Molecular Cloning and Expression Analysis of a Putative Terpene Synthase Gene from Citrus

Ying Jia, Dianren Xia, and E.S. Louzada

Texas A&M University–Kingsville, Citrus Center, 312 N. International Boulevard, Weslaco, TX 78596

Abstract. A cDNA coding for a putative terpene synthase (Grtps) was isolated from ‘Rio Red’ grapefruit (Citrus paradisi Macf.) mature fruit by differential display RT-PCR and the corresponding full-length cDNA and genomic clone were subsequently obtained. The isolated cDNA clone was 1644 bp in length encoding a protein of 548 amino acids with a predicted molecular mass of 64 kDa and of pI 5.38. The genomic clone was 3203 bp in length with 6 introns and 7 exons. This Grtps appears to be a sesquiterpene synthase based on molecular weight, genomic organization, and similarity with the other terpene synthases. Both RT-PCR and Northern blot expression analysis indicated that Grtps is not expressed in immature fruits, roots, or leaves, but only in mature fruits. Southern blot analysis of genomic DNA demonstrated that Grtps is one of the members in the family of terpene synthases.

Terpenoids are the largest class of natural products, containing well over 30,000 individual compounds (Buckingham, 1998), and are present and usually abundant in all phyla, and have diverse functions in primary metabolism and in ecological interactions (Trapp and Croteau, 2001). Recently, a family of 40 terpenoid synthase genes was isolated from Arabidopsis thaliana (L.) Heynh., which is the largest and most diverse group of TPS gene so far isolated from one plant species (Aubourg et al., 2002). This large family may be involved in the production of terpenoids that are accumulated in small amounts into the plant, released from plant into the environment, or rapidly metabolized to other substances, since A. thaliana lacks the anatomic structures associated with the accumulation of terpenoids (Füldt et al., 2002). All terpene synthases are very similar in physical and chemical properties.

Sequence comparison of terpene synthases within and between subfamilies and characterization of the corresponding enzymes demonstrated that sequence relatedness allows identification of members of the three major classes of terpene synthases, monoterpene, sesquiterpene, and diterpene synthases (Bohlmann et al., 1998). Trapp and Croteau (2001) compared the plant terpene synthases according to their genomic structure, including intron number, exon size, and intron position and classified them into three classes, 12–14 introns (class I), 9 introns (class II), or 6 introns (class III).

The terpenoids, monoterpene(C10), sesquiterpene(C15), and diterpene(C20), are derived from the prenyl diphosphate substrates geranyldiphosphate (GPP), farnesyl diphosphate (FPP), and geranylgeranyldiphosphate (GGPP), respectively, by numerous variations on a common electrophilic reaction mechanism (Gershenzon and Croteau, 1993). They play numerous functional roles in primary metabolism and in ecological function for communication or defense (Back and Chappell, 1995; Hahlbrock et al., 1995; Harborne, 1991). The majority of terpenoids have traditionally been referred to as secondary metabolites or as natural products. Terpenoid secondary metabolites are abundant in many essential oils (Lawrence, 1992), resins (Martin et al., 2002), and floral scents (Dudareva and Pichersky, 2000), and several are of pharmacological significance, including the monoterpoid dietary anticarcinogen limonene (Crowell and Gould, 1994) and the terpenoid anticancer drug Taxol (Holmes et al., 1995).

Terpenoid synthases (monoterpene, sesquiterpene, and diterpene synthases) serve to convert the central precursor isopentenyl diphosphate to the parent skeletons of the various terpene types (Ramos-Valdivia et al., 1997; Trapp and Croteau, 2001), since the products of the reactions are most often cyclic the terpene synthase are also called cyclases. This family of enzymes not only catalyzes the committed step in the biosynthesis of most terpenoids but also is largely responsible for the great structural diversity encountered in this class of natural products (Davis and Croteau, 2000).

In this paper, we describe the isolation, molecular characterization, and the expression pattern of a putative grapefruit terpene synthase (Grtps) from juice vesicles of ‘Rio Red’ grapefruit.

Materials and Methods

Plant materials. Immature fruits of ≈4.5 cm in diameter with no color developed, and mature fruits of ≈10 cm that had the red internal color fully developed were harvested from ‘Rio Red’ grapefruit tree. Juice vesicles were collected from the fruits, frozen in liquid nitrogen, and maintained at −80 °C for future use. Juice vesicles were also collected from mature fruits of ‘White Marsh’ grapefruit and ‘Marrs’ sweet orange [Citrus sinensis (L.) Osb.]. Leaves and roots were collected from ‘Rio Red’ grapefruit, frozen in liquid nitrogen, and maintained at −80 °C for future use.

RNA and DNA isolation. Total RNA was isolated from leaves, roots, and juice vesicles of immature and mature fruits using RNAmix or ToTally RNA kits (Ambion, Austin, Tex.) following the manufacturer’s instruction. DNA contamination was eliminated by a DNAse I RNAse free (Amersham Biosciences, Piscataway, N.J.) treatment. Genomic DNA was isolated using plant DNAzol Reagent according to the instruction (Invitrogen, Carlsbad, Calif.).

Differential display RT-PCR and isolation of full-length cDNA. Differential display RT-PCR experiments were performed using RNAmix kit (GeneHunter Corp., Nashville) according to the manufacture instructions, as briefly described. Two hundred nanograms of total RNA from juice vesicles of immature and
mature 'Rio Red' grapefruit fruits were reverse transcribed using Omniscript reverse transcriptase (Qiagen, Valencia, Calif.) and 1 µL of either one base anchor primers (HT11A, HT11C, and HT11G). Aliquots of 2 µL of the obtained cDNA were amplified using each anchor primers (1 µL) along with 1 µL of either one of the AP random primers and HotStart Taq DNA polymerase (Qiagen). PCR products were labeled by α-[32P]dATP and separated in 6% denaturing polyacrylamide sequence gels, Sequagel-6 (National Diagnostic, Atlanta). After separation, the gels were dried without fixing and exposed to Kodak Biomax MR (Eastman Kodak Co., Rochester, N.Y.) film at –80 °C. Bands from transcripts up-regulated in mature ‘Rio Red’ grapefruit were reamplified, validated by reverse northern blot, cloned into PCR-TRAP vector (GeneHunter Corp.), and sequenced.

Rapid amplification of cDNA ends (RACE) and isolation of genomic clone. Full-length Grtps cDNA clones was obtained by RLM-5’ RACE using the GeneRacer Kit (Invitrogen), following the manufacturers instructions. Briefly, 5 µg of total RNA from mature ‘Rio Red’ grapefruit was reverse transcribed using the GeneRacer oligo dT primer and SuperScript II reverse transcriptase (Invitrogen). In subsequent PCRs, a gene-specific reverse primer (5’-CGT CAAAATGGAACGTGGTCTCCTA-3’) was synthesized based on the sequence of Grtps partial cDNA clone in combination with the GeneRacer 5’ primer (5’-CGACTGGAG-GCAAGGCACACT GA-3’). The RACE-PCR products were separated on 1.2% agarose gel and cloned into TOPO TA cloning system (Invitrogen) according to the manufacturer instructions.

The genomic clone was obtained by PCR using genomic DNA as a template and primers flanking the Grtpls open reading frame. Sequencing was performed by primer walking. Similarity search was performed at the GenBank database using blastn and blastx, and alignment was performed using ClustalX (European Bioinformatics Institute, 2004) and boxshade software (Institute Pasteur, 2004).

Northern blot and reverse-transcription-PCR. Total RNA (15 µg) from juice vesicles of mature and immature ‘Rio Red’ grapefruit, and from roots and leaves were electrophoresed on 1.2% agarose/formaldehyde gels and blotted onto nylon membranes (Hybond N+; Amersham Biosciences). The partial Grtps clone fragment of 320 bp (includes the 3’ UTR plus 62 nucleotides upstream of the stop codon) was radiolabeled by the Nick-translation method (Invitrogen) and used as a probe. Hybridization was carried out at 65 °C overnight in a modified Church buffer (Church and Gilbert, 1984) containing 7% SDS, 0.5 M EDTA, 0.5 M sodium phosphate, and 1% BSA. Blots were washed twice with 2× SSC, 0.1% SDS for 20 min, 1× SSC, 0.1% SDS for 10 min at 65 °C and exposed to Biomax MS film (Eastman Kodak Co.) with intensifying screen at –80 °C.

Southern hybridization. Genomic DNA (15 µg) was digested with EcoRI, BamHI, and HindIII, electrophoresed in 0.8% TAE agarose gels, and transferred onto Hybond N+ nylon membrane (Amersham Biosciences) by alkaline blotting (0.4 × NaOH). The membrane was hybridized with α-[32P]dCTP-labeled Grtps partial clone (320 bp) and with the Grtps entire coding region cDNA at 65 °C overnight in a modified Church buffer (Church and Gilbert, 1984) containing 7% SDS, 0.5 M EDTA, 0.5 M sodium phosphate, and 1% BSA. Blots were washed twice with 2× SSC, 0.1% SDS for 20 min, 1× SSC, 0.1% SDS for 10 min at 65 °C and exposed to Biomax MS film (Eastman Kodak Co.) with intensifying screen at –80 °C.

Expression of Grtps in Escherichia coli. PCR-generated cDNA encompassing the complete coding sequence of Grtps was inserted into the cloning site of the protein expression vector pET101/D-TOPO (Invitrogen). The construct pET101/D-TOPO-Grtps was introduced into E. coli BL21Star (DE3) cells and cultured at 37 °C in Luria-Bertani (LB) medium containing 100 µg·mL–1 ampicillin. When the OD600 of the culture reached 0.5–0.8 they were induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 1 mM. The cultivation was continued and cells were harvested in different times intervals and suspended in tris-glycine SDS buffer (2×) (Invitrogen). The proteins were separated by Novex tris-glycine polyacrylamide gel (Invitrogen) according to the manufacturer instructions and staining with Coomassie blue. In parallel, lysates of E. coli cells transformed with pET101/D-TOPO containing no insert was prepared in the same way and used as the control.

Results and Discussion

Isolation and characterization of Grtps cDNA. Using differential display RT-PCR and reverse northern blot we identified one fragment of 320 bp expressed uniquely in mature fruit of ‘Rio Red’ grapefruit, which was cloned and sequenced. Northern blot analysis (Fig. 1) indicated that this partial transcript is not expressed in immature fruits, leaves or roots of ‘Rio Red’ grapefruit. Similarity search at the GenBank database revealed no similarity at the nucleotide level but a high similarity at the deduced amino acid level with a terpene synthase from Citrus junos Sieb. ex Tan. We designated it as partial Grtps, for partial grapefruit terpene synthase cDNA.

To obtain the full-length Grtps cDNA, RLM-5’ RACE was performed as previously described. The full-length Grtps obtained is composed by 1877 nucleotides, including an open reading frame (ORF) of 1644 bp (GeneBank accession no. AF411120), encoding a polypeptide of 548 amino acid with a calculated molecular mass of 64 kDa and a pl of 5.38.

In general terpene synthases encode proteins of 550–850 amino acids and molecular masses of 50–100 kDa. Monoterpenes synthases are in the range of 600–650 amino acids and are smaller than diterpene synthases by ≈210 amino acids, and larger than sesquiterpene synthases by 50–70 amino acids (Bohmann et al., 1998; Chen et al., 1995; Facchini and Chappell, 1992). The length of the deduced Grtps protein (548 amino acids) and the predicted molecular weight (64,000) is in the range of other sesquiterpene synthases.

To verify the expression of the full-length Grtps in different varieties and organs, we used RT-PCR. RNA was isolated from mature fruit of ‘White Marsh’ and ‘Rio Red’ grapefruit; ‘Marrs’ sweet orange; and from immature fruit, leaves, and roots of ‘Rio Red’ grapefruit. Five micrograms of total RNA was reverse tran-
PCR amplification was performed with the SuperScript II (Invitrogen) and Oligo(dT)12-18. The PCR product was used as a probe for hybridization to the orfs of the mature, immature, and leaf RNA samples. The hybridization was performed using the HotStarTaq DNA polymerase and primers specific for the orf of the target transcripts. The hybridization was detected as a single band on the gel stained with ethidium bromide.

The results indicate (Fig. 2) that the transcript is expressed exclusively in mature fruits and is not specific for the red grapefruit but also for the white grapefruit, as well as for orange. As seen previously, no expression was detected in immature fruits, leaves, and roots of the 'Rio Red' grapefruit. The main reason to use RT-PCR instead of Northern blot is that, as will be discussed later, the transcript is part of a terpene synthase family in citrus and it seems to share many conserved motifs with other enzymes in the same group, therefore, using the ORF as probe would allow hybridization to many other terpene synthases in addition to the target transcript. Additionally, the southern blot using the partial ORF as probe would allow hybridization to many other transcripts, which may complicate the analysis.

For functional expression, the cDNA was subcloned into pET101/D-TOPO vector and expressed in E. coli BL21 at a final concentration of 1 mM IPTG. The desired protein was detected as a single band on the gel stained with ethidium bromide. Additionally, the expression was confirmed by protein gel staining with ethidium bromide as a control. The protein was purified from the recrystallization buffer, and its molecular weight was determined by SDS-PAGE. The protein was then analyzed by MALDI-TOF mass spectrometry to confirm the identity of the protein.

The alignment of the deduced amino acid sequence of the target transcript with other known terpene synthase sequences from various plant species is shown in Fig. 4. The sequence shows high similarity with terpenoid synthases from several plant species, including the highly conserved asparagine-rich motif (DDXXD), which is found in all terpenoid synthases, and is recognized as the site for the divalent metal ion of the substrate diphosphate-metal ion complex (Tarshis et al., 1994). The conserved histidine, cysteine, and arginine residues, involved in catalysis by terpene synthases, are also conserved in the target transcript.

The activity of the expressed protein was detected by reverse-transcription-PCR (RT-PCR) of 5 µg total RNA isolated from 'White Marsh' grapefruit mature fruit (W), 'Maris' sweet orange mature fruit (MO), 'Rio Red' grapefruit mature fruit (M), immature fruit (I), roots (R), and leaves (L), using primers flanking the Grtps ORF. RNAs were visualized by staining with ethidium bromide as equal loading controls.
synthases from both angiosperms and gymnosperms (Munck and Croteau, 1990; Rajaonarivony et al., 1992; Savage et al., 1994, 1995), are present in
Grtps. The conserved motif RRx8W, which is the characteristic of most of monoterpene synthases (Aubourg et al., 2002; Bohlmann et al., 1998), is missing in the
Grtps cloned from Citrus (Fig. 4).

The comparison of Grtps amino acid sequence with sequences of other terpene synthase indicated that this Grtps cDNA is truncated at the 5’ terminus (Fig. 4), the truncation represents 18 amino acid residues of a presumptive transit peptide region at the N-terminus of the deduced protein. This peptide is found in both monoterpene and diterpene synthases and supposedly it facilitates the import of these nuclear encoded gene products into plastids, a process that involves cleavage of the pre-proteins to the mature active enzyme (Colby et al., 1993; Mau and West, 1994; Vogel et al., 1996). This also supports the assumption that Grtps may be a sesquiterpene synthase.

Fig. 4. Alignment of deduced amino acid sequence of Tps cloned from citrus with other terpene synthase. 1) Citrus paradisi, terpene synthase (AF411120); 2) C. jen os, terpene synthase (AF288465); 3) C. jen os, beta-farnesene synthase (AF374462); 4) Lycopersicon hirsutum Dunal, sesquiterpene synthase (AF2794); 5) Arabidopsis thaliana, monoterpene synthase (At4g16740). This alignment was created using Clustal X and boxshade. The absolutely conserved aspartate rich region DDXD of terpene synthases is marked. The conserved histidine, cysteine, and arginine residues are marked with arrows.

Southern blot analysis of Citrus genomic DNA digested with EcoRI, BamHI, and HindIII enzymes and hybridized with a probe containing Grtps ORF revealed a large number of bands, indicating that a high number of Tps genes are present in the citrus genome (Fig. 5b). Since the Northern blot analysis (Fig. 1), using the partial Grtps clone (320 bp, comprising the far 3’ end ORF plus 3’ UTR) as probe, indicated that this gene is specifically expressed in mature fruit, we decided to perform another Southern blot using this fragment as probe, in high stringent conditions. Figure 5a shows the results for the Southern blot of genomic DNA digested with EcoRI, BamHI, and HindIII and hybridized with the partial Grtps. The results revealed a single band for each of the enzyme digests, indicating that Grtps is a single-copy gene in the citrus genome.

Genomic Organization. The genomic clone (3 kb) corresponding to Grtps was obtained by PCR amplification using primers flanking the ORF. The intron junctions, mapped by nucleotide sequence comparison of the cDNA and the genomic clone, revealed that Grtps gene is structured as seven exon segments interrupted by six introns with a range in size from 91 to 497 bp (Fig. 6). Trapp and Croteau (2001) classified the terpene synthase genes into three classes (class I, class II, and class III), which appear to have evolved sequentially from class I to class III by intron loss and loss of the conifer diterpene internal sequence domain. Class I Tps comprise 12–14 introns and 13–15 exons, class II Tps genes comprise 9 introns and 10 exons, and class III Tps genes comprise 6 introns and 7 exons. According to this classification the
**Grtps gene belongs to class III.** According to Trapp and Croteau (2001) this terpene synthase class comprises only angiosperm monoterpene, sesquiterpene, and diterpene synthases involved in the secondary metabolism.

**Literature Cited**

Aubourg, S., A. Lecharny, and J. Bohlmann. 2002. Genomic analysis of the terpenoid synthase (AtTPS) gene family of Arabidopsis thaliana. Mol. Genet. Genomics 267:730–745.

Back, K. and J. Chappell. 1995. Cloning and bacterial expression of a sesquiterpene cyclase from Hyoscyamus muticus and its molecular comparison to related terpene cyclases. J. Biol. Chem. 270:7375–7381.

Bohlmann, J., G. Meyer-Gauen, and R. Croteau. 1998. Plant terpenoid synthases: Molecular biology and phylogenetic analysis. Proc. Natl. Acad. Sci. USA 95:4126–4133.

Buckingham, J. 1998. Dictionary of natural products on CD-ROM, version 6.1. Chapman and Hall, London.

Church, G.M. and W. Gilbert. 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA 81:1991–1995.

Colby, S.M., W.R. Alonso, E.J. Katahira, D.J. McGarvey, and R. Croteau. 1993. 4S-limonene synthase from the oil glands of spearmint (Mentha spicata). cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpene cyclase. J. Biol. Chem. 268:23016–23024.

Crowell, P.L. and M.N. Gould. 1994. Chemoprevention and therapy of cancer by d-limonene. Critical Rev. Oncogenesis 5:1–22.

Davis, E.M. and R. Croteau. 2000. Cyclization enzymes in the biosynthesis of diterpenoids. Arch. Biochem. Biophys. 324:255–266.

Dudareva, N. and E. Pichersky. 2000. Biochemical and molecular genetic aspects of floral scents. Plant Physiol. 122:627–633.

European Bioinformatics Institute. 2004. ClustalW submission form. 12 Dec. 2004. <http://bioweb.pasteur.fr/seqanal/interfaces/boxshade.html>.

Facchini, P.J. and J. Chappell. 1992. Gene family for an elicitor-induced sesquiterpene cyclase in tobacco. Proc. Natl. Acad. Sci. USA 89:11088–11092.

Fäldt, J., G. Arimura, J. Gershenzon, J. Takabayashi, and J. Bohlmann. 2002. Functional identification of AtTPS03 as (E)-farnesyl cyclase: A monoterpene synthase catalyzing jasmonate- and wound-induced volatile formation in Arabidopsis thaliana. Planta 216:745–751.

Gershenzon, J. and R. Croteau. 1993. Terpenoids biosynthesis: The basic pathway and formation of monoterpens, sesquiterpens, and diterpens. p. 339–388. In: T.S. Moore, Jr. (ed.). Lipid metabolism in plants. CRC Press, Boca Raton, Fla.

Hahlbrock, K., D. Scheel, E. Logemann, T. Nurnberger, M. Parniske, S. Reinold, W.R. Sachs, and E. Schmelzer. 1995. Oligopeptide elicitor-mediated defense gene activation in cultured parsley cells. Proc. Natl. Acad. Sci. USA 92:4150–4157.

Harborne, J.B. 1991. Recent advances in the ecological chemistry of plant terpenoids, p. 396–426. In: J.B. Harborne and F.A. Tomas-Barberan (eds.). Ecological chemistry and biochemistry of plant terpenoids. Clarendon Press, Oxford.

Holmes, F.A., A.P. Kiudelka, J.J. Kavanagh, M.H. Huber, and J.A. Ajant. 1995. Current status of clinical trials with paclitaxel and docetaxel, p. 31–57. In: G.I. Georg, T.T. Chen, I. Ojima, and D.M. Vyas (eds.). Taxane anticancer agents: Basic science and current status. Amer. Chem. Soc. Symp. Ser. 583, Washington, D.C.

Institute Pasteur. 2004. Boxshade: Printouts from multiple-aligned protein or DNA sequences. 12 Dec. 2004. <http://bioweb.pasteur.fr/seqanal/interfaces/boxshade.html>.

Lawrence, B.M. 1992. Essential oils as sources of natural aroma chemicals. Perfumer Flavorist 17:15–28.

Martin, D., D. Tholl, J. Gershenzon, and J. Bohlmann. 2002. Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis and terpenoid accumulation in developing xylem of norway spruce (Picea abies) stems. Plant Physiol. 129:1003–1018.

Mau, C.J.D. and C.A. West. 1994. Cloning of casbene synthase cDNA: Evidence for conserved structural features among terpenoid cyclases in plants. Proc. Natl. Acad. Sci. USA 91:8497–8501.

Munck, S.L. and R. Croteau. 1990. Purification and characterization of the sesquiterpene cyclase patchoulol synthase from Pogostemon cablin. Arch. Biochem. Biophys. 282:58–64.

Rajaonarivony, J.I.M., J. Gershenzon, and R. Croteau. 1992. Characterisation and mechanism of (4S)-limonene synthase, a monocyclase from the glandular trichomes of peppermint (Mentha piperita). Arch. Biochem. Biophys. 296:49–57.

Ramos-Valdivia, A.C., R.V.D. Heiden, and R. Verpoorte. 1997. Isopentenyl diphosphate isomerase: A core enzyme in isoprenoid biosynthesis. A review of its biochemistry and function. Natural Product Rpts. 14:591–603.

Savage, T.J., M.W. Hatch, and R. Croteau. 1994. Monoterpene synthases of Pinus contorta and related conifers. J. Biol. Chem. 269:4012–4020.

Savage, T.J., H. Ichii, S.D. Hume, D.B. Little, and R. Croteau. 1995. Monoterpene synthases from gymnosperms and angiosperms: Stereo-specificity and inactivation by cysteiny1- and arginyl-directed modifying reagents. Arch. Biochem. Biophys. 320:257–265.

Tarshis, L.C., M. Yan, C.D. Poulter, and J.C. Sacchettini. 1994. Crystal structure of recombinant farnesyl diphosphate synthase at 2.6 Å. Biochemistry 33:10871–10877.

Trapp, S.C. and R. Croteau. 2001. Genomic organization of plant terpene synthases and molecular evolutionary implications. Genetics 158:811–832.

Vogel, B.S., M. Wildung, G. Vogel, and R. Croteau. 1996. Abietadiene synthase from grand fir (Abies grandis). J. Biol. Chem. 271:23262–23268.