



## Terpene production in the peel of sweet orange fruits

Marco A. Takita<sup>1,2</sup>, Irving J. Berger<sup>1</sup>, Ana Carolina Basílio-Palmieri<sup>1</sup>, Kleber M. Borges<sup>1</sup>, Juliana M. de Souza<sup>1</sup> and Maria L.N.P. Targon<sup>1</sup>

<sup>1</sup>Centro APTA Citros Sylvio Moreira, Instituto Agronômico de Campinas, Cordeirópolis, SP, Brazil.

<sup>2</sup>Centro de Pesquisa e Desenvolvimento de Recursos Genéticos Vegetais, Instituto Agronômico de Campinas, Campinas, SP, Brazil.

### Abstract

Terpenoids constitute the largest and most diverse class of natural products. They are important factors for aroma and flavor, and their synthesis is basically done from two compounds: isopentenyl diphosphate and dimethylallyl diphosphate. Isopentenyl diphosphate is synthesized through two different pathways, one that occurs in the cytoplasm and one in the plastid. With the sequencing of ESTs from citrus, we were able to perform *in silico* analyses on the pathways that lead to the synthesis of terpenes as well as on the terpene synthases present in sweet orange. Moreover, expression analysis using real-time qPCR was performed to verify the expression pattern of a terpene synthase in plants. The results show that all the pathways for isopentenyl diphosphate are present in citrus and a high expression of terpene synthases seems to have an important role in the constitution of the essential oils of citrus.

*Key words:* EST, fruit, terpenoids, orange, essential oil.

Received: September 21, 2006; Accepted: July 13, 2007.

### Introduction

Citriculture plays a fundamental role in Brazilian agrifood. The main activity is the production of concentrated frozen juice where the State of Sao Paulo is the main producer, processor, and exporter. It is responsible for an annual income of around 1.5 billion dollars in exports, which is not restricted to orange juice since pectin, animal feed, and essential oils are an important part of these revenues. The essential oil of sweet orange is a byproduct of the juice production and therefore, makes Brazil the main producer in the world.

Plants synthesize a wide range of low molecular weight compounds that do not show any obvious function in growth and development, called secondary metabolites. These metabolites are also called natural products and their studies led to a current focus on the discovery of new drugs, antibiotics, insecticides, and herbicides (Croteau *et al.*, 2000). The natural products can be divided into three different groups: alkaloids, phenylpropanoids and phenolic compounds, and terpenoids, which constitute the largest and the most diverse class. The vast majority of the terpenoids are classified as secondary metabolites whose functions can be

ecological, providing defense against herbivores or pathogens, attracting animals that disperse pollen and seeds, or inhibiting germination of neighbor plants (Langenheim, 1994; Bouwmeester *et al.*, 1999; Dicke, 1999; Pichersky and Gershenzon, 2002). The biosynthesis of terpenes is divided into four steps: synthesis of the precursors, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) through two different pathways; repetitive addition of the precursors to form a series of homologues of prenyl diphosphate, which are the immediate precursors of the different classes of terpenes; elaboration of the terpenes backbones by the activity of specific synthases; and secondary enzymatic modification of these backbones that will give the functional properties and family diversity. The two pathways responsible for the synthesis of IPP and DMAPP are known as the acetate-mevalonate pathway and non-mevalonate pathway and are localized in different cellular compartments: one in the cytoplasm and the other in plastids, respectively (Adam *et al.*, 1999; Lichtenthaler, 1999). In general, the former leads to the synthesis of some sesquiterpenes and triterpenes (sterols) while the later is responsible for the synthesis of monoterpenes, diterpenes, tetraterpenes (carotenoids) and polyterpenes.

Since several different enzymes participate in the synthesis of terpene precursors and terpenes, in this work we analyzed the presence of transcripts for most of these

enzymes within the expressed sequence tags (EST) dataset of the CitEST Project, the Brazilian initiative for *Citrus* EST sequencing coordinated by the Centro APTA Citrus Sylvio Moreira. First we identified transcripts for all enzymes involved in the synthesis of geranyl diphosphate, farnesyl diphosphate, and geranylgeranyl diphosphate, in the fruit epicarp of sweet orange. We also identified terpene synthases within the whole *Citrus sinensis* CitEST database and evaluated the ones expressed in the pericarp of the fruits since this is the major source for essential oil in citrus. We also confirmed some of the data for the relative abundance of one terpene synthase in different tissues by quantitative real time PCR.

