JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

J. Agric. Food Chem. **2007,** *55,* 7405−7417 **7405**

Carotenoid Biosynthetic Pathway in the *Citrus* Genus: Number of Copies and Phylogenetic Diversity of Seven Genes

ANNE-LAURE FANCIULLINO,[†,⊥] CLAUDIE DHUIQUE-MAYER,[‡,⊥] FRANÇOIS LURO,[§,⊥] RAPHAËL MORILLON,[†,⊥] AND PATRICK OLLITRAULT*[,†]

Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), UPR amélioration génétique des espèces à multiplication végétative, Avenue Agropolis - TA A-75 / 02, 34398 Montpellier cedex 5, France, CIRAD, UMR QUALISUD, Avenue J. F. Breton - TA B-95 / 16, 34398 Montpellier Cedex 5, France, and Institut National de la Recherche Agronomique (INRA), UR GEQA Site de San Giuliano, 20230 San Giuliano, France

The first objective of this paper was to analyze the potential role of allelic variability of carotenoid biosynthetic genes in the interspecific diversity in carotenoid composition of *Citrus* juices. The second objective was to determine the number of copies for each of these genes. Seven carotenoid biosynthetic genes were analyzed using restriction fragment length polymorphism (RFLP) and simple sequence repeats (SSR) markers. RFLP analyses were performed with the genomic DNA obtained from 25 *Citrus* genotypes using several restriction enzymes. cDNA fragments of *Psy*, *Pds*, *Zds*, *Lcy-b*, *Lcy-e*, *Hy-b*, and *Zep* genes labeled with [α-^{32}P]dCTP were used as probes. For SSR analyses, two primer pairs amplifying two SSR sequences identified from expressed sequence tags (ESTs) of *Lcy-b* and *Hy-b* genes were designed. The number of copies of the seven genes ranged from one for *Lcy-b* to three for *Zds*. The genetic diversity revealed by RFLP and SSR profiles was in agreement with the genetic diversity obtained from neutral molecular markers. Genetic interpretation of RFLP and SSR profiles of four genes (*Psy1*, *Pds1*, *Lcy-b*, and *Lcy-e1*) enabled us to make inferences on the phylogenetic origin of alleles for the major commercial citrus species. Moreover, the results of our analyses suggest that the allelic diversity observed at the locus of both of lycopene cyclase genes, *Lcy-b* and *Lcy-e1*, is associated with interspecific diversity in carotenoid accumulation in *Citrus*. The interspecific differences in carotenoid contents previously reported to be associated with other key steps catalyzed by PSY, HY-b, and ZEP were not linked to specific alleles at the corresponding loci.

KEYWORDS: ***Citrus*; carotenoids; biosynthetic genes; allelic variability; phylogeny**

INTRODUCTION

Carotenoids are pigments common to all photosynthetic organisms. In pigment−protein complexes, they act as light sensors for photosynthesis but also prevent photo-oxidation induced by too strong light intensities. In horticultural crops, they play a major role in fruit, root, or tuber coloration and in nutritional quality. Indeed some of these micronutrients are precursors of vitamin A, an essential component of human and animal diets (*1*). Carotenoids may also play a role in chronic disease prevention (such as certain cancers), probably due to their antioxidant properties (*2*).

The carotenoid biosynthetic pathway is now well established (*3*, *4*; **Figure 1**). Carotenoids are synthesized in plastids by nuclear-encoded enzymes (*4*). The immediate precursor of carotenoids (and also of gibberellins, plastoquinone, chlorophylls, phylloquinones, and tocopherols) is geranylgeranyl diphosphate (GGPP). In light-grown plants, GGPP is mainly derived from the methylerythritol phosphate (MEP) pathway (*5*; **Figure 1**). The condensation of two molecules of GGPP catalyzed by phytoene synthase (PSY) leads to the first colorless carotenoid, 15-*cis*-phytoene. Phytoene undergoes four desaturation reactions catalyzed by two enzymes, phytoene desaturase (PDS) and ζ-carotene desaturase (ZDS), which convert phytoene into the red-colored poly-*cis*-lycopene. Recently, Isaacson et al. (*6*) and Park et al. (*7*) isolated from tomato and *Arabidopsis thaliana*, respectively, the genes that encode the carotenoid isomerase (CRTISO) which, in turn, catalyzes the isomerization of poly-*cis*-carotenoids into all-*trans*-carotenoids. CRTISO acts on prolycopene to form all-*trans* lycopene, which undergoes cyclization reactions. Cyclization of lycopene is a branching

