway of the isoprene emitter, kudzu (Pueraria montana), was compared with similar genes in Arabidopsis (Arabidopsis thaliana), which does not make isoprene. The MEP pathway genes in kudzu were similar to the corresponding Arabidopsis genes. Isoprene synthase genes of kudzu and aspen (Populus tremuloides) were cloned to compare their divergence with the divergence seen in MEP pathway genes. Phylogenetic analysis of the terpene synthase gene family indicated that isoprene synthases are either within the monoterpane synthase clade or sister to it. In Arabidopsis, the gene most similar to isoprene synthase is a myrcene/ocimene (acyclic monoterpenes) synthase. Two phenylalanine residues found exclusively in isoprene synthases make the active site smaller than other terpene synthase enzymes, possibly conferring specificity for the five-carbon substrate rather than precursors of the larger isoprenoids. Expression of the kudzu isoprene synthase gene in Arabidopsis caused Arabidopsis to emit isoprene, indicating that whether or not a plant emits isoprene depends on whether or not it has a terpene synthase capable of using dimethylallyl diphosphate.

Some plants (about one-third of angiosperms) can emit a significant fraction of recently fixed carbon as isoprene. Isoprene emission is an important biological process because it plays a large role in atmospheric chemistry (Trainer et al., 1987; Fehsenfeld et al., 1992; Thompson, 1992). It has been hypothesized that plants make isoprene for thermotolerance (Sharkey and Singsaas, 1995), especially, the activity of one of the last enzymes, isopentenyl diphosphate (IPP) isomerase (Idi), which converts IPP to DMAPP, shows a strong correlation with isoprene emission (Brüggemann and Schnitzler, 2002).

In metabolic pathways that diverge to produce different products, as is true of terpenoid synthesis in plastids, genes for enzymes that act early in the pathway are often more conserved than genes for enzymes acting later in the pathway (Rauscher et al., 1999). There are several examples of biosynthetic pathways where the terminal steps involve enzymes that are members of large gene families (Degtyarenko and Archakov, 1993; Mehta et al., 1993; Toh et al., 1993). In chloroplasts, many different terpenes are made from the products of the MEP pathway. Isoprene synthase has been cloned from mRNA of hybrid poplar (Populus × canescens). Escherichia coli transformed with this gene produces isoprene, along with insignificant amounts of monoterpenes (Miller et al., 2001). The isoprene synthase protein sequence is similar to that of the large family of terpene synthases (Bohlmann et al., 1999).

¹ This work was supported by the National Science Foundation (grant no. IBN–0212204).
* Corresponding author; e-mail tsharkey@wisc.edu; fax 608–262–7509.
[w] The online version of this article contains Web-only data.

¹ This work was supported by the National Science Foundation (grant no. IBN–0212204).
* Corresponding author; e-mail tsharkey@wisc.edu; fax 608–262–7509.
Isoprene Synthase

1998), raising the possibility that isoprene synthases may evolve much more rapidly than enzymes of the MEP pathway.

Isoprene-emitting plant species have been found in most of the major groups of land plants including eudicots, monocots, gymnosperms, pteridophytes, and mosses, listed in Kesselmeier and Staudt (1999; see also http://www.es.lancs.ac.uk/cnhgroup/isoemissions.pdf). Isoprene-emitting species show a scattered distribution across the phylogeny of land plants, suggesting multiple gains and/or losses of isoprene emission. Harley et al. (1999) hypothesized that isoprene synthase may have evolved many times and that this explains the distribution of isoprene emission among plants. Hanson et al. (1999) reasoned that loss of function was more likely than gain of function and hypothesized that the trait of isoprene emission was lost many times, accounting for the erratic distribution of the trait. More recently, Lerdau and Gray (2003) have proposed that isoprene emission may have originated independently in gymnosperms and angiosperms, but only once in the angiosperms, with multiple losses accounting for the distribution of isoprene emission among angiosperms.

The isoprene biosynthetic pathway was investigated by cloning genes encoding enzymes in the MEP pathway and isoprene synthase from kudzu (Pueraria montana). The MEP pathway genes were compared with the homologs from Arabidopsis (Arabidopsis thaliana), a nonemitter, and isoprene synthase was compared with aspen (Populus tremuloides) isoprene synthase. We examined specifically: (1) whether genes encoding enzymes acting at early steps in the isoprene biosynthetic pathway are more conserved than genes encoding enzymes acting at later steps; (2) the relationship between isoprene synthases from different species based on sequence and gene structure; and (3) the relationship of isoprene synthase to other terpene synthases. We also investigated what features might be necessary for activity as an isoprene synthase.

RESULTS

Characterization of MEP Pathway Genes

Three cDNA clones that encode DXS were isolated from a cDNA library prepared from heat- and light-treated kudzu leaves because heat and light induce isoprene synthesis. The clones were identical except for differing lengths of the 5’ untranslated region. The longest DXS cDNA (GenBank accession no. AY315652) was 2,679 bp in length and encoded a protein of 717 amino acids with a predicted molecular mass of 77.6 kD. The coding region of the kudzu DXS was 75% identical at the nucleotide level to the coding region of Arabidopsis DXS (At4g15560). Using prediction programs (PSORT, ChloroP, and TargetP; Emanuelsson et al., 1999, 2000), it was predicted that the kudzu gene product contains a transit peptide of 45 amino acids and is targeted to the chloroplast. When the nucleotides encoding the transit peptide sequence were removed from the analysis, the nucleotide identity between the Arabidopsis and kudzu DXS sequences increased to 89% (Table I). Alignment of the predicted protein sequence with other DXS protein sequences revealed that the proteins in the gene family were highly conserved, both within the plant kingdom and between plants and bacteria. Phylogenetic analysis based on the alignment of the coding regions of these genes revealed a relationship between the phylogeny of DXS sequences and phylogenetic distance (Fig. 2).