## Material and Methods

### cDNA libraries

EST data were generated in the CitEST Project (citest.centrodecitricultura.br). The dataset included all the *Citrus sinensis* ESTs present in the CitEST databank. For fruit libraries, the pericarp of fruits measuring 1, 2.5, 5, 7, 8, and 9 cm in diameter was used. In brief, total RNA was extracted from 1 g of tissue using the TRIzol Reagent (Invitrogen Corp., Carlsbad, CA). Poly A+ RNA was isolated from 0.5 mg of the total RNA using the mRNA Isolation System (Promega Corporation, Madison, WI). The cDNA libraries were constructed with the SuperScript Plasmid System with Gateway Technology for cDNA Synthesis and Cloning kit (Invitrogen), which uses a primer consisting of an oligo (dT) sequence with a *NotI* restriction site for the cDNA synthesis together with *SaII* adapters to directional cloning. The cDNA fragments were cloned into the pSPORT 1 vector and transformed into *E. coli* DH5 $\alpha$  cells.

The plasmids DNA was extracted by the boiling method (Marra *et al.*, 1999). The sequencing reactions were done using the Big Dye Terminator kit v3 and v3.1 (Applied Biosystems). Sequencing was carried in ABI 3700 and 3730 automatic DNA analyzers (Applied Biosystems).

The electropherograms obtained from the sequencer were submitted to the database through a web interface and stored for analysis.

### Searches for sequences in the database and processing

The search for sequences encoding components of the terpene biosynthesis pathway was performed using the tblastn tool (Altschul *et al.*, 1997). These searches were carried out using sequences from *Arabidopsis thaliana* that encode the different proteins of the terpene biosynthesis pathways available in GenBank (Table 1). For assembly, the sequences were clustered by the local ongoing project manager "Gene Projects," a tool available in CitEST that allows searching, clustering, and analysis of reads and the formed contigs. Clusters that did not show similarity to en-

**Table 1** - Enzymes of *Arabidopsis thaliana* used for the search of the CitEST database. The enzymes listed in the first column were searched in the CitEST database using the genes of *Arabidopsis thaliana*, listed in the last column. In column 2 are listed the EC numbers for each enzyme obtained from KEGG.

Name	EC number	<i>A. thaliana</i> accession number
acetoacetyl CoA thiolase	2.3.1.9	At5g48230 At5g47720
hydroxymethylglutaryl-CoA synthase	2.3.3.10	At4g11820
hydroxymethylglutaryl-CoA reductase	1.1.1.34	At1g76490 At2g17370
mevalonate kinase	2.7.1.36	At5g27450
phosphomevalonate kinase	2.7.4.2	At1g31910
diphosphomevalonate decarboxylase	4.1.1.33	At2g38700 At3g54250
1-deoxy-D-xylulose-5-phosphate synthase	2.2.1.7	At3g21500 At4g15560 At5g11380
1-deoxy-D-xylulose-5-phosphate reductoisomerase	1.1.1.267	At5g62790
2-C-methyl-D-erythritol 4-phosphate cytidyltransferase	2.7.7.60	At2g02500
4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase	2.7.1.148	At2g26930
2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	4.6.1.12	At1g63970
4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	1.17.4.3	At5g60600
4-hydroxy-3-methylbut-2-enyl diphosphate reductase	1.17.1.2	At4g34350
isopentenyl-diphosphate delta-isomerase	5.3.3.2	At3g02780
farnesyltransferase	2.5.1.29	At2g18620
geranyltransferase	2.5.1.10	At3g14530
dimethylallyltransferase	2.5.1.1	At3g14550 At3g20160 At3g29430 At3g32040 At4g36810 At4g17190 At5g47770

zymes related to the synthesis of terpene were manually discarded. For identification of the terpene synthases, we performed searches using keywords like terpene or terpenoid in a assembly of all the reads from sweet orange using CAP3 (Huang and Madan, 1999).

### Gene expression analysis

RNA was purified from leaves (pool of young leaves from a 6 months old plant grown in greenhouse), fruit (pericarp of a fruit of 6 cm in diameter from a plant grown in the field), flower (one fully developed flower isolated from a plant grown in the field), bark (pool of tissue from a 6 months old plant grown in the greenhouse) and root (same

plant used for bark preparation) of sweet orange plants (*Citrus sinensis* L. Osbeck var. Pêra IAC). The plant material was thoroughly ground in a mortar, the RNA was extracted using the TRIzol reagent and DNase-treated (DNase I Amp Grade, Invitrogen) according to the manufacturer's instructions. The cDNA was prepared using the M-MLV Reverse Transcriptase (Invitrogen). Briefly, 5 µg of RNA was mixed to 150 ng of random hexamers and 10 mM dNTP mix in a 12 µL reaction. After incubation for 5 min at 65 °C the RT buffer was added to 1X, DTT to 0.1 M and 20 U of RNase OUT (Invitrogen). The mixture was incubated for 2 min at 37 °C and 200 U of M-MLV reverse transcriptase were added to the reaction. The tube was further incubated at 25 °C for 10 min and 37 °C for 50 min. The enzyme was inactivated at 70 °C for 15 min. Eighty nanograms of the cDNA were used for the real time qPCR assay. One terpene synthase was selected for primer design using the Primer-Select software from the Lasergene99 suite (DNASTAR Inc., Madison, WI). Ten picomol of the primers (d-limonene-1F TGCATTAATCCCTCAACCTTGCTACCT, d-limonene-1R ATGATTCTGATGGCTGCTCTATTTG TTGC) were used in real time qPCR analysis run in an ABI 7000 Sequence Detection System (Applied Biosystems, Foster City, CA). The amplification was done using SYBR Green PCR Master Mix (Applied Biosystems). After incubation at 50 °C for 2 min and 95 °C for 10 min, we ran 40 amplification cycles of denaturation at 95 °C for 30 s, and annealing/amplification at 60 °C for 1 min. A dissociation analysis was performed to detect eventual abnormalities in the amplifications. As an internal control, a gene encoding β-tubulin (β-Tubulina-F - GGATCTTGAACCCGGTAC CA, β-Tubulina-R - ATCAATTCGGCGCCTTC AG) was used. The relative expression was calculated in the ABI Prism 7000 SDS software with RQ Study (Applied Biosystems).

## Results

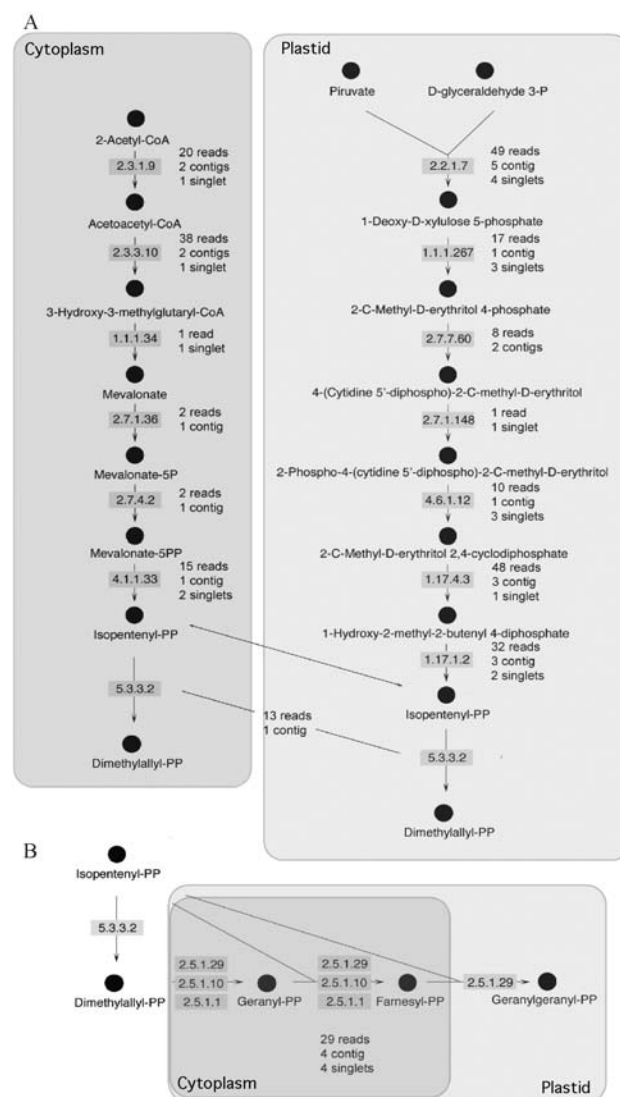
### *In silico* analysis

In plants, IPP is synthesized via two well-known pathways and isomerized to DMAPP by an enzyme known as IPP isomerase (E.C. 5.3.3.2). These two compounds are used for the synthesis of geranyl diphosphate, farnesyl diphosphate and geranylgeranyl diphosphate. In order to confirm the existence of the components of these pathways in sweet orange, the CitEST database was screened for sequences containing similarity to the different enzymes that participate in the synthesis of IPP. This search was done using the tblastn tool and an e-value of  $10^{-4}$  as a cut-off. As expected, in this analysis we identified all enzyme-encoding genes in the peel of sweet orange fruits as shown in Figure 1A. Three similar enzymes (farnesyltransferase, geranyltransferase, and dimethylallyltransferase - E.C.s 2.5.1.1, 2.5.1.10, and 2.5.1.29) process IPP and DMAPP to produce geranyl

diphosphate, farnesyl diphosphate and geranylgeranyl diphosphate. Genes encoding these enzymes were also found in the CitEST database (Figure 1B).

The actual number of reads varied for each gene. The highest number of reads identified for a particular enzyme was 49 for 1-deoxy-D-xylulose-5-phosphate synthase. These reads produced, after clusterization, 5 contigs and 4 singlets, which also represented the highest number of clusters observed for a particular enzyme.

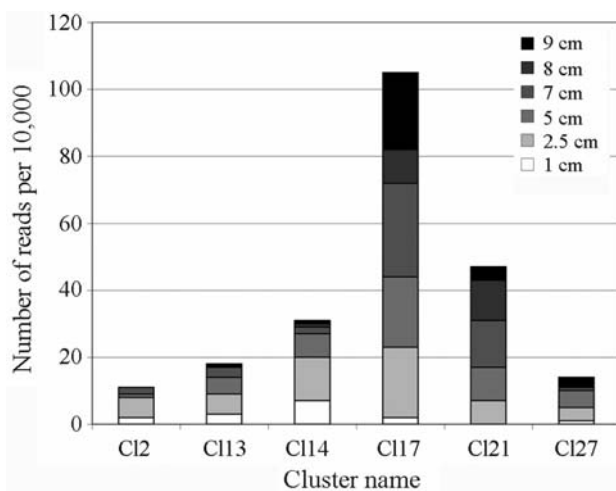
Production of terpene also depends on the activity of terpene synthases, enzymes that use geranyl diphosphate,



**Figure 1** - Pathways for the synthesis of terpenes. Scheme of the reactions that lead to the synthesis of terpenes in plants. The marked enzymes were found in the CitEST database and the number of reads for each enzyme is shown close to the EC number, as well as the number of contigs and singlets obtained after the assembly. A. The two distinct pathways represented in the specific compartments as they occur. B. IPP and DMAPP are converted in geranyl diphosphate, farnesyl diphosphate and geranylgeranyl diphosphate. The first two compounds are formed in both compartments, the cytoplasm and the plastid, while geranylgeranyl diphosphate is produced in the plastid.

farnesyl diphosphate and geranylgeranyl diphosphate as substrate to produce different terpenes. Monoterpenes are derived from geranyl diphosphate, sesquiterpenes from farnesyl diphosphate, and diterpenes from geranylgeranyl diphosphate by the action of terpene synthases or cyclases. In a global assembly of all the sweet orange reads, we were able to find 292 sequence tags obtained from the peel of fruits that formed 44 clusters with sequences related to terpene synthases. These 44 clusters encompassed a total of 344 reads (292 from fruit libraries and 52 from other sweet orange libraries) and were constituted by 27 contigs and 17 singlets. Eight of these 27 contigs (CI14, CI15, CI17, CI18, CI19, CI20, CI22, and CI25) were composed exclusively of reads from the fruit libraries. We compared the clusters that had reads from the fruit libraries with the Swissprot database using the blastp tool (WU-Blast2 search at EMBL). The obtained clusters were classified in different classes of enzymes: E.C.s 5.4.99.8, 4.2.3.8, 4.2.3.9, 2.5.1.-, 4.2.3.13 and E.C. 4.2.3.20. The last two were the most represented ones in the database, with 82 and 193 reads respectively (Table 2). The clusters that showed more than 10 reads were further analyzed. Since these reads came from libraries prepared from the peel of sweet orange fruits of different sizes, we evaluated the expression pattern in the different developmental stages of the fruit. Therefore, we counted the number of reads from each stage and normalized by 10,000 reads (Figure 2). The clusters 17 and 21 showed an expression more uniform throughout the development while the others showed a more concentrated expression in the three first developmental stages.

We also made a comparison of the number of reads from libraries prepared from different tissues. In this analysis, we compared the expression pattern of the terpene



**Figure 2** - Gene expression of terpene synthases in fruits. The most expressed terpene synthases found in fruits were evaluated for their expression pattern in fruit development. The number of reads for each transcript was counted for each of the libraries prepared, which represent different developmental stages. The reads were counted and normalized for 10,000 reads.

**Table 2** - Clusters of terpene synthase genes. The clusters containing reads from the fruit libraries were analyzed in relation to the number of reads and their composition as well as to the class of enzyme they encode. EC 4.2.3.20 is (R)-limonene synthase, EC 4.2.3.13 is (+)- $\delta$ -cadinene synthase, EC 5.4.99.8 is cycloartenol synthase, EC 2.5.1.32 is phytoene synthase, EC 4.2.3.9 is aristolochene synthase and EC 4.2.3.8 is casbene synthase.

Cluster name	Nr	N fruit libraries	N leaf libraries	N flower library	N bark library	Enzyme
CI2	12	11	1	0	0	EC 4.2.3.20
CI3	20	4	16	0	0	EC 4.2.3.13
CI4	7	3	4	0	0	EC 4.2.3.13
CI6	2	1	1	0	0	EC 4.2.3.13
CI7	3	1	2	0	0	EC 4.2.3.13
CI8	3	1	2	0	0	EC 5.4.99.8
CI9	2	1	1	0	0	EC 4.2.3.13
CI10	2	1	0	1	0	EC 4.2.3.20
CI11	3	2	0	1	0	EC 4.2.3.20
CI12	4	3	1	0	0	EC 5.4.99.8
CI13	19	18	1	0	0	EC 4.2.3.13
CI14	31	31	0	0	0	EC 4.2.3.20
CI15	2	2	0	0	0	EC 4.2.3.20
CI16	4	2	1	0	1	EC 4.2.3.13
CI17	106	106	0	0	0	EC 4.2.3.20
CI18	2	2	0	0	0	EC 4.2.3.20
CI19	4	4	0	0	0	EC 2.5.1.32
CI20	3	3	0	0	0	EC 4.2.3.20
CI21	57	47	9	1	0	EC 4.2.3.13
CI22	9	9	0	0	0	EC 4.2.3.20
CI23	3	2	1	0	0	EC 5.4.99.8
CI24	3	1	0	2	0	EC 5.4.99.8
CI25	2	2	0	0	0	EC 4.2.3.20
CI26	4	1	3	0	0	EC 2.5.1.32
CI27	15	14	0	1	0	EC 4.2.3.20
CI28	3	2	1	0	0	EC 4.2.3.13
CI34	2	1	1	0	0	EC 4.2.3.13
CI43	1	1	0	0	0	EC 4.2.3.20
CI44	1	1	0	0	0	EC 5.4.99.8
CI45	1	1	0	0	0	EC 4.2.3.20
CI46	1	1	0	0	0	EC 4.2.3.13
CI47	1	1	0	0	0	EC 4.2.3.20
CI48	1	1	0	0	0	EC 4.2.3.9
CI49	1	1	0	0	0	EC 4.2.3.20
CI50	1	1	0	0	0	EC 4.2.3.20
CI51	1	1	0	0	0	EC 4.2.3.13
CI52	1	1	0	0	0	EC 5.4.99.8
CI53	1	1	0	0	0	EC 4.2.3.20
CI54	1	1	0	0	0	EC 4.2.3.20
CI55	1	1	0	0	0	EC 4.2.3.20
CI56	1	1	0	0	0	EC 4.2.3.20
CI57	1	1	0	0	0	EC 4.2.3.8
CI58	1	1	0	0	0	EC 4.2.3.20
CI59	1	1	0	0	0	EC 5.4.99.8
Total	344	292	45	6	1	

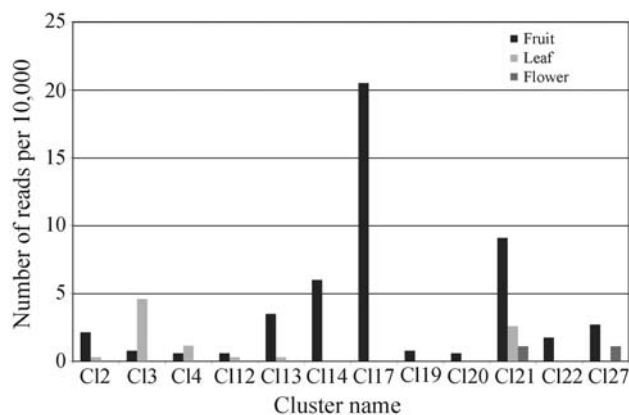
Nr: total number of reads; N fruit libraries: number of reads in the fruit libraries; N leaf libraries: number of reads in the leaf libraries; N flower library: number of reads in the flower library; N bark library: number of reads in the bark library.

synthases that were expressed in the fruit with other libraries. The reads were counted, and the expression was normalized by 10,000 reads (Figure 3). Besides the fruits, these genes were expressed in leaves, flowers and barks. For this analysis, it is clear that most of these clusters showed higher expression in fruits but others like C13 and C14 had higher expression in leaves.

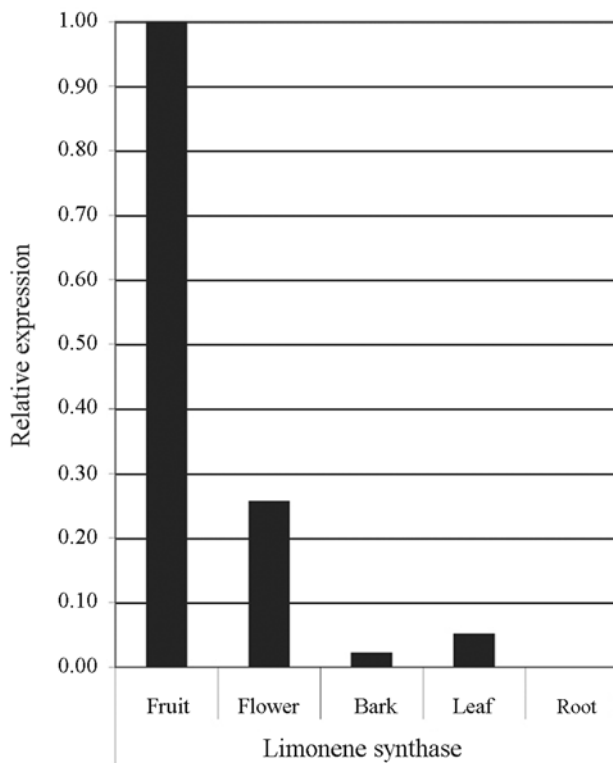
### Gene expression analysis

Limonene is the major component of citrus essential oil. One cluster in particular, C117, seems to be the most important for limonene synthesis in fruits because it presented the highest number of reads besides having a high expression throughout development. To understand more about the enzyme responsible for the synthesis of this terpene, we evaluated the expression pattern of this putative limonene synthase. Moreover, all the data evaluated was obtained *in silico*. To confirm the data, it was necessary to take a biological approach. Therefore, we decided to validate the observations using the quantitative real time PCR method. Gene-specific primers were used for the amplification of a particular region of the transcript of the most expressed limonene synthase. The specificity was certified by a blastn search against the CitEST database to be unique for the studied genes. The expression pattern in fruit, leaf, flower, bark, and root (additional to the *in silico* studied tissues) of sweet orange cv. Pêra was analyzed in triplicate (Figure 4).

Using fruit as a calibrator, we were able to compare the expression pattern in the different tissues for the gene encoding this particular limonene synthase. In this analysis, we were able to confirm a higher expression of this gene in fruits as observed in the *in silico* analysis. It presents an expression around 4 times higher than flower, the second most abundant tissue. Detection in root was obtained but at a very high Ct, demonstrating the low amount of transcripts present in this tissue.



**Figure 3** - Gene expression of terpene synthases in sweet orange. The terpene synthases expressed in fruits were analyzed *in silico* for their expression in other tissues. The reads were counted and normalized to 10,000 reads.



**Figure 4** - Expression of the limonene synthase in sweet orange. The tissue expression pattern of a putative limonene synthase was evaluated by real-time qPCR. The expression in each tissue was compared to the expression in fruit of sweet orange.

### Discussion

Citrus production in Brazil is a major economic activity. It accounts for a gross income of US\$ 1.5 billion mainly because of the production of concentrated juice for exportation. As a byproduct of this activity, the industry also produces essential oil obtained from the peel of the fruit. This oil is the main component of the essential oil exports for Brazil. Together with oils from other citrus, it accounts for 85% of the exports. Terpenes are the main component of the citrus essential oil and it is composed mainly of mono (around 90%) and sesquiterpenes that accumulate especially in oil glands present in the epicarp. The terpenes that constitute the essential oil of sweet orange are  $\alpha$ -pinene, camphene, sabinene,  $\beta$ -pinene, myrcene,  $\delta$ -3-carene, d-limonene, terpinolene,  $\alpha$ -copaene,  $\beta$ -elemene,  $\beta$ -caryophyllene,  $\alpha$ -cadinene,  $\alpha$ -humulene,  $\beta$ -farnesene, valencene, and  $\delta$ -cadinene. The major terpene by far is d-limonene, which accounts for almost 95% of the volatile fraction of the oil.

In this work, we evaluated the genes that encode the components involved in the synthesis of terpenes found in the fruit of sweet orange.

Synthesis of terpenes relies on two different pathways that lead to the synthesis of IPP. These two pathways occur in different compartments, the cytoplasm and the plastid.

Both of the pathways have enzymes encoded in the nucleus and their transcripts were present in the libraries constructed for the sweet orange flavedo, as expected. The expression level of the enzymes was very variable with genes showing up to 49 transcripts (the 1-deoxy-D-xylulose-5-phosphate synthase) and genes showing just one, the hydroxymethylglutaryl-CoA (HMG) reductase. This enzyme is important in the first stages of fruit development (Narita and Gruissem, 1989; Cowan *et al.*, 1997; Kato-Emori *et al.*, 2001; Kobayashi *et al.*, 2002) when it correlates with cell division. In later stages of fruit development, this enzyme does not seem to be active (Narita and Gruissem, 1989; Cowan *et al.*, 1997; Kato-Emori *et al.*, 2001). Also noticeable, 1-deoxy-D-xylulose-5-phosphate synthase (DXPS) was detected as the most expressed gene within the mevalonate-independent pathway. That could correlate with the role of DXPS on catalyzing the first committed step of isoprenoid (non-mevalonate derived) synthesis within plastids, a pathway responsible for supplying all the required IPP into the compartment and its partial exportation to the cytosol (Lichtenthaler, 2000).

Several terpene synthase encoding genes are present in the citrus genome. Transcripts for terpene synthases were detected in fruit, leaf, flower, and bark tissues, which probably reflects the general role of terpenes as pollinator attracters or repellents against herbivores or pathogens (Langenheim, 1994; Bouwmeester *et al.*, 1999; Dicke, 1999; Pichersky and Gershenzon, 2002; Dudareva *et al.*, 2003). We focused our analysis on the terpene synthases that were identified in the epicarp of sweet orange fruits since citrus essential oil is extracted from the flavedo where they accumulate in specialized secretion cavities (Lucker *et al.*, 2002). In this analysis, we found 44 clusters containing reads from the fruit libraries. Most of the 27 contigs that formed these 44 clusters also had reads from other libraries, although in a lesser amount. These clusters encoded in their majority enzymes of E.C. 4.2.3.13 ( $\delta$ -cadinene synthase) or E.C. 4.2.3.20 (limonene synthase). Most of the reads belonged to clusters that showed similarity to limonene synthase. Limonene is the major component of citrus essential oil and therefore it is not a surprise that genes encoding these enzymes showed high expression in fruits of sweet orange. Interestingly, one of the clusters, Cl17 is the one that show the highest number of sequences and *in silico* analysis of expression show that this limonene synthase is expressed quite equally throughout the development of the fruit in high levels. This could be the form responsible for most of the limonene synthesized in the fruits of sweet orange. This is also true for Cl21, which encodes an enzyme of E.C. 4.2.3.13. The expression pattern of the other clusters throughout fruit development showed that the terpene synthases are predominantly expressed at the beginning of the fruit development. It remains to be seen if this expression correlates with production of essential oil and terpene more specifically.

Our qPCR analysis showed higher expression of the limonene synthase encoding gene in fruits when compared to flower, leaf, bark, and root tissues. This result was expected, since it agreed with our *in silico* results. The qPCR results corroborate our EST data, reinforcing the usefulness of sequencing cDNA stretches on inferring gene expression patterns when non-normalized libraries are considered (Ewing *et al.*, 1999). The limonene synthase gene showed expression in all analyzed tissues, fruit and flower with the highest levels. This probably reflects the role of terpenes as an important group of secondary metabolites involved in fruit and flower aroma (Sharon-Asa *et al.*, 2003). In agreement with *in silico* data, expression of the limonene synthase gene was higher in fruit tissue. The second highest level of gene expression for limonene synthase was detected in flower, but interestingly, not a single sequence tag was detected within the respective cluster in that tissue in a previous *in silico* analysis (Figure 3). This probably can be explained by the reduced number of sequenced flower-derived reads in the CitEST Project in comparison to fruit (there are 5 times more reads for fruits). Moreover, the high sensitivity of the real time quantitative PCR analysis could also reveal marked differences that are not observed in less sensitive methods (Lossos *et al.*, 2004). Another difference is in relation to the source for the experiments. For the qPCR study of flowers, a unique flower was used for RNA extraction while, for the library construction, a pool of flowers in different developmental stages was utilized. In *Citrus limon*, two different limonene synthases were identified (Lucker *et al.*, 2002). In the CitEST database, there are many more sequences encoding putative limonene synthases. One possibility is that the clustering was too stringent and some of the sequences should have been clustered together. Nevertheless, the high level of expression of one of these genes can help to explain the reason why limonene is more than 90% of the terpenes in the composition of citrus essential oil (Sawamura *et al.*, 2004; Song *et al.*, 2006).

Further gene characterization, validation of their expression and enzyme activity, and elucidation of specific roles for terpene synthases in terpene production are still in demand and remain a challenge in generating new information and supporting their application for development of novel technologies in plant breeding, drug discovery, and *in vitro* synthesis.

## Acknowledgements

The authors would like to thank Silvia O. Dorta for technical assistance in sequencing the libraries, Marcelo Reis for bioinformatic assistance, and Marcos A. Machado for coordinating the CitEST Project. CNPq/ Millennium Institute (62.0054/01-8) and FAPESP (2006/06995-0) financially supported this work. MAT is a CNPq fellow.

## References

- Adam KP, Thiel R and Zapp J (1999) Incorporation of 1-[1-(13)C]Deoxy-D-xylulose in chamomile sesquiterpenes. *Arch Biochem Biophys* 369:127-132.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997) Gapped BLAST and psi-blast: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402.
- Bouwmeester HJ, Verstappen FW, Posthumus MA and Dicke M (1999) Spider mite-induced (3S)-(E)-nerolidol synthase activity in cucumber and lima bean. The first dedicated step in acyclic C11-homoterpene biosynthesis. *Plant Physiol* 121:173-180.
- Cowan AK, Moore-Gordon CS, Bertling I and Wolstenholme BN (1997) Metabolic control of avocado fruit growth (isoprenoid growth regulators and the reaction catalyzed by 3-hydroxy-3-methylglutaryl coenzyme A reductase). *Plant Physiol* 114:511-518.
- Croteau R, Kutchan TM and Lewis NG (2000) Natural products (secondary metabolites). In: Buchanan BB, Gruissem W and Jones RL (eds) *Biochemistry & Molecular Biology of Plants*. American Society of Plant Physiologists, Rockville, pp 1250-1318.
- Dicke M (1999) Specificity of herbivore-induced plant defences. *Novartis Foundation Symp* 223:43-54.
- Dudareva N, Martin D, Kish CM, Kolosova N, Gorenstein N, Faldt J, Miller B and Bohlmann J (2003) (E)-beta-ocimene and myrcene synthase genes of floral scent biosynthesis in snapdragon: Function and expression of three terpene synthase genes of a new terpene synthase subfamily. *Plant Cell* 15:1227-1241.
- Ewing RM, Ben Kahla A, Poirot O, Lopez F, Audic S and Claverie JM (1999) Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Res* 9:950-959.
- Huang X and Madan A (1999) CAP3: A DNA sequence assembly program. *Genome Res* 9:868-877.
- Kato-Emori S, Higashi K, Hosoya K, Kobayashi T and Ezura H (2001) Cloning and characterization of the gene encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase in melon (*Cucumis melo* l. *Reticulatus*). *Mol Genet Genomics* 265:135-142.
- Kobayashi T, Kato-Emori S, Tomita K and Ezura H (2002) Detection of 3-hydroxy-3-methylglutaryl-coenzyme A reductase protein cm-hmgr during fruit development in melon (*Cucumis melo* l.). *Theor Appl Genet* 104:779-785.
- Langenheim JH (1994) Higher plant terpenoids: A phyto-centric overview of their ecological roles. *J Chem Ecol* 20:1223-1280.
- Lichtenthaler HK (1999) The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annu Rev Plant Physiol Plant Mol Biol* 50:47-65.
- Lichtenthaler HK (2000) Non-mevalonate isoprenoid biosynthesis: Enzymes, genes and inhibitors. *Biochem Soc Trans* 28:785-789.
- Lossos IS, Czerwinski DK, Alizadeh AA, Wechsler MA, Tibshirani R, Botstein D and Levy R (2004) Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *N Engl J Med* 350:1828-1837.
- Lucker J, El Tamer MK, Schwab W, Verstappen FW, van der Plas LH, Bouwmeester HJ and Verhoeven HA (2002) Monoterpene biosynthesis in lemon (*Citrus limon*). cDNA isolation and functional analysis of four monoterpene synthases. *Eur J Biochem* 269:3160-3171.
- Marra MA, Kucaba TA, Hillier LW and Waterston RH (1999) High-throughput plasmid DNA purification for 3 cents per sample. *Nucleic Acids Res* 27:e37.
- Narita JO and Gruissem W (1989) Tomato hydroxymethylglutaryl-coA reductase is required early in fruit development but not during ripening. *Plant Cell* 1:181-190.
- Pichersky E and Gershenzon J (2002) The formation and function of plant volatiles: Perfumes for pollinator attraction and defense. *Curr Opin Plant Biol* 5:237-243.
- Sawamura M, Thi Minh Tu N, Onishi Y, Ogawa E and Choi HS (2004) Characteristic odor components of *Citrus reticulata* Blanco (ponkan) cold-pressed oil. *Biosci Biotechnol Biochem* 68:1690-1697.
- Sharon-Asa L, Shalit M, Frydman A, Bar E, Holland D, Or E, Lavi U, Lewinsohn E and Eyal Y (2003) Citrus fruit flavor and aroma biosynthesis: Isolation, functional characterization, and developmental regulation of Cstps1, a key gene in the production of the sesquiterpene aroma compound valencene. *Plant J* 36:664-74.
- Song HS, Lan Phi NT, Park YH and Sawamura M (2006) Volatile profiles in cold-pressed peel oil from Korean and Japanese shiranui (*Citrus unshiu* Marcov. X *C. sinensis* Osbeck x *C. reticulata* Blanco). *Biosci Biotechnol Biochem* 70:737-739.

## Internet Resources

- CitEST Database, <http://citest.centrodecitricultura.br/> (March 25, 2006).
- Genbank, [www.ncbi.nlm.nih.gov/Genbank/](http://www.ncbi.nlm.nih.gov/Genbank/) (March 25, 2006).
- BLAST EMBL, <http://dove.embl-heidelberg.de/Blast2/> (February 25, 2007).

Associate Editor: Reinaldo Montrazi Barata