A method for the recovery of full-length cDNAs from predicted terpene synthase genes containing introns is described. The approach utilizes Agrobacterium-mediated transient expression coupled with a reverse transcription-polydeoxyribonucleotide chain reaction assay to facilitate expression cloning of processed transcripts. Subsequent expression of intronless cDNAs in a suitable prokaryotic host provides for direct functional testing of the encoded gene product. The method was optimized by examining the expression of an intron-containing β-glucuronidase gene agroinfiltrated into petunia (Petunia hybrida) leaves, and its utility was demonstrated by defining the function of two previously uncharacterized terpene synthases. A tobacco (Nicotiana tabacum) terpene synthase-like gene containing six predicted introns was characterized as having 5-epi-aristolochene synthase activity, while an Arabidopsis (Arabidopsis thaliana) gene previously annotated as a terpene synthase was shown to possess a novel sesquiterpene synthase activity for α-barbatene, thujopsene, and β-chamigrene biosynthesis.

Sesquiterpene synthases are a class of enzymes that catalyze the conversion of prenyl diphosphates to mono-, sesqui-, and diterpenoid compounds (Chappell, 1995; Davis and Croteau, 2000). The reactions catalyzed by many of these enzymes are stereospecific and complex, often generating one to upwards of 30 reaction products (Steele et al., 1998), many of which may contain one or more ring structures. Terpene synthases are of particular interest to phytochemists for their role in the synthesis of natural products, including flavors and fragrances, such as nootkatone and patchouli alcohol and antimicrobial phytoalexins like capsidiol (Chappell, 1995). Sesquiterpene synthases, a subset of terpene synthases catalyzing the biosynthesis of products derived from farnesyl diphosphate, have been characterized from many plant species including tobacco (Nicotiana tabacum; Facchini and Chappell, 1992), tomato (Lycopersicon esculentum; Colby et al., 1998), Hyoscyamus (Back and Chappell, 1995), cotton (Gossypium hirsutum; Chen et al., 1995), Arabidopsis (Arabi -

1 This work was supported by the Kentucky Tobacco Research and Development Center and the National Science Foundation (to J.C.), by the National Institutes of Health (R.M.C.), and by the Kentucky Agricultural Experiment Station.

2 These authors contributed equally to the paper.

3 Present address: Biochemistry Department, University of Nebraska-Omaha, Omaha, NE, 68182.

4 Present address: Allylix Inc., A165 ASTeCC, University of Kentucky, Lexington, KY 40546–0286.

* Corresponding author; e-mail chappell@uky.edu; fax 859–257–7125.

[1w] The online version of this article contains Web-only data. Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.105.059386.
predict the product specificity of the individual terpene synthase enzymes. For instance, monoterpene synthases utilize 10-carbon geranyl diphosphate (GPP) as substrate, while sesquiterpene synthases use 15-carbon \((E, E)\)-farnesyl diphosphate (FPP). Moreover, terpene synthase reaction product specificity relies on a very complex series of partial reactions that can result in stereo- and regio-specific cyclization events giving rise to mono-, di-, tri-, and even tetracyclic structures, and can include other modifications such as methyl/methylene migrations and double bond formation or loss (Croteau, 1987; Cane, 1990). The importance of accurate functional identification for terpene synthases is also evident from the observation that many of these genes belong to gene families within a particular plant species (Facchini and Chappell, 1992; Trapp and Croteau, 2001; Martin et al., 2004). Many synthases are multi-functional (Steele et al., 1998), and single amino acid substitutions are sufficient to alter the reaction product specificity for a terpene synthase (Cane et al., 1997; Rising et al., 2000; Rynkiewicz et al., 2002; Seemann et al., 2003; Deligeorgopoulou and Allemann, 2004; Kollner et al., 2004).

As part of our ongoing investigation into the structure and catalytic function of terpene synthases, the objective of this work was to develop a strategy for terpene synthase cDNA recovery in which a priori knowledge of gene expression patterns was not necessary, and a method in which the number of molecular manipulations was kept to a minimum before functional expression of the resulting cDNA. Our approach was to combine Agrobacterium-mediated transient expression in infiltrated leaf mesophyll (agroinfiltration) with a reverse transcription (RT)-PCR assay that facilitated cloning into expression vectors (Scheme I). In this way, we should be able to dictate the conditions of gene expression in a plant cell, exploit native genetic machinery to produce fully processed mRNA, and recover cognate cDNA. We refer to this strategy as surrogate splicing for functional genomics and demonstrate its utility by examining the function of two previously uncharacterized sesquiterpene synthase-like genes, one from tobacco and the other from Arabidopsis.

**RESULTS**

**Optimization of the Surrogate Splicing Method**

We initially assessed the surrogate splicing method using a previously constructed \(\beta\)-glucuronidase (GUS) marker gene containing an intron, GUSi (Vancanneyt et al., 1990), inserted into a pBI Ti-plasmid vector (Bevan, 1984) and agroinfiltrated into several plant species including tobacco and petunia (Petunia \(\times\) hybrida) leaves. Van der Hoorn et al. (2000) had previously used this approach to assess applicability of the agroinfiltration method for gene expression studies per se. Our initial experiments monitored GUS enzyme activity (a measure of transcript processing) quantitatively in extracts of the infiltration zones over time in relationship to visual inspection of tissue integrity. Under the conditions employed, tobacco tissues displayed water-soaked lesions and tissue collapse within 4 d, while petunia tissues maintained normal appearances for 8 or more days. Tissue collapse was dependent on the Agrobacterium concentration used for infiltration and could be minimized by using low concentrations. Low Agrobacterium concentrations, however, translated into significantly lower levels of GUS enzyme activity associated with the infiltration zones and often lower levels than could be accounted for by the dilution factor of the infiltrated Agrobacterium cultures. Given the ultimate aim of generating the greatest levels of processed RNA transcripts to facilitate the next step of reverse transcription polydeoxyribonucleotide chain reaction (RT-PCR), emphasis...
was placed on maximizing expression level of GUS activity without compromising tissue integrity. Expression levels and tissue quality were routinely superior with petunia versus tobacco under all conditions tested.

The time course for petunia mesophyll cells taking up and expressing T-DNA-borne transgenes following agroinfiltration was determined in leaf discs collected at daily intervals and tested quantitatively for GUS enzyme activity (Fig. 1A). Detached petunia leaves were infiltrated with a suspension of *Agrobacterium tumefaciens* carrying the intron-containing GUS gene (GUSi) driven by the cassava vein mosaic virus (CsVMV-GUSi) promoter, a promoter previously characterized for its ability to direct strong constitutive expression in leaf tissue (Verdaguer et al., 1998). GUS activity was absent or barely above background levels for the first 2 d after infiltration, then increased dramatically and almost linearly over the next 4 d. Maximum GUS activity was observed by 6 d post-agroinfiltration and declined rapidly thereafter. The time course for transient expression of GUS activity in petunia is consistent with those previously reported for other plant species (Van der Hoorn et al., 2000). The absolute level of GUS activity in these transient expression studies was readily measured and directly comparable to the levels observed in stable transgenic lines (Verdaguer et al., 1998).

Recovery of a full-length cDNA from petunia leaf tissue agroinfiltrated with the GUSI gene was used to assess further the utility of this system for the generation of properly processed transcripts (Fig. 1B). Total RNA was isolated using a standard isolation procedure and 5 μg used for first-strand cDNA synthesis with an oligo(dT) primer. An aliquot of the first-strand synthesis reaction was then used in combination with primers designed to bracket the start and stop codons of the GUS gene and containing convenient restriction sites for future insertion of the PCR fragments into suitable prokaryotic expression vectors. Single-primer, RNA-only, and templateless controls showed no amplification products (lanes 2–5), while the complete experimental reaction yielded a reaction product that was approximately 190 bp smaller than the positive control product amplified directly from the GUSi gene (compare lane 6 to lane 7). The amplification product of lane 6 was subsequently cloned and sequenced. The sequence revealed that it was identical to the original GUS gene minus the 189-bp artificial intron (Vancanneyt et al., 1990) that had been properly removed at the 5′ (TAC/GTAA) and 3′ (GCAG/CT) splice sites.

**Functional Identification of a Putative Tobacco Sesquiterpene Synthase Gene**

General applicability of the surrogate splicing method for functional analysis of an unknown gene containing several predicted introns was initially assessed using a putative tobacco terpene synthase genomic clone referred to as g110 (Fig. 2). This genomic clone, along with approximately 30 other clones, was obtained when a tobacco cv Xanthi genomic library was screened with a probe corresponding to the first two exons of the 5-epi-aristolochene synthase 4 gene (EAS4; Facchini and Chappell, 1992). Sequence analysis of g110 (GenBank accession AY313939) revealed that it was 90% identical to the EAS4 gene at the nucleotide level (after insertion of 14 gaps to optimize for sequence alignment) and, like EAS4, was predicted to have seven exons punctuated by six introns (Fig. 2A). A single nucleotide deletion in the first exon at position +21, relative to the start ATG codon, resulted
in a frame shift of the predicted g110-encoded protein. Whether this nucleotide deletion represents a genuine missense mutation in the tobacco genome, or an artifact of library construction and cloning, did not bear directly on the evaluation of our splicing method and was not resolved. Instead, the missing nucleotide was restored during PCR amplification of the cDNA for insertion into the bacterial expression vector. The restored open reading frame resulted in a conceptual translation product 95% identical to that for EAS4.

Leaf discs of detached petunia leaves infiltrated with *A. tumefaciens* harboring the CsVMV-g110 construct into detached petunia leaves. Total RNA was isolated from leaf discs 4 d postinfiltration and used for RT-PCR assays. Lane 1, RT-PCR assay without template RNA or PCR primers added; lanes 2 and 9, molecular size standards; lane 3, RT-PCR assay without forward PCR primer added; lane 4, RT-PCR assay without reverse PCR primer added; lane 5, RT-PCR assay without reverse transcriptase added; lane 6, positive size control PCR assay using the tobacco EAS4 cDNA as the template with appropriate primers; lane 7, complete RT-PCR assay for g110 transcript; lane 8, positive size control PCR assay using the intron-containing CsVMV-110 gene construct as template. C, The RT110 cDNA generated by surrogate splicing was expressed in *E. coli*, the encoded protein purified, incubated with the substrate FPP, and the pentane extractable reaction product(s) analyzed by gas chromatography. D, Mass spectrum of the product peak (8.76 min) in C. E, Mass spectrum of an authentic 5-epi-aristolochene standard.
was purified by chromatography over a nickel-affinity column to approximately 95% purity based on Coomassie Blue staining of a denaturing polyacrylamide gel (Mathis et al., 1997). The purified protein was then incubated with (E,E)-FPP, the substrate of known sesquiterpene synthases, under standard conditions (Rising et al., 2000). Gas chromatography showed a single major pentane-extractable product from this reaction (Fig. 2C). The mass spectrum of this major product (Fig. 2D) was identical to a genuine 5-epi-aristolochene standard (Fig. 2E), thus demonstrating that the RT110 cDNA encodes for a functional 5-epi-aristolochene synthase activity.

Functional Characterization of At5g44630 as a Unique Terpene Synthase Gene

To evaluate the general applicability of the surrogate splicing methodology to genes from nonsolanaceous plants, Arabidopsis genes annotated as terpene synthases were considered. Of the 34 to 40 sequences annotated (Aubourg et al., 2002; Chen et al., 2003) and examined (see Supplemental Figs. 1 and 2), nine of these genes have been characterized using either a genetic approach or a variety of methods to isolate full-length cDNAs for heterologous host expression studies. The characterized genes include six monoterpene synthases (Bohmann et al., 2000; Chen et al., 2003, 2004; Fäldt et al., 2003), one sesquiterpene synthase (Chen et al., 2003), and two diterpene synthases (Sun and Kamiya, 1994; Yamaguchi et al., 1998). The remaining genes were then screened for those with the highest sequence similarity to sesquiterpene synthases but containing differences to previously characterized synthases. Gene At5g44630 was noteworthy in this regard because it exhibited similarities to different catalytic classes of sesquiterpene synthases (CADS from cotton, a cadinene-type synthase [Chen et al., 1995]; EAS from tobacco, an eremophilene-type synthase [Facchini and Chappell, 1992]; and HPS from Hyoscyamus, a spirovetivene-type synthase [Back and Chappell, 1995]), as well as features common to diterpene synthases such as abietidiene synthase from Abies grandis (Vogel et al., 1996) and the monoterpene synthase for limonene from mint (Colby et al., 1993). Although monoterpene and sesquiterpenes biosynthesis by Arabidopsis are well documented, no cadinene-type sesquiterpenes or labdane-type diterpenes (other than gibberellins) have been reported (Aharoni et al., 2003; Chen et al., 2003; Steeghs et al., 2004).

The At5g44630 gene was first PCR amplified from genomic DNA using PCR primers bracketing the putative (Fig. 3) start and stop codons, and recombined into a pBluescript II vector modified with a suitable recombination cloning cassette (Hartley et al., 2000). Agrobacteria harboring the cauliflower mosaic virus (CaMV)-At5g44630 construct were subsequently infiltrated into petunia leaves and leaf discs collected 4 to 7 d postinfiltration. Using total RNA isolated from the combined leaf discs, a single RT-PCR product approximately 1,650 bp in size was obtained (Fig. 3B, lane1), which was significantly smaller than the amplification product generated from the original At5g44630 genomic construct (2,327 bp, lane 2). DNA sequencing of the recovered RT-PCR amplification product demonstrated that six intervening sequences had been removed to create an open reading frame of 1,674 bases encoding for a protein of 557 amino acids (GenBank accession no. AY876386). The conceptual cDNA translational product exhibited about the same sequence identity (30%-35%) and similarity (49%-54%) to several well-characterized plant sesquiterpene synthases including EAS (Facchini and Chappell, 1992), HPS (Back and Chappell, 1995), CADS (Chen et al., 1995), and amorpha-4,11-diene synthase (ADS; Chang et al., 2000; Mercke et al., 2000; Wallaart et al., 2001), but much less so to plant monoterpene and diterpene synthases, or sesquiterpene synthases of microbial origins such as trichodiene synthase (Rynkiewicz et al., 2002).

The isolated RT-At5g44630 cDNA was inserted into the pET28a expression vector and transformed into E. coli host cells supplemented with additional tRNA genes considered limiting for eukaryotic gene expression (Rossetta cells, Novagen, Madison, WI). Lysates of IPTG-induced E. coli bearing the pET28a-RT-At5g44630 cDNA construct were recovered after sonication and centrifugation and initially assessed for total monoterpene, sesquiterpene, and diterpene synthase activities. Significant hexane-extractable, radioactive products were observed from incubations with high specific activity [1-^3H]FPP but not with either [1-^3H]GPP or [1-^3H]GGPP (Fig. 3C). Unfortunately, further attempts to purify the Arabidopsis sesquiterpene synthase via an amino-terminal fused His-tag failed. Enzyme activity was completely lost, indicative of the labile nature of this enzyme and/or these preparations. Hence, the sesquiterpene reaction products were initial verified by comparison of lysates from noninduced and IPTG-induced bacteria cultures incubated with nonradioactive FPP and the organic solvent extractable products were evaluated by gas chromatography-mass spectrum (GC-MS; Supplemental Fig. 3). Approximately 16 unique sesquiterpene hydrocarbon reaction products were reproducibly observed in association with extracts from the induced cell extracts, but not the extracts from noninduced cultures. To more fully document the reaction products, we partially purified the synthase activity by rapid anion-exchange chromatography method (Vogeli et al., 1990) and following incubation with FPP, the organic solvent extractable reaction products purified by silica gel chromatography before examining them by GC-MS (Fig. 3, D-G). The three dominant peaks were identified by MS (and quantified by total ion chromatography (TIC)) as α-barbate (constituting 27.3% of the total reaction products), thujopsene (17.8% of the total products), and β-chamigrene (9.9% of the total products). We have no evidence at this time concerning the absolute configuration of the three sesquiterpenes products. Both enantiomers of thujopsene
Figure 3. Surrogate splicing and functional characterization of a putative Arabidopsis terpene synthase gene. A, A cartoon depiction of At5g44630, a terpene synthase gene predicted from the DNA sequence of the Arabidopsis genome to contain six introns (Aubourg et al., 2002). B, RT-PCR recovery of a processed transcript for this terpene synthase after infiltration of A. tumefaciens carrying the CsVMV-At5g44630 construct into detached petunia leaves. Total RNA was isolated from leaf discs collected 4 to 7 d postinfiltration and used for RT-PCR assays. Lane 1, complete RT-PCR assay for the At5g44630 transcript; lane 2, positive size control PCR assay using the intron-containing At5g44630 gene construct as template; and lane M, molecular size standards. C, The At5g44630 cDNA generated by surrogate splicing was expressed in E. coli, and extracts incubated with [3H]GPP, [3H]FPP, and [3H]GGPP to obtain a general measure of mono-, sesqui-, and diterpene synthase activity. D, Synthase activity partially purified from E. coli extracts by anion-exchange chromatography was incubated with FPP and the organic extractable products examined by GC-MS. A total ion chromatogram is shown. The mass spectra for reaction product peaks 1, 2, and 3 are shown relative to published spectra for (E) α-barbatene, (F) thujopsene, and (G) β-chamigrene.
and β-chamigrene have been reported to occur naturally. However, it seems reasonable to suppose that they are the same as those established recently for the same compounds detected in volatile emissions from Arabidopsis flowers (Chen et al., 2003; Scheme II).

The remaining 13 reaction products were qualified as sesquiterpenes by having parent ions of 204 and fragmentation patterns typical for sesquiterpenes and individually accounted for 0.25% to 5% of the total products. Given the dominant α-barbatene reaction product, the At5g44630 locus can now be referred to as an Arabidopsis α-barbatene synthase (AtBS) gene.

**DISCUSSION**

Validation of the Surrogate Splicing Method

Agroinfiltration or in planta transient expression have been used in several novel ways recently to investigate various aspects of the transformation process and gene function. In studies evaluating the contribution of the Ti-plasmid vir genes to the T-DNA transfer process, Narasimhulu et al. (1996) used an RT-PCR amplification assay for an intronless GUS mRNA to establish a time course for the transfer and expression of an intron-containing GUS gene from Agrobacterium to cell cultures of tobacco and maize. Mixed cultures of Agrobacterium lines harboring known tomato resistance genes and the corresponding Cladosporium avirulence genes have also been infiltrated into leaves of plants from several families as a measure of the conservation of downstream components necessary for the development of hypersensitive-type lesions (Van der Hoorn et al., 2000, 2001). More recently, Mitsui et al. (2003) employed a particle bombardment-mediated transient expression system for the study of intron removal efficiency in planta. The agroinfiltration technique has been extended in this report by employing it to remove introns from plant genes and to generate full-length cDNAs (Scheme I). Petunia leaves, which withstood infiltration with high titers of Agrobacterium and supported higher levels of transgene expression than tobacco, were infiltrated with suspensions of *A. tumefaciens* carrying putative terpene synthase genomic clones. The correspondingly processed transcripts were readily recovered as RT-PCR products 4 to 6 d after infiltration and subsequently cloned into a convenient bacterial expression vector. The *E. coli* expressed proteins were functionally characterized in vitro and demonstrated to be sesquiterpene synthases capable of converting farnesyl diphosphate to 5-epi-aristolochene for the tobacco gene and a multifunctional α-barbatene synthase from Arabidopsis.

Splicing of RNA transcripts in plant cells comprises several distinct steps, reviewed by Brown and Simpson (1998), during which a ribonucleoprotein complex, called the spliceosome, is assembled. These steps proceed via the recognition of introns by trans-acting proteins giving rise to (+)-β-chamigrene or thujopsene/β-chamigrene biosynthesis. A methylene migration would generate the chamigrenyl intermediate, which could undergo either a direct proton abstraction to form (+)-β-chamigrene, or a ring expansion followed by a homoallyl-cyclopropylcarbinyl cyclization and proton elimination to yield (+)-thujopsene. The alternative branch pathway for the cuprenyl cation consists of sequential methyl migrations to generate a bazzanenyl intermediate. Instead of deprotonation, a key branch point for the diversion of reaction intermediates to either α-barbatene or thujopsene/β-chamigrene biosynthesis. A chemical rationalization for the major sesquiterpene products generated by the Arabidopsis (+)-α-barbatene synthase. Allylic rearrangement of the diphosphate moiety of all-trans farnesyl diphosphate, (E,E)-FPP, and formation of nerolidyl diphosphate, (3R)-NPP, allows for rotation about the single 2,3 bond. Reionization of the diphosphate group in a cisoid conformation of NPP and presumably concerted anti,endo S_n1 cyclization generates the (3R)-bisabolyl cation. Anti-Markovnikov π-cyclization onto the 10, 11 double bond creates the 5-membered B ring and results in the centering of the secondary carboxylation at C10 of the B ring. A 1,4-hydride shift from the A ring repositions the reactive carboxylation back onto the A ring, which creates a key branch point for the diversion of reaction intermediates to either α-barbatene or thujopsene/β-chamigrene biosynthesis. A methylene migration would generate the chamigrenyl intermediate, which could undergo either a direct proton abstraction to form (+)-β-chamigrene, or a ring expansion followed by a homoallyl-cyclopropylcarbinyl cyclization and proton elimination to yield (+)-thujopsene. The alternative branch pathway for the cuprenyl cation consists of sequential methyl migrations to generate a bazzanenyl intermediate. Instead of deprotonation at C12 as observed for trichodiene biosynthesis (Cane, 1990), another π-cyclization event generates the tricyclic barbatenyl cation that undergoes a final endocyclic deprotonation giving rise to (+)-α-barbatene. Precedent for these predictions is provided by the extensive characterization of trichodiene synthase (Cane, 1990; Rynkiewicz et al., 2002) and in vivo labeling studies of β-barbatene biosynthesis by liverwort cell cultures (Nabeta et al., 1998; Warmers and König, 2000). The absolute configurations of the three sesquiterpene products are assumed to be the same as those established for the same compounds detected in volatile emissions from Arabidopsis flowers (Chen et al. 2003).
The definition of 5’ and 3’ intron splice junctions, the cleavage of the 5’ junction, formation of the “lariat” structure at a nucleotide site within the intron, and finally the cleavage at the 3’ junction and ligation of neighboring exons. Conserved sequence features are posited to interact with the RNA component of snRNPs to distinguish intron-splice junctions, though the absolute sequence requirements remain to be established (Brown and Simpson, 1998). Several studies have been undertaken with the goal of identifying sequences required for 5’ and 3’ splice site definition, either by the comparison of gene sequences (Brown et al., 1996) or by the experimental manipulation of intron or exon sequences to observe the effects on splicing function (Baynton et al., 1996; Egoavil et al., 1997; McCullough and Schuler, 1997; Latijnhouwers et al., 1999). Algorithms have been developed using these data to identify exon/intron boundaries (Hebsgaard et al., 1996; Brendel and Kleffe, 1998; Pavy et al., 1999) for the purpose of predicting gene structure. Similarly, experiments have been undertaken to uncover internal sequences that define the intron branch point needed for lariat formation (Liu and Filipowicz, 1996; Simpson et al., 1996).

Special instances are likely in which surrogate splicing may not result in the generation of mRNAs that are reflective of the native expressed forms. Such could be the case for transcripts that are alternatively spliced to produce distinct mRNA isoforms (Li and Howe, 2001; Mano et al., 1999) or those that use the infrequent U12 splice junction. Several studies have also suggested that monocots and dicots may differ significantly in the factors necessary for splicing (Keith and Chua, 1986; Goodall and Filipowicz, 1991), which argues that a monocot host for surrogate splicing of monocot genes will likely have to be developed. Nonetheless, the surrogate splicing method described here does demonstrate good fidelity with regards to splice site recognition. The generally agreed upon consensus sequences for 5’ and 3’ splice sites in plants are AG/GTAAGT (62%, 79%, 100%, 99%, 70%, 58%, 49%, 53% frequency of specified base) and TGCAG/G (64%, 42%, 95%, 100%, 100%, 57% frequency of specified base), respectively (with “/” representing the splice site; Lorkovic et al., 2000). The six intervening sequences properly processed from the tobacco epi-aristolochene synthase gene in this study are consistent with these consensus sequences, including typical deviations within these sites. The composite 5’ and 3’ sites for the six splice sites are DR/GTRHGW and HXBAG/R. The six intervening sequences processed from the Arabidopsis α-barbatene synthase gene also represent typical splicing events with a composite 5’ sequence of VK/GTRHNH and the 3’ splice junctions consisting of DNYAG/D. Hence, while surrogate splicing will be subject to caveats of developmental and environmental triggered alternative splicing, the method accurately processes multiple introns having a typical range of splice-site sequences found in plant genes, which reflects the importance of other factors including other sequences and contextual positioning of these sequences for high fidelity processing (Lorkovic et al., 2000; Brown et al., 2002).

There are many methods for obtaining expressible cDNA clones for putative genomic genes including construction and screening of cDNA libraries and PCR screens for expressed mRNAs. The surrogate splicing process described here represents another relatively rapid method that only requires 4 to 5 d from infection to RT-PCR recovery of the cDNA. The greater part of the effort before and after these steps entails the manipulation of genomic or cDNA clones into various sequencing and expression vectors. The current method has also been improved by use of a recombination-based cloning system (Hartley et al., 2000) rather than the traditional endonuclease-based cloning system. Adoption of the recombination system alleviates difficulties arising from incompatible restriction sites necessary for ligation steps, and allows for genes to be quickly moved into an Agrobacterium Ti vector, and the resultant cDNA mobilized into an expression vector with minimal manipulations. However, other experiments to better optimize the surrogate splicing method are also warranted. These include a broad screen for other surrogate splicing host plants to accommodate diverse genetic materials and efforts to identify suitable Agrobacterium strains for each plant host. Moreover, the surrogate splicing method needs to be evaluated for its utility to sort out other problematic regions of genomes, especially those purported to be involved in developmental programs and for which no catalytic function can be inferred (i.e. F-box genes).

Use of the Surrogate Splicing Method to Isolate Novel Terpene Synthases

Several studies have demonstrated how difficult gene identification based solely on sequence identity may be (Lehfeldt et al., 2000; Schoch et al., 2001). This is even more problematic when attempting to sort out enzymes like terpene synthases. These are enzymes that utilize a range of allylic diphosphate substrates (GPP, FPP, and GGPP) and can catalyze quite distinct reactions from one another, although still maintaining a very high degree of sequence identity (Trapp and Croteau, 2001; Martin et al., 2004). The availability of the surrogate splicing method described here provides yet another tool for the preparation of processed mRNAs that can be readily cloned and used for functional testing in appropriate host systems.

These results ascribe catalytic function to two genomic clones. While identification of the tobacco g110 gene as an epi-aristolochene synthase is not unexpected given its high degree of similarity to EAS and the existence of 12 to 16 related terpene synthase genes in the tobacco genome (Facchini and Chappell, 1992), the AtBS gene represents a unique catalytic activity. Expression of the At5g44630 gene was previously demonstrated by RT-PCR assays to occur almost exclusively
sesquiterpenes dominated by bidopsis flowers emit a suite of monoterpenes and Aharoni et al. (2003) also demonstrated that Ara-
in floral tissues (Chen et al., 2003). Chen et al. (2003) and Aharoni et al. (2003) also demonstrated that Ara-
and the nature of these particular floral derived compounds and the existence of more than 20 unchar-
characterized sesquiterpene synthase-like genes (Supplemental Fig. 1), it was conceivable that individual
synthases could be responsible for the synthesis of each compound. However, the description of At5g44630 as encoding for a synthase generating α-barbatene, thujopsene, and β-chamigrene in a ratio approximating that observed in floral emissions suggests that a single
enzyme could be responsible for these particular floral sesquiterpenes. Measuring the floral emissions from the available insertion mutants in the At5g44630 locus (SALK insertional mutants 126868, 151777, 147855) will provide a direct test of this assertion.

While precedence for the generation of multiple reaction products of single terpene synthases has been firmly established (Steele et al., 1998), the diverse nature of the bicyclic and tricyclic products of AtBS is quite novel (see Scheme II). Moreover, although α-barbatene, thujopsene, and β-chamigrene have been isolated from a variety of diverse natural sources including marine organisms (Blunt et al., 2004), liver-
noctores the aim to eliminate select, partial reactions, particularly those at branch point positions, thus creating synthases that catalyze the accumulation of reaction intermediates or their der-
vatives (Cane et al., 1997a, 1997b; Rising et al., 2000; Deligeorgopoulou and Allemann, 2003). Such an
approach could also benefit from a comparison of the higher plant enzyme (AtBS) to those enzymes that must exist in the evolutionarily distant, but related liverworts (http://tolweb.org/tree/phylogeny.html) using a combination of sequence comparisons and three-dimensional modeling to identify conserved and potentially functionally important residues (Takahashi et al., 2005).

MATERIALS AND METHODS

Plant Materials and Agroinfiltration

Petunia (Petunia × hybrida) plants were started from commercially avail-
able seed. Plants were grown in a greenhouse with supplemental light provided by sodium vapor lamps. Prior to experimentation, a population of plants was generated by crossing two parental lines chosen for broad rosette leaves suitable for infiltration. Seeds generated from this cross were collected and subsequently maintained as population A. Leaves for experimental infiltration were chosen on the basis of size with a 5-cm width minimum. Leaves were either left on the plant or cut from plants and rinsed in tap water to remove any adhering debris. A brief submersion was also helpful to promote hydration of the leaf, as fully turgid leaves were preferable for infiltration. Immediately prior to infiltration, detached leaves were placed on dampened paper towels in plastic boxes on the lab bench.

Agrobacterium tumefaciens strain GV3850 was transformed by electropora-
tion and maintained under kanamycin and rifampicin selection. Overnight cultures for infiltration were concentrated by centrifugation, resuspended in a 10% Suc solution, recombined, and finally resuspended in 10% Suc to a final concentration of OD600 equal to 0.5 ± 0.05. Addition of 20 mM acetoxyringone 3 h prior to infiltration enhanced in planta expression, but was not necessary. Petunia leaves were nicked on the lower leaf surface, and the bacterial suspension introduced using a needle-less syringe. Infiltrated plants were maintained in the greenhouse while infiltrated leaves were measured in an open plastic container on wet paper towels for up to 1 week.

Leaf discs for GUS enzyme assays (Wallach, 1992) or for RNA isolation were cut with a 1-cm diameter cork borer and stored at −80 °C until processing. RNA isolation was performed using Trizol according to the manufacturer’s instructions (Invitrogen, Carlsbad, CA).

Molecular Manipulations

GUSI (Vancanneyt et al., 1990) and the tobacco g110 locus were PCR amplified with primers harboring convenient restriction sites, digested accordingly, and ligated into a Ti plasmid derived from pBI101 (CLONTECH, Palo Alto, CA), 5’ to the CaMV promoter. The Arabidopsis (Arabidopsis thaliana) At5g44630 gene was amplified with primers containing recombination cloning sites and introduced into a modified pBluescript vector (CLONTECH) where a recombin-

attained GUSi (Vancanneyt et al., 1990) and the tobacco g110 locus were PCR amplified with primers harboring convenient restriction sites, digested accordingly, and ligated into a Ti plasmid derived from pBI101 (CLONTECH, Palo Alto, CA), 5’ to the CaMV promoter. The Arabidopsis (Arabidopsis thaliana) At5g44630 gene was amplified with primers containing recombination cloning sites and introduced into a modified pBluescript vector (CLONTECH) where a recombin-

Received for publication June 17, 2005 and in revised form July 14, 2005.

www.plantphysiol.org on October 30, 2017 - Published by Downloaded from
was filtered through a 0.45-micron filter and the His-tagged protein purified. The lysate was recovered after a 20-min centrifugation at 38,000 g.


Bacterial Expression Studies

The RT-110 BamHI-SstI digestion fragment was ligated into the corresponding sites within the pET-28a vector in-frame with a N-terminal hexahistidyl tag (Novagen, Madison, WI), pBluescript (Strategene, San Diego) for DNA sequence analysis. All the DNA sequencing reactions were performed using the BigDye Terminator Cycle sequencing kit (Perkin-Elmer, Wellesley, MA) with the BigDye Terminator Cycle sequencing kit (Applied Biosystems, Foster City, CA).

Enzyme Assays and Reaction Product Analyses

Initial terpene synthase activities were measured with radioactive [3H]GPP (ARC, St. Louis, 20 Ci/mmole, [3H]GPP (Perkin-Elmer, Boston, 61.1 Ci/mmole), or [14]GPP (Perkin-Elmer, Boston, 23 Ci/mmole) using standard reaction conditions (Rising et al., 2000). Typically, cell lysates, or partially purified enzymes, containing up to 30 μg of protein, were incubated with 0.2 to 0.4 μM of high specific labeled (16-20 Ci/mmole) substrates for 5 to 20 min in 50-μL reactions containing 20 mM Tris-HCl, pH 7.4, 20 mM MgCl₂, and 0.2 mM MnCl₂ at 30°C. The radioactive reaction products were extracted with three reaction volumes of hexane or pentane, and the organic extracts were then mixed with a small amount of silica gel to remove oxygenated products. Aliquots of the organic extracts were then quantified by liquid scintillation counting.

The product profile for the Ni²⁺ affinity column-purified RT-110 enzyme was determined by incubating approximately 100 nM enzyme with 46 μM FPP in a 2.5-μL reaction. Reactions were incubated 30 min, then extracted twice with 2 mL of pentane, which was concentrated under nitrogen to approximately 50 μL for GC-MS analysis. GC-MS analysis was performed with an HP-GC plus equipped with a DB-5ms capillary column (30 m x 0.25 mm, 0.25-μm phase thickness) and run with He as the carrier gas at 1 mL/min. Sample injections were splitless with an injection port temperature of 250°C. The oven was programmed to hold at 100°C for 1 min and then increased to 270°C with an 8°C/min ramp. The dominant product was compared to an authentic sample of s-epi-athulolutechene (Zhao et al., 2004).

The reaction product profile for the anion-exchange purified A15y44630 enzyme was determined similarly. Partially purified enzyme (4 mg) was incubated with 11.5 μM FPP in a total volume of 10 mL of reaction buffer (20 mM Tris-HCl, pH 7.4, 20 mM MgCl₂, 0.2 mM MnCl₂) for 30 min before sequential extractions with 10 mL of hexane. The hexane extracts were concentrated to 1 mL, passed over a 1-mL silica gel column, and the hexane eluate was reconstituted to 30 μL. Reaction products were identified by MS using a Thermo Finnigan DSQ GC-MS system equipped with a Restec Rtx-5 capillary column (30 m x 0.32 mm, 0.25-μm phase thickness). One-microliter samples were injected in the splitless mode at 250°C with an initial oven temperature of 70°C for 1 min followed by a 4°C per min gradient to 230°C. Mass spectra were recorded at 70 eV, scanning from 35 to 300 atomic mass units, and compared to NIST and MassFinder library standards for verification. Tentative compound identifications are based on standard NIST and MassFinder computer algorithms matching unknowns to library standards with similarity values in excess of 30%, visual comparison of the mass spectral patterns, and retention time comparisons to those reported in Chen et al. (2003).

Distribution of Materials

Upon request, all materials described in this publication will be made available in a timely manner for noncommercial research purposes, subject to the requisite permission from any third-party owners of all or parts of the material. Obtaining all such permission will be the responsibility of the requestor.

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under accession numbers AY313939 (g110) and AY856386 (RT-At5g44630).

Received January 6, 2005; returned for revision February 23, 2005; accepted February 24, 2005; published online June 17, 2005.

LITERATURE CITED


Blunt JW, Copp BR, Munro MHG, 730–745.


AAB TAT GAA GAA GA-3

GTT GC


Keith B, Chua NH (1986) Monocot and dicot pre-messenger RNAs are processed with different efficiencies in transgenic tobacco. EMBO J 5: 2419–2425


Copyright © 2005 American Society of Plant Biologists. All rights reserved.