Table 1. Identity of DNA or amino acids comparing genes and proteins from kudzu and Arabidopsis

<table>
<thead>
<tr>
<th></th>
<th>DXS</th>
<th>DXR</th>
<th>Idi</th>
</tr>
</thead>
<tbody>
<tr>
<td>cDNA</td>
<td>75</td>
<td>73</td>
<td>65/71</td>
</tr>
<tr>
<td>Amino acids</td>
<td>89</td>
<td>80</td>
<td>87/90</td>
</tr>
</tbody>
</table>
There appeared to be two major lineages of DXS genes, but the kudzu gene was most closely related to one of the *Medicago truncatula* genes (MtDXS1; Fig. 2). *Medicago* does not emit isoprene but is in the same family as kudzu.

A cDNA that coded for DXR (GenBank accession no. AY315651) was also isolated from the kudzu cDNA library. Sequence analysis of this clone indicated that the open reading frame (ORF) was relatively short compared to the DXR sequences from other plant species. Partial sequencing of the corresponding kudzu genomic DNA revealed a different translational stop site. This information was used to correct the cDNA sequence in the region in question. The corrected DXR cDNA was 1,788 bp in length. The protein is predicted to contain 465 amino acids and have a molecular mass of 50.6 kD. This is somewhat smaller than DXR proteins from other plant species. A transit peptide of 44 amino acids was predicted to target the protein to the chloroplast.

Removing the putative transit peptide, the kudzu DXR nucleotide coding region was 80% identical to the Arabidopsis homolog (At5g62790) coding region. The protein sequences that were available were easily aligned, although the kudzu protein contains some amino acid residues not found in other species. Some of these changes result in a substitution of a similar amino acid, but there are five positions where amino acids with different properties are introduced. These include positions 104 (nonpolar to Ser), 166 (nonpolar to Ser), 231 (nonpolar to Ser), 319 (polar to Ile), and 365 (nonpolar to Ser; numbering based on the kudzu sequence). At these positions, the change was found only in kudzu except for position 231, where Ser was also found in *E. coli*. There were fewer plant DXR sequences available, but of those sequences available, a phylogenetic analysis indicated that all of the Asterid sequences were more closely related to each other than to the sequences from Rosids, and again, the kudzu sequence was most closely related to a member of the Fabales (*Glycine max*; Fig. 3).

Five Idi clones were isolated from the kudzu cDNA library. The sequences were identical, and the longest cDNA (GenBank accession no. AY315650) was 1,252 bp in length and encoded a protein 301 amino acids in length with a predicted molecular mass of 34.2 kD. A transit peptide of 71 amino acids was predicted to target the protein to the chloroplast. Alignment of the plant and *E. coli* Idi proteins showed that some of the plant sequences did not have a transit peptide; therefore, these protein products are probably cytosolic. At the nucleotide level, the kudzu Idi coding region was 65% and 71% identical to the Arabidopsis genes *Idi1* (At5g16440) and *Idi2* (At3g02780). When the transit
peptide sequence was removed from the analysis, the predicted mature proteins from Arabidopsis were 87% (Idi1) and 90% (Idi2) identical to the predicted kudzu mature protein. Phylogenetic analysis indicated that kudzu Idi was most closely related to G. max Idi of those sequences available (Fig. 4).

Cloning and Characterization of Isoprene Synthase Genes

Genomic and cDNA clones encoding isoprene synthase were isolated from kudzu libraries. The kudzu genomic clone (GenBank accession no. AY316691) was 3,139 bp in length and included 278 bp upstream of the ATG start site and 178 bp downstream of the stop codon (Fig. 5). The ORF in the cDNA was 1,824 bp in length and encoded a protein of 608 amino acids (Table I) with a predicted molecular mass of 70 kD. The first 44 amino acids were predicted by the ChloroP program to function as a transit peptide, directing the protein to the chloroplast.

The protein without the putative transit peptide sequence was expressed in E. coli and shown to have isoprene synthase activity. The $k_{cat}$ of the enzyme from

Figure 3. Phylogenetic analysis of DXR genes. A rooted phylogram was generated using a ClustalX alignment of the nucleotide coding region of DXR genes using the neighbor-joining method. Bootstrap analysis was done with 1,000 replicates and some of the bootstrap values are indicated. The tree was rooted with monocot DXR sequences.

Figure 4. Phylogenetic analysis of Idi genes. A rooted phylogram was generated using a ClustalX alignment of the nucleotide coding region of Idi genes using the neighbor-joining method. Bootstrap analysis was done with 1,000 replicates and some of the bootstrap values are indicated. The tree was rooted with C. reinhardtii Idi. To improve clarity, not all genes used in the analysis are shown.
E. coli was 0.088 mol mol\(^{-1}\) s\(^{-1}\). The velocity as a function of substrate was sigmoidal (Fig. 6), indicating cooperativity. A Hill plot (not shown) indicated that the Hill coefficient was greater than 3 and a coefficient of 4.1 was needed to best fit a model to the data. By modeling, a \(K_m\) of 7.7 mM was estimated.

Genomic and cDNA clones were also isolated from aspen libraries. The aspen genomic clone (GenBank accession no. AY341431) was 5,324 bp in length and included 1,067 bp upstream of the ATG start site and 1,086 bp downstream of the stop codon (Fig. 5). The ORF in the cDNA was 1,785 bp in length and encoded a protein of 595 amino acids with a predicted molecular mass of 69 kD. The aspen cDNA was 99% identical at the nucleotide level to the isoprene synthase sequence from hybrid poplar, which was shown to encode a protein with isoprene synthase activity (Miller et al., 2001). According to the ChloroP program, the first 45 amino acids function as a transit peptide directing the protein to the chloroplast. The aspen isoprene synthase nucleotide coding sequence was 65% identical to the kudzu gene, while the protein sequences were 54% identical. When the putative transit peptides were removed from the analysis, 57% of the amino acids in the proteins were identical.

The gene structure was determined by comparing the genomic and cDNA sequences for both kudzu and aspen (Fig. 5). Both isoprene synthase genes had seven exons and six introns. The size of the kudzu exons was similar to the size of the aspen exons. The intron sizes were also similar except for the first intron, which was significantly larger in aspen (927 bp) than in kudzu (357 bp). The exon and intron sizes were similar to the Class III terpene synthase genes (Table II; Trapp and Croteau, 2001b). In addition, the phase of the introns matched those of the Arabidopsis Class III terpene synthase genes (Aubourg et al., 2002).

Unlike aspen and kudzu, Arabidopsis does not emit isoprene and does not appear to contain isoprene synthase genes. The most closely related gene in the Arabidopsis genome is At2g24210, known to be a myrcene/ocimene synthase (Bohlmann et al., 2000). This gene was not closely related to isoprene synthases in a phylogenetic analysis. Kudzu isoprene synthase protein was expressed in E. coli and used to make antibodies specific for isoprene synthase (Fig. 7A). No immunoreactive protein was found in extracts of wild-type Arabidopsis (Fig. 7B).

Phylogenetic Analysis of Isoprene Synthase and Terpene Synthase Genes

To determine the relationship of the isoprene synthase genes to the terpene synthase gene family, a preliminary phylogenetic analysis was performed using the predicted protein sequences of the isoprene synthase genes and 177 terpene synthase genes (available as an MSF file, Supplemental Data File 1). The isoprene synthases were closely related, falling into a clade with sequences from Melaleuca alternifolia (GenBank accession no. AY279379), Malus × domestica (AY182241), and Cinnamomum tenuiplum (AJ457070). This clade appears to be situated as sister to a much larger clade that includes many well-characterized angiosperm monoterpene synthases (Tps-b; Bohlmann et al., 1998). Although M. alternifolia has not been tested for isoprene emission, isoprene emission has been shown in another Melaleuca species. The Malus × domestica sequence used in this analysis encodes an α-farnesene synthase (an acyclic sesquiterpene; Pechous and Whittaker, 2004). This unusual protein contains the RR(X)\(_W\) domain that is conserved in monoterpene synthases but is not usually found in sesquiterpene synthases. The finding of farnesene synthases with the RR(X)\(_W\) motif nested within monoterpene synthases has also been reported among gymnosperm terpene synthases (Martin et al., 2004). The C. tenuiplum sequence encodes a geraniol synthase according to the information included with the GenBank submission; geraniol is an acyclic oxygenated monoterpene.

Further analysis was performed using the nucleotide sequences of the coding regions of the Tps-b genes and selected Tps-a and Tps-g genes using maximum likelihood analysis. The isoprene synthase genes formed a monophyletic group closely associated with the monoterpene synthases (Fig. 8). The Tps-b genes tended to group together based on plant species as opposed to gene product function. This
supports the hypothesis that much gene duplication has occurred relatively recently (van der Hoeven et al., 2000). Unlike the analyses of DXS, DXR, and Idi, isoprene synthase from kudzu did not group together with terpene synthase sequences of other Rosids (Fig. 8).

What Are the Amino Acid Residues Likely To Be Involved in Isoprene Synthase Activity?

The cloning of isoprene synthase genes from two different species made it possible to identify conserved residues that might be important for isoprene synthase function. We focused on amino acid residues at four positions in the carboxy-terminal part of the enzyme based on their location in the active site and uniqueness among terpene synthase genes.

The 4 conserved residues are Phe-343, Gly-453, Phe-493, and Cys-496 (based on kudzu isoprene synthase numbering) and are highlighted in Supplemental Figure 2. Only the isoprene synthases and the Melaleuca gene product have a Gly at position 453. Fourteen of the other terpene synthases have either the first or second Phe but only the isoprene synthases and Melaleuca gene have both of these Phes. Cys-496 is unique to isoprene synthases; the Melaleuca gene product has a Thr at that position. To investigate the effect of the amino acid changes on the active site, the amino acid sequence for 5-epi-aristolochene synthase was modified by substituting isoprene synthase-specific residues at these positions. The predicted structure was then modeled based on the structure of Nicotiana epi-aristolochene synthase (5EAS; GenBank accession no. L04680) using SWISS-MODEL (Guex and Peitsch, 1997; http://www.expasy.org/spdbv/). The results were interpreted using the structure of EAS (Starks et al., 1997) and bornyl diphosphate synthase (Whittington et al., 2002).

The results of the protein modeling are shown in Figure 9. The two Phes are at opposite ends of the active site and cause it to be much smaller than the active site of 5EAS. The Gly unique to isoprene synthases is below Phe-440. (Numbering for the protein modeling is based on the 5EAS sequence. Position 440 in the 5EAS sequence corresponds to position 493 in the kudzu sequence.) EAS has a Tyr in the position (404) where isoprene synthases have Gly. This change may allow the Phe to be accommodated in the active site. The Cys found in isoprene synthases is one turn of the helix above Phe-440. Therefore, the active site Cys (Cys-440 in 5EAS) is still present in isoprene synthases but has been displaced by a distance corresponding to one helical turn (Supplemental Fig. 2, arrow). The changes in these positions (404 and 440 of the 5EAS sequence) would affect the position of the diphosphate end of DMAPP. Phe-343 (294 in the 5EAS sequence) would be located on the other side of the active site and near the dimethyl end of bound DMAPP. A Phe in this position could preclude binding of substrates larger than DMAPP.
Isoprene Synthase Is Sufficient to Confer the Trait of Isoprene Emission

To test whether expression of isoprene synthase was sufficient to allow a plant to emit isoprene, the kudzu genomic sequence (including 276 bp upstream from the ATG start codon) was transformed into Arabidopsis. Immunoreactive protein could be detected in two independent lines of transgenic plants and isoprene emission could be detected (Fig. 10). The occurrence of isoprene emission was confirmed by mass spectrometry and there was no evidence for monoterpene synthesis as a result of this transformation. The isoprene synthase in Arabidopsis was at the same apparent molecular mass (70 kD) as the native kudzu enzyme on western blots. This is consistent with the transcript being correctly spliced and the transit sequence being cleaved after import into chloroplasts. Immunoreactive protein could be recovered from purified chloroplasts of transgenic plants (not shown).

DISCUSSION

MEP Pathway Genes

Because isoprene synthesis can be limited by the rate at which the MEP pathway supplies substrate for isoprene synthase (Rosenstiel et al., 2003; Wolfertz et al., 2003), the MEP pathway must have a substantially greater capacity in plants that emit isoprene at high rates than in other plants. This might have led to rapid evolution of MEP pathway genes in kudzu, but this was not apparent in the sequences. Some plants
have two versions of DXS and there appears to be some specialization of function (Walter et al., 2002). We recovered only one DXS gene from kudzu and that appears related to MtDXS1 of Medicago, the version that is expressed in tissues other than roots. DXS2 may be related to monoterpene synthesis in some gland cells (Walter et al., 2002), but we did not find a DXS2 in kudzu.

It remains a possibility that small changes in amino acid sequence could give rise to large changes in kinetic properties, but it is also possible that all of the required extra flux through the MEP pathway is accommodated by regulation (Wolfertz et al., 2004). The Arabidopsis genome has only one functionally characterized copy of DXS (there are two additional candidate genes) and one DXR gene in contrast to the presumed monoterpene synthases, where six genes are present. The DXS and DXR genes of kudzu are similar to the Arabidopsis genes. Thus, the MEP pathway genes required for isoprene synthesis appear highly conserved. In contrast, the two known isoprene synthase genes, while more similar to each other than to other terpene synthase genes, have a much lower degree of amino acid identity than found among DXS, DXR, and IDI. In phylogenetic analyses, the MEP pathway genes of kudzu always were most closely related to other Fabales species (Figs. 2, 3, and 4).

**Isoprene Synthase**

Kudzu isoprene synthase exhibited cooperativity with respect to its substrate DMAPP. The $K_m$ of 7.7 mM is high but within the range of values reported previously for isoprene synthase activity (Silver and Fall, 1995; Wildermuth and Fall, 1996, 1998; Lehning et al., 1999; Wolfertz et al., 2004). There was no evidence for significant activity of this enzyme as a monoterpene synthase and its $k_{cat}$ was relatively high. We conclude that this is an isoprene synthase, giving the second isoprene synthase sequence known. Together with the intron structure of both the kudzu gene and the aspen gene reported here, we were able to make some predictions about the evolution of the capacity for isoprene synthesis. However, this analysis is limited by the fact that it is based on sequence from only two angiosperm species and that comparisons to genes in databases rely on annotations in the databases, most of which have not been confirmed.

Isoprene synthases plus three other terpene synthases formed a monophyletic group most closely related to Tps-b terpene synthases but possibly representing a new group. IDI sequences of kudzu, Populus, and Melaleuca are unrelated to one another (Fig. 4), but the terpene synthases from these three species form a monophyletic clade. With so few sequences available, it is not possible to distinguish whether this relationship is based on common ancestry of the gene or convergent evolution based on gene function.

Two major hypotheses concerning the evolution of the trait of isoprene emission from land plants are: (1) evolution once, early in land plant evolution, with multiple losses of function (Hanson et al., 1999); and (2) frequent evolution of genes with the relevant activities (Harley et al., 1999) analogous to $C_4$ and Crassulacean acid metabolism evolution (Sage and Pearcy, 2000; Keeley and Rundel, 2003). If loss of function is more likely than gain of function, then hypothesis 1 is favored. The clustering of the two isoprene synthase genes is most consistent with a single origin of isoprene synthase within the angiosperms. However, in gymnosperms, the genomic structure of monoterpene synthases is very different from the seven exon/six intron gene family found in...
angiosperms. No gymnosperm terpene synthase genes are known and we speculate that isoprene synthases from gymnosperms will be most similar to monoterpene synthases of gymnosperms. If true, this would support the hypothesis that isoprene synthase arose more than once. The finding that the isoprene synthases of kudzu and aspen are less than 60% identical and are nearly as closely related to Tps-b genes coding for monoterpene synthases allows a hybrid hypothesis. The terpene synthase gene family may have evolved once, early in land-plant evolution. The terpene synthase gene family, through rapid evolution, then provided a source of genes that could become isoprene synthases. This favors the hypothesis that isoprene emission evolved independently many times, though it may have evolved only once in the angiosperms. The data presented here and this interpretation are consistent with the view put forward by Lerdau and Gray (2003) that isoprene synthesis is a primitive condition for angiosperms but a derived capability in gymnosperms.

Based on the similarity of specific amino acid residues between the isoprene synthases and the terpene synthase gene of M. alternifolia (GenBank accession no. AY279379), we hypothesize that AY279379 is also an isoprene synthase. The sequence from Malus × domestica was determined to be an α-farnesene synthase (Pechous and Whitaker, 2004) and AJ457070 is reported to be a geraniol synthase. Farnesene, myrcene, ocimene, and geraniol are acyclic terpenes, and so the chemistry of their formation would be similar to that of isoprene (except for the introduction of oxygen into geraniol). Martin et al. (2004) also reported the clustering of some of these same genes and also a clustering of farnesene synthases from Norway spruce and loblolly pine (Phillips et al., 2003) together with a linalool synthase, nested well within gymnosperm monoterpene synthase genes. Martin et al. (2004) interpreted this to indicate independent evolution of α-farnesene synthases from monoterpene synthases in both angiosperms and gymnosperms. This is similar to what we are speculating for the evolution of isoprene synthases in angiosperms and gymnosperms.

The terpene synthase gene family may have undergone extensive evolution throughout the radiation of the seed plants. The genes involved in GA biosynthesis are more conserved and tend to be single copies (Aubourg et al., 2002), presumably because changes in these genes would have deleterious effects (Yamaguchi et al., 1998). However, the vast majority of the terpene synthases produce secondary compounds that are not vital for plant function but may be important in ecological interactions with other organisms (Harborne, 1991). The diversity of the family may be the result of repeated duplication of an ancestral gene and divergence by functional and structural specializations (Trapp and Croteau, 2001b). Radiation of some of the terpene synthase genes occurred recently within a species, as in the case of tomato where the terpene synthase genes appeared to have radiated after the Lycopersicon genus diverged from the rest of Solanum (van der Hoeven et al., 2000). In Arabidopsis, there are eight pseudogenes and multiple tandem repeats indicative of a high rate of evolution of terpene synthases (Aubourg et al., 2002). The phylogram in Figure 7 is consistent with all of the monoterpene synthases of Arabidopsis and Perilla frutescens having diverged after speciation (although the density of species sampling is still limited). Therefore, the terpene synthase gene family could provide a reservoir from which isoprene synthase genes could arise more than once.

Comparison of isoprene synthase gene sequences with monoterpene synthase gene sequences of closely related species can help determine how recently isoprene synthases evolved. The one case where this is possible is the isoprene synthase of hybrid poplar and gene O32C12 of poplar. Gene O32C12 is more closely related to monoterpene synthases of other species than to hybrid poplar isoprene synthase, indicating that the two poplar genes likely have independent origins.

Another piece of evidence for frequent evolution of isoprene synthases from a reservoir of monoterpene synthase genes is the phylogenetic relationship of the myrcene synthase gene from oak (Quercus spp.). Mediterranean oak species emit monoterpenes under the same conditions that stimulate isoprene emission in North American oaks (Loreto et al., 1998a, 1998b). The myrcene synthase isolated from this oak species thought to be responsible for the monoterpene emission (Fischbach et al., 2001) was not part of the isoprene synthase clade and does not make isoprene when given DMAPP. Because the function of the monoterpene synthases in the Mediterranean oaks may be the same as isoprene in North American oaks, the isoprene synthase and these monoterpene synthase genes may have evolved independently from different monoterpene synthases to have the same function.

In conclusion, the molecular evidence indicates that MEP pathway genes are more highly conserved than isoprene synthase genes. Isoprene synthases from angiosperms form a monophyletic group but it is not yet possible to determine if this reflects evolution from a common ancestral gene or convergence based on function. The data supports a hybrid hypothesis for the evolution of the capacity for isoprene emission (Lerdau and Gray, 2003) that isoprene synthase could be ancestral and has evolved once (Hanson et al., 1999) in angiosperms but that isoprene synthesis in gymnosperms represents convergent evolution. Several amino acid residues that could be critical to specifying isoprene synthesis rather than synthesis of monoterpenes have been identified and can be used to predict additional isoprene synthase sequences. Finally, adding an isoprene synthase gene to a species that does not emit isoprene causes plants to emit isoprene. The trait of isoprene emission appears to be easily acquired by plants, although the regulation of DMAPP concentration may have to be altered to support high rates of isoprene emission seen in some plants.
MATERIALS AND METHODS

Growth and Inductive Conditions of Kudzu and Aspen

Kudzu (*Pueraria montana* var. *lobata* [Willd.] Maesen and S. Almeida) plants were grown from stem cuttings in 10-L pots filled with vermiculite/peat moss-based growing medium (Metro-Mix 360, Grace Sierra, Milpitas, CA). The plants were grown in a reach-in controlled-environment chamber (model E15, Conviron, Winnipeg, Canada) at 12-h nights at 18°C and 12-h days at 20°C with a light level of 200 μmol photons m⁻² s⁻¹. Plants were watered with Miracle-Gro Water Soluble All Purpose Plant Food at a concentration of 0.675 g L⁻¹, pH 6, as recommended by the manufacturer’s instructions (Scotts, Marysville, OH). Isoprene synthesis was induced in individual leaves (attached to the plant) using a laboratory-based gas-exchange system (Tennessen et al., 1994) to increase the temperature and light to 35°C and 1,000 μmol photons m⁻² s⁻¹ for at least 6 h. The kudzu leaves were monitored for isoprene emission with an analytical gas chromatograph with photoionization detection (Shimadzu, Kyoto) calibrated with an isoprene standard (Sigma-Aldrich, St. Louis) made daily as described by Loreto and Sharkey (1993). Ten-milliliter gas samples were taken before, during, and after induction to ensure that isoprene synthesis was induced.

Aspen (*Populus tremuloides* Michx.) trees were grown in temperature-controlled greenhouses in the Biotron facility at the University of Wisconsin, Madison with a daytime temperature of 25°C and nighttime temperature of 16°C. The daylength was extended to 16 h using high pressure sodium vapor lamps. The plants were watered with half-strength Hoagland solution (Hoagland and Arnon, 1938). Isoprene emission was induced in aspen trees by moving the trees into a second greenhouse room maintained at 35°C air temperature and 1,000 μmol photons m⁻² s⁻¹ for at least 6 h. For both kudzu and aspen, leaf samples were collected by quickly cutting at the petiole and immediately freezing in liquid nitrogen. Tissue samples were stored at −80°C until used.

Analysis of Transgenic Arabidopsis

Arabidopsis (*Arabidopsis thaliana*) ecotype Wassilewskija was used in these experiments. Seeds were surface sterilized (95% ethanol for 5 min and 0.5% [v/v] sodium hypochlorite for 5 min) and plated on germination medium (Murashige and Skoog salts and vitamins [Murashige and Skoog, 1962] supplemented with 10 g/L Suc, 0.5 g/L MES, pH 5.6 to 5.7, and 8 g/L agar). The plates were kept at 4°C for 2 d and then transferred to 22°C, 16-h-light (approximately 125 μmol photons m⁻² s⁻¹) and 8-h-dark (defined as long-day) conditions, and 70% to 80% relative humidity in a controlled environment chamber (Enconmic Ecological Chambers, Winnipeg, Canada). After 1 week, seedlings were transferred into a 2.1 (v/v) mix of peat-based potting mixture (Jiffy Mix, Jiffy Products of America, Batavia, IL) and Perlite (Midwest Perlite, Appleton, WI) and were grown to maturity under short-day conditions on a 10-h-light/14-h-dark cycle.

Kudzu cDNA Library Construction and Screening

Total RNA was extracted from tissues as described by Logemann et al. (1987) with minor modifications. The poly(A) RNA was isolated from total RNA using the PolyATtract mRNA Isolation System (Promega, Madison, WI). Approximately 1 mg of poly(A) enriched RNA isolated from kudzu leaves induced to make isoprene was used to generate a cDNA library using the Uni-Zap XR Library Construction kit according to the manufacturer’s instructions (Stratagene, La Jolla, CA). Approximately 113,400 recombinant phage were present in the unamplified cDNA library.

After amplification, the cDNA library was plated and screened on nitrocellulose membranes (Osmonics, Minnetonka, MN) as described by Sambrook et al. (1989) using probes generated from Arabidopsis sequences. cDNA clones were obtained from the Arabidopsis Biological Resource Center (ABRC) stock center for Arabidopsis DXS (H2A12T7), DXR (120E17T), and Idi (168F15T) genes. 32P-oligotet labeled probes were generated from the Arabidopsis sequences by random priming and hybridized with the nitrocellulose membranes using standard procedures (Sambrook et al., 1989). The membranes were exposed to Kodak X-Omat AR x-ray film (Eastman-Kodak, Rochester, NY) for 12 to 48 h to detect hybridization. Candidate clones were excised and subcloned into the pBluescript SK (−) plasmid (Stratagene) according to the manufacturer’s instructions. Each clone was sequenced in both directions by using cycle sequencing with Big Dye terminators (Perkin Elmer Applied Biosystems, Wellelsley, MA). The sequences were analyzed at the University of Wisconsin-Madison Biotechnology sequencing facility. GenBank was searched using BLAST programs (Altschul et al., 1990, 1997) to determine the identity of the clones.

The kudzu isoprene synthase gene was isolated by screening the cDNA library with a probe consisting of 300 bp of aspen isoprene synthase sequence. Aspen sequence information was provided by Dr. Ray Fall (University of Colorado, Boulder). This information was used to design gene-specific primers and a 300-bp cDNA fragment was obtained by reverse transcription-PCR from RNA isolated from induced leaves of aspen using the Omniscript kit according to the manufacturer’s instructions (Qiagen, Valencia, CA). The cDNA sequence was used to generate a 32P-labeled probe, which was then used to screen the kudzu library and isolate a cDNA clone representing part of the kudzu gene. Primers were designed based on the sequence of this clone and were used along with Takala Ex Taq DNA Polymerase and the manufacturer’s instructions (PanVera, Madison, WI) to amplify the rest of the cDNA sequence from the kudzu library. A boiled aliquot (1 μL) of the cDNA library was used as the template in the 50-μL PCR reaction. The PCR reaction cycle was: 94°C for 5 min; 30 cycles of 94°C for 30 s, 55°C for 1 min and 72°C for 1 min; 72°C for 7 min. The PCR-generated fragments were cloned into the pCR4-TOPO vector using the TOPO TA cloning kit for sequencing and transformed into TOP10 competent Escherichia coli cells (Invitrogen, Carlsbad, CA). Both strands of the resulting plasmids were sequenced.

Cloning of the cDNA and Genomic Sequences for the Aspen Isoprene Synthase

An aspen genomic library was constructed as described for the kudzu genomic library. Sequence corresponding to the aspen isoprene synthase gene was obtained by PCR amplification using the same strategy outlined for the kudzu gene. Primers based on the 300 bp sequence obtained from R. Fall were used to obtain the 5′ and 3′ flanking sequences. New primers were designed based on these sequences and were used to obtain an overlapping fragment representing the middle of the gene. The PCR-generated fragments were cloned into the pCR4-TOPO vector using the TOPO TA cloning kit for sequencing and transformed into TOP10 competent Escherichia coli cells (Invitrogen, Carlsbad, CA). Both strands of the resulting plasmids were sequenced.

Expression and Purification of Kudzu Isoprene Synthase in E. coli

Oligonucleotides specific for kudzu isoprene synthase and flanked by NdeI (sense primer, 5′-ACACACATATGCTTGGAAATACCTTA-3′) and BamHI (antisense primer, 5′-GGTACGATCCCACGTCACATTGTTAGT-3′) restriction sites were used to amplify isoprene synthase cDNA sequence from the kudzu cDNA library. The amplified sequence encoded a truncated form of the protein because the 5′ primer was designed close to the conserved double Arg region (positions R61, R62). The PCR product was cloned into a modified pBluescript SK (−) vector with T overhang. The
fragment was released from the pBlueScript SK (+) vector using the NdeI and BamHI restriction enzymes (Promega) and cloned into the pET15b expression vector (Novagen, Madison, WI), which adds a His-tag. The new construct, named pET15b-ISP5, was transformed into expression host E. coli cells, BL21(DE3) (Novagen).

Selected, verified, and transformed cells were grown at 37°C by inoculating 450 mL of Luria-Bertani (supplemented with 100 μg/mL ampicillin) with a 30 mL overnight starter culture of the transformed E. coli. Protein expression was induced by adding 0.4 mM isopropyl-1-thio-β-D-galactopyranoside when the growing culture had an OD600 of 0.8. After 4 to 5 h, cells were collected by centrifuging at 4,500 g for 10 min and were frozen at −20°C overnight. The frozen cells were ruptured by thawing on ice for 15 min, resuspension in 4 mL lysis buffer (50 mM NaH2PO4, pH 8.0; 300 mM NaCl; 10 mM imidazole; 1 mg/mL lysozyme [Sigma-Aldrich]), followed by incubation on ice for 30 min and then sonication. The lysates were centrifuged at 10,000g for 30 min to pellet bacterial debris and leave proteins in the supernatant. The recombinant isoprene synthase was purified from the supernatant using a Ni-NTA agarose column according to the manufacturer’s instructions (QiaGen).

Isoprene synthase activity was measured in a pH 8.0 buffer (50 mM bicine, 50 mM MgCl2, 5 mM KCl, 2 mM NaF, MgCl2, and 5% glycerol and 5 mM dithiothreitol). Reactions were carried out in 5.5-mL sealed vials at 35°C. After 15 min, 3 mL of the head space gas was removed by syringe and injected on a gas chromatograph with photoionization detection for quantification. Reactions were typically run with 12 mM DMAPP and DMAPP was synthesized as described by Davison et al. (1985).

Generation of Kudzu Isoprene Synthase Antibodies and Western Blotting

Recombinant isoprene synthase protein was purified following separation of product eluted from the Ni-NTA column (described above) using 10% denaturing PAGE (Bio-Rad Laboratories, Hercules, CA). A single major protein band at the expected Mw was taken as the recombinant isoprene synthase protein and was excised from the gel. Approximately 3 mg of purified protein was sent to Caprologics (Hardwick, MA) to immunize 2 rabbits for antibody production.

Total soluble protein was extracted from plant tissues as described by Heck et al. (1996), separated on a NuPAGE NOVEX 4% to 12% BisTris gel, and transferred to polyvinylidene difluoride membrane (Amersham) using the NOVEX XCell II Mini-Cell system according to the manufacturer’s instructions (Invitrogen). The membranes were incubated with either immune or preimmune serum that was diluted 1:25,000. The immunoreactive protein was visualized using the ECL western-blotting system according to manufacturer’s instructions (Amersham). The secondary antibodies linked with horseradish peroxidase was diluted 1:5,000. The blots were linked with horseradish peroxidase was diluted 1:5,000. The blots were linked and visualized using the Chemiluminescent Detection System (Tropix). Two rabbits for antibody production.

Transformation of Arabidopsis with the Kudzu Genomic Isoprene Synthase Gene

The kudzu isoprene synthase genomic sequence was amplified using specific oligonucleotides (sense primer, 5′-GTCGCTCCCGCGGTGTTGACTGCTGGTATTAGCCATGAANTA-3′; antisense primer, 5′-ACACCCGCCGATTTAGATCGCTTTCGTTATTTTA-3′) that were flanked with Smal restriction sites (underlined). The amplified fragment included 270 bp upstream of the ATG translational start site and ended near the translational stop codon. The amplified DNA was cloned into the pGEM-T vector (Promega). The insert was released with the Smal restriction enzyme and cloned into the pPPZ212 Agrobacterium tumefaciens binary vector (Hajdukiewicz et al., 1994) at the Smal site to produce pGIPS. The pGIPS plasmid was transformed into the A. tumefaciens strain GV3101.

Arabidopsis ecotype Wassilewskija was transformed using the whole-plant floral dipping (Clough and Bent, 1998) method. Transformed individuals (T1 generation) were selected on germination medium supplemented with 75 μg/mL kanamycin before transfer to soil. The number of unlabeled T1 DNA loci within each line was determined by analyzing the segregation of kanamycin resistance in the progeny (T2 generation). The rate of isoprene emission from transformed Arabidopsis was determined by enclosing plants in a static chamber and sampling the headspace after 10 min. Additional experiments were done by enclosing detached leaves in small vials for 30 to 60 min followed by sampling of the head space gas for isoprene analysis. Isoprene emission was confirmed by mass spectrometry performed at the Wisconsin State Laboratory of Hygiene (Madison, WI) by collecting on 4-mm i.d. solid adsorbent tubes (Dynamax MX-06-2313 sorbent tubes packed with 20:35 mesh Tenax-TA/60:80 mesh Carboxen 1000/60:80 mesh Carbosieve SII).

Computer Analysis and Prediction Programs Used to Characterize the Sequences

The EDITSEQ and SEQMAN programs (DNASTAR, Madison, WI) were used to assemble the gene sequences. PSORT (Nakai and Horton, 1999; http://psort.imst.u-tokyo.ac.jp/), TargetP (Nielsen et al., 1997; Emanuelsson et al., 2000; http://www.cbs.dtu.dk/services/TargetP/) and ChloroP (Emanuelsson et al., 1999; http://www.cbs.dtu.dk/services/ChloroP/) were used to identify possible targeting sequences in the gene products. Analysis of the promoter sequences for any known motifs and elements was performed using the PLACE signal scan search (Higo et al., 1999; http://www.dna.affrc.go.jp/htdocs/PLACE/).

Compilation of the Terpene Synthase Sequences

Terpene synthase sequences from other species were obtained by several different methods. The GenBank database was searched by PubMed using “terpene synthase” as a keyword as well as by using BLASTP and TBLASTN programs (Altschul et al., 1997) with isoprene synthase sequences from kudzu and aspen. The Pfam database was searched (Sonhammer et al., 1998; http://pfam.wustl.edu/) with both isoprene synthase genes to obtain protein sequences that contained terpene synthase domain(s). The Arabidopsis terpene synthase sequences were obtained by searching (using BLASTP programs) the Arabidopsis genome (http://mpsg.gsl.de/proj/thal/db/search/search_frame.html) with both terpene synthase domains obtained from Pfam. Both nucleotide and predicted protein sequences were obtained. BLASTN and TBLASTN searches of the hybrid aspen (Populus tremula × Populus tremuloides) sequencing project (http://poppel.fysbot.umu.se/info.html) using the isoprene synthase genes were performed to obtain nucleotide sequences that were translated for the analysis.

Multiple Sequence Alignment and Phylogenetic Analysis

ClustalX (Higgins et al., 1996; downloaded from ftp://ftp.ibcgscn.u-strasbg.fr/pub/ClustalX/) was used to perform the alignment of the sequences. The nucleotide sequences were aligned using the International Union of Biochemistry matrix with gap opening penalty of 10 and gap extension penalty of 0.10. The protein sequences were aligned using the Gonnet 250 matrix and gap opening penalty of 35 and gap extension penalty of 0.75. The sequence identity comparison was performed using the sequence distance analysis of MEGALIGN (DNASTAR) on the aligned sequences.

Protein sequences were used instead of the nucleotide sequences for the preliminary terpene synthase tree analysis. The nucleotide sequences were translated into protein sequences and aligned. The alignment of all the known terpene synthase genes and the isoprene synthase genes was used to generate a phylogram using PAUP* (Swofford, 1998) with the neighbor-joining method (Saitou and Nei, 1987). Three related fungal genes (AB013295, Y15013, and AB003395) and a related yeast gene (U60996) were included in the alignment and were defined as outgroups.

The nucleotide sequences of the Tps-b and isoprene synthase genes were aligned along with the nucleotide sequences of selected Tps-a and Tps-g genes. The alignments were used to generate a second phylogram that was rooted using the Tps-a gene sequences. Maximum likelihood analysis, as implemented in PAUP, was used with the nucleotide sequences. ModelTest version 3.06 (Posada and Crandall, 1998) was used to evaluate 56 maximum likelihood models using hierarchical likelihood ratio tests. A heuristic maximum likelihood search with tree bisection-reconnection branch-swapping was conducted using parameters determined for the best model of sequence evolution by ModelTest. The sequences for the other DXS, DXR, and Idi genes were obtained by searching GenBank with the kudzu nucleotide and protein sequences. The gene sequences from Chlamydomonas reinhardtii were used to root the DXS.
tree. The DXR tree was rooted with the monocot gene sequences. The trees were bootstrapped with 1,000 replicates.

Sequence data for the genes reported here were obtained from the EMBL/GenBank libraries under the following accession numbers: poplar clone DXP, AY315652; poplar clone DXR, AY315651; poplar clone DXP, AY315650; poplar clone IspS, AY316691; and aspen IspS, AY341431.

ACKNOWLEDGMENTS

We thank Ray Fall for providing a 300-bp sequence of aspen isoprene synthase, which we used to clone the isoprene synthase sequences reported here, and David Baum for comments on the manuscript. The Wisconsin State Laboratory of Hygiene made the mass spectrum analysis to confirm the presence of isoprene in the head space of vials with transformed Arabidopsis leaves.

Received October 5, 2004; returned for revision November 5, 2004; accepted November 8, 2004.

LITERATURE CITED


Sharkey et al.


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