* Author to whom correspondence should be addressed. Tel: +33(0)467615971. Fax: +33(0)467615666. E-mail: patrick.ollitrault@cirad.fr.
[†] Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), UPR amélioration génétique des espèces à multiplication végétative.
[⊥] E-mail addresses: fanciullino@corse.inra.fr, claudie.dhuique-mayer@cirad.fr, luro@corse.inra.fr, raphael.morillon@cirad.fr.
[§] Institut National de la Recherche Agronomique (INRA).
[‡] CIRAD, UMR QUALISUD.

10.1021/jf070711h CCC: $37.00
Published on Web 08/11/2007

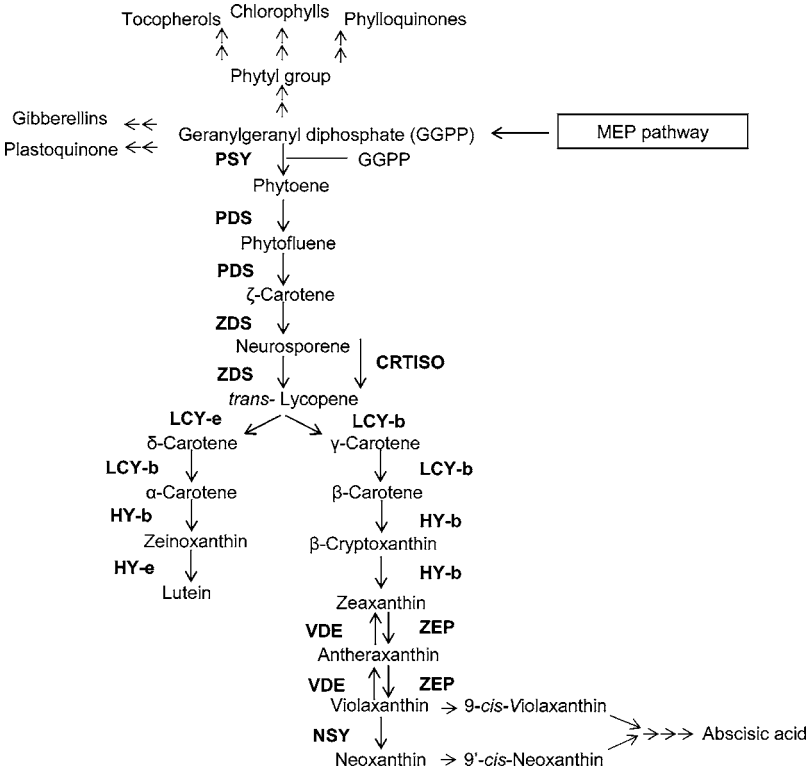

Figure 1. Carotenoid biosynthetic pathway in plants: MEP pathway, methylerythritol phosphate pathway; PSY, phytoene synthase; PDS, phytoene desaturase; ZDS, ζ-carotene desaturase; CRTISO, carotenoid isomerase; LCY-e, lycopene ϵ-cyclase; LCY-b, lycopene β-cyclase; HY-b, β-carotene hydroxylase; HY-e, ϵ-carotene hydroxylase; ZEP, zeaxanthin epoxidase; VDE, violaxanthin de-epoxidase; NSY, neoxanthin synthase.

point: one branch leads to β-carotene (β,β-carotene) and the other to α-carotene (β,ϵ-carotene). Lycopene β-cyclase (LCY-b) then converts lycopene into β-carotene in two steps, whereas the formation of α-carotene requires the action of two enzymes, lycopene ϵ-cyclase (LCY-e) and lycopene β-cyclase (LCY-b). α-Carotene is converted into lutein by hydroxylations catalyzed by ϵ-carotene hydroxylase (HY-e) and β-carotene hydroxylase (HY-b). Other xanthophylls are produced from β-carotene with hydroxylation reactions catalyzed by HY-b and epoxydation catalyzed by zeaxanthin epoxidase (ZEP). Most of the carotenoid biosynthetic genes have been cloned and sequenced in *Citrus* varieties (*8*−*12*). However, our knowledge of the complex regulation of carotenoid biosynthesis in *Citrus* fruit is still limited. We need further information on the number of copies of these genes and on their allelic diversity in *Citrus* because these can influence carotenoid composition within the *Citrus* genus.

Citrus fruit are among the richest sources of carotenoids. The fruit generally display a complex carotenoid structure, and 115 different carotenoids have been identified in *Citrus* fruit (*9*−*11*, *13*−*18*). The carotenoid richness of *Citrus* flesh depends on environmental conditions, particularly on growing conditions, and on geographical origin (*13*). However the main factor influencing variability of carotenoid quality in juice has been shown to be genetic diversity. Kato et al. (*15*) showed that mandarin and orange juices accumulated high levels of β-cryptoxanthin and violaxanthin, respectively, whereas mature lemon accumulated extremely low levels of carotenoids. Goodner et al. (*16*) demonstrated that mandarins, oranges, and their hybrids could be clearly distinguished by their β-cryptoxanthin contents. Juices of red grapefruit contained two major carotenoids: lycopene and β-carotene (*17*). More recently, we conducted a broad study on the organization of the variability of carotenoid contents in different cultivated *Citrus* species in relation with the biosynthetic pathway (*18*). Qualitative analysis of presence or absence of the different compounds revealed three main clusters: (1) mandarins, sweet oranges, and sour oranges; (2) citrons, lemons, and limes; (3) pummelos and grapefruit. Our study also enabled identification of key steps in the diversification of the carotenoid profile. Synthesis of phytoene appeared as a limiting step for acid *Citrus* (cluster 2), while formation of β-carotene and α-carotene from lycopene were dramatically limited in cluster 3 (pummelos and grapefruit). Only varieties in cluster 1 were able to produce violaxanthin. In the same study (*18*), we concluded that there was a very strong correlation between the classification of *Citrus* species based on the presence or absence of carotenoids (below, this classification is also referred to as the organization of carotenoid diversity) and genetic diversity evaluated with biochemical or molecular markers such as isozymes (*19*) or randomly amplified polymorphic DNA (RAPD) (*20*). We also concluded that, at the interspecific level, the organization of the diversity of carotenoid composition was linked to the global evolution process of cultivated *Citrus* rather than to more recent mutation events or human selection processes. Indeed, at interspecific level, a correlation between phenotypic variability and genetic diversity is common and is generally associated with generalized gametic disequilibrium resulting from the history of cultivated *Citrus* (*19*). Thus from numerical taxonomy based on morphological traits or from analysis of molecular markers (*19*−*21*), all authors agreed on the existence of three basic taxa (*C. reticulata*, mandarins; *C. medica*, citrons; and *C. maxima*, pummelos) whose differentiation was the result of allopatric evolution. All other cultivated *Citrus* species (*C. sinensis*, sweet oranges; *C. aurantium*, sour oranges; *C. paradisi*, grapefruit; and *C. limon*, lemons) resulted from hybridization events within this basic pool except for *C. aurantifolia*, which may be a hybrid between *C. medica* and *C. micrantha* (*20*). Below, all the above-

Table 1. Genotypes Used for RFLP and SSR Analyses

no.	common name	species	ICVN number[a]
1	Willowleaf mandarin	*C. reticulata* Blanco	ICVN 0100133
2	Wase Satsuma	*C. reticulata* Blanco	ICVN 0100230
3	Hansen mandarin	*C. reticulata* Blanco	ICVN 0100357
4	Seedless pummelo	*C. maxima* (Burm.) Merr.	ICVN 0100710
5	Deep Red pummelo	*C. maxima* (Burm.) Merr.	ICVN 0100757
6	Chandler pummelo	*C. maxima* (Burm.) Merr.	ICVN 0100608
7	Etrog citron	*C. medica* L.	ICVN 0100130
8	Diamante citron	*C. medica* L.	ICVN 0100540
9	Marsh grapefruit	*C. paradisi* Macf.	ICVN 0100188
10	Star Ruby grapefruit	*C. paradisi* Macf.	ICVN 0100293
11	Ray Ruby grapefruit	*C. paradisi* Macf.	ICVN 0100604
12	Shamouti orange	*C. sinensis* (L.) Osb.	ICVN 0100299
13	Sanguinelli orange	*C. sinensis* (L.) Osb.	ICVN 0100243
14	Cara Cara navel orange	*C. sinensis* (L.) Osb.	ICVN 0100666
15	Morocco sour orange	*C. aurantium* L.	ICVN 0110033
16	Bouquetier de Nice	*C. aurantium* L.	ICVN 0100688
17	Myrtle-leaf orange	*C. aurantium* L.	ICVN 0100708
18	Eureka Frost lemon	*C. limon* (L.) Burm. f.	ICVN 0100004
19	Volkamer lemon	*C. limonia* Osbeck	ICVN 0100729
20	Meyer lemon	*C. meyeri* Yu. Tan.	ICVN 0100549
21	Rangpur lime	*C. limonia* Osbeck	ICVN 0110050
22	Mexican lime	*C. aurantifolia* (Christm.) Swing.	ICVN 0100140
23	Palestine sweet lime	*C. limettioides* Tan.	ICVN 0100802
24	Clementine	*C. clementina* hort. ex Tan.	ICVN 0100092
25	Huang pi Chen orange	*C. sinensis* (L.) Osb.	ICVN 0100567

[a] International Citrus Variety Numbering.

mentioned species are referred to as secondary species. Moreover it appears that only mutational (or epigenetic) events were involved in the diversification of secondary species such as *C. sinensis*, *C. aurantium*, or *C. paradisi* (*19*).

Our previous results (*18*) and data on *Citrus* evolution lead us to propose the hypothesis that the allelic variability supporting the organization of carotenoid diversity at interspecific level preceded events that resulted in the creation of secondary species. Such molecular variability may have two different effects: on the one hand, non-silent substitutions in coding region affect the specific activity of corresponding enzymes of the biosynthetic pathway, and on the other hand, variations in untranslated regions affect transcriptional or post-transcriptional mechanisms.

There is no available data on the allelic diversity of *Citrus* genes of the carotenoid biosynthetic pathway. The objective of this paper was to test the hypothesis that allelic variability of these genes partially determines phenotypic variability at the interspecific level. For this purpose, we analyzed the RFLPs around seven genes of the biosynthetic pathway of carotenoids (*Psy*, *Pds*, *Zds*, *Lcy-b*, *Lcy-e*, *Hy-b*, *Zep*) and the polymorphism of two SSR sequences found in *Lcy-b* and *Hy-b* genes in a representative set of varieties of the *Citrus* genus already analyzed for carotenoid constitution. Our study aimed to answer the following questions: (a) are those genes mono- or multilocus, (b) is the polymorphism revealed by RFLP and SSR markers in agreement with the general history of cultivated *Citrus* thus permitting inferences about the phylogenetic origin of genes of the secondary species, and (c) is this polymorphism associated with phenotypic (carotenoid compound) variations?

MATERIALS AND METHODS

Plant Materials. Leaves of 25 *Citrus* genotypes (**Table 1**) and fruits of Satsuma Wase mandarin were harvested from trees grown at the *Citrus* Gerplasm Bank (Station de Recherche Agronomique, San Giuliano, Corsica). Total DNA extraction was performed according to Doyle and Doyle (*22*) from 0.5 g of leaf tissue. The DNA was quantified by measuring the increased fluorescence of the Hoechst dye 33258. This plant material used for RFLP analysis was the same as that used for SSR analysis.

Total RNA was isolated from the pulp of Satsuma mandarin fruits as described by Manning (*23*). UV absorption spectrophotometry and gel electrophoresis were performed to test RNA quality as described by Sambrook et al. (*24*). First-strand cDNAs were synthesized from 1 μg of total RNA using First Strand cDNA synthesis kit (Fermentas). Satsuma mandarin cDNA pool was used as template for probe amplifications.

RFLP Analysis. Total DNA (10 μg) was digested with *Eco*RV, *Bam*HI, and *Hin*dIII, electrophoresed on 1% agarose gels, and blotted onto nylon membranes (Hybond-N, Amersham, UK) according to the manufacturer's recommendations.

To prepare probes, cDNA fragments of carotenoid biosynthetic genes *Psy*, *Pds*, *Zds*, *Lcy-b*, *Hy-b*, *Zep*, and *Lcy-e* were amplified by PCR using cDNA from Satsuma mandarin and specific primers. Primer pairs were designed from *Citrus Psy*, *Pds*, *Zds*, *Lcy-b*, *Hy-b*, *Zep*, and *Lcy-e* full coding sequences deposited in the database (GenBank accession numbers AF220218, AB046992, AB072343, AY166796, AF315289, AB075547, and AY533827) (for primer sequences and cDNA fragment length, see **Table 2**). The amplified fragments were separated by electrophoresis on 1% agarose gel and purified with GFX PCR DNA and gel band purification kit (Amersham, UK). The identity of all cDNA fragments was confirmed by sequencing (MWG, Martinsried, Germany). The cDNA fragments were labeled with [α-^{32}P]dCTP (Megaprime DNA Labeling System, Amersham, UK). In order to screen the presence of intronic sequences and restriction sites in genomic DNA corresponding to RFLP probes, amplifications were performed with genomic DNA using primer pairs described in **Table 2**. Aliquots of PCR products were purified with GFX PCR DNA and gel band purification kit (Amersham, UK) and digested with *Eco*RV, *Bam*HI, and *Hin*dIII. Digestions were carried out overnight at 37 °C in a total volume of 200 μL with 40 units of restriction enzyme (Invitrogen), the 1× corresponding reaction buffer, and 10 μg of DNA. PCR and digestion products were separated by electrophoresis on 1% agarose gel.

Membranes were prehybridized at 65 °C for 4−5 h in 50 mM Tris HCl (pH 8), 10 mM EDTA (pH 8), 5× standard saline citrate (SSC), 0.2% sodium dodecyl sulfate (SDS), 1× Denhardt's solution, and 100 μg mL^{-1} denatured salmon sperm DNA. Hybridization was carried out at 65 °C overnight in 50 mM Tris HCl (pH 8), 10 mM EDTA (pH 8), 5× SSC, 0.2% SDS, 1× Denhardt's solution, 100 μg mL^{-1} denatured

Table 2. Primers Used for cDNA Amplification and RFLP Analysis

gene	primer	amplicon size (bp)
*Psy*_1 forward	5′-GGGTTGTATCACCTAAC-3′	397
*Psy*_1 reverse	5′-CTTAAAGTTCCGGGT-3′	
*Pds*_1 forward	5′-CAACTTGCGATATGGTT-3′	588
*Pds*_1 reverse	5′-TTCACTTTCTCCGGC-3′	
*Zds*_1 forward	5′-TTCTTCAGTTCTGTTTCCT-3′	557
*Zds*_1 reverse	5′-GAGCATTTCTTGCTTTATC-3′	
*Lcy-b*_1 forward	5′-TCTTGCCCCAAGTTC-3′	498
*Lcy-b*_1 reverse	5′-TTTTCGACTTCAGCAAC-3′	
*Hy-b*_1 forward	5′-ATGCACGAGTCTCACC-3′	382
*Hy-b*_1 reverse	5′-TTCTCTTACTGATCTCCTTCT-3′	
*Zep*_1 forward	5′-CAGTTAATCTTTCAACAGC-3′	635
*Zep*_1 reverse	5′-CTTATCTCCATGATCCTTA-3′	
*Lcy-e*_1 forward	5′-GCTCCAAACTATGCTTC-3′	343
*Lcy-e*_1 reverse	5′-GTCTGATAAGGCACTTTCT-3′	

salmon sperm DNA, and 10% dextran sulfate. Membranes were washed once at room temperature in 0.5× SSC and 0.1% SDS, once at 65 °C for 30 min in 0.5× SSC and 0.1% SDS, and once at 65 °C for 30 min in 0.25× SSC and 0.1% SDS. Membranes were exposed to X-ray film (Biomax MS Film, Kodak).

SSR Analysis. SSR sequences were screened from a *Citrus* expressed sequence tag (EST) library, which contained carotenoid biosynthetic genes (*25*). Two EST SSR sequences were found: one corresponded to the *Lcy-b* gene and another one to the *Hy-b* gene. Two primer pairs, flanking the SSR motifs, were designed, 1210 and 1388 (see **Table 3**) (*25*). These two primer pairs were used for SSR analysis and nuclear SSR amplifications. PCR amplifications of the samples were performed using a PTC-200 thermocycler (MJ Research Inc.) in 15 μL final volume containing 0.8 U of *Taq* DNA polymerase (Eurogentec) and corresponding 1× reaction buffer, 10 ng of *Citrus* genomic DNA, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 0.2 μM forward primer, and 0.2 μM reverse primer. The following PCR program was applied: denaturation at 94 °C for 5 min; 35 repeats of the cycle 30 s at 94 °C, 1 min at 55 or 60 °C, 45 s at 72 °C; and a final elongation step of 4 min at 72 °C. Samples were then kept at 4 °C prior to analysis. After adding 15 μL of loading buffer [98% formamide, 10 mM EDTA, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol], the mixture was denatured at 92 °C for 3 min and kept at 70 °C during gel loading. Six microliters of each sample was loaded in 5% denaturing polyacrylamide (19:1 acrylamide/bisacrylamide) gels with 7.5 M urea in 0.5% TBE buffer prior to electrophoresis at 60 W for 2 h. Gels were silver stained using an improved method adapted from Beidler et al. (*26*).

Data Analysis. The data matrix was composed of 65 fragments as variables and 25 genotypes. Fragments were scored as 1 for the presence or 0 for the absence. Three representations were obtained from the presence or absence of the variables and matrices of Dice dissimilarities using @DARwin 4.0 software (CIRAD Montpellier, France). Dice's dissimilarities were calculated as follows:

$$d(i,j) = (b + c)/(2a + b + c)$$

where $d(i,j)$ = dissimilarity between genotypes i and j, a = number of variables where x_i = presence and x_j = presence; b = number of variables where x_i = presence and x_j = absence; c = number of variables where x_i = absence and x_j = presence.

The first tree was based on the presence or absence of the 58 fragments isolated from the RFLP profiles of 7 genes, whereas the second tree was based on the presence or absence of the 9 fragments isolated from the RFLP and SSR profiles of the *Hy-b* gene, and the third one was based on the presence or absence of the 7 fragments isolated from the RFLP profiles of the *Zep* gene. The trees were constructed with the weighted neighbor-joining method.

RESULTS AND DISCUSSION

Global Diversity of the Genotype Sample Observed by RFLP Analysis. RFLP analyses were performed using probes defined from expressed sequences of seven major genes of the carotenoid biosynthetic pathway (**Figure 1**). One or two restriction enzymes were used for each gene. None of these enzymes cut the cDNA probe sequence except *Hin*dIII for the *Lcy-e* gene. Intronic sequences and restriction sites on genomic sequences were screened with PCR amplification using genomic DNA as template and with digestion of PCR products. The results indicated the absence of an intronic sequence for *Psy* and *Lcy-b* fragments. The absence of intron in these two fragments was checked by cloning and sequencing corresponding genomic sequences (data not shown). Conversely, we found introns in *Pds*, *Zds*, *Hy-b*, *Zep*, and *Lcy-e* genomic sequences corresponding to RFLP probes. *Eco*RV did not cut the genomic sequences of *Pds*, *Zds*, *Hy-b*, *Zep*, and *Lcy-e*. In the same way, no *Bam*HI restriction site was found in the genomic sequences of *Pds*, *Zds*, and *Hy-b*. Data relative to the diversity observed for the different genes are presented in **Table 4**. A total of 58 fragments were identified, six of them being monomorphic (present in all individuals). In the limited sample of the three basic taxa, only eight bands out of 58 could not be observed. In the basic taxa, the mean number of bands per genotype observed was 24.7, 24.7, and 17 for *C. reticulata*, *C. maxima*, and *C. medica*, respectively. It varies from 28 (*C. limettioides*) to 36 (*C. aurantium*) for the secondary species. The mean number of RFLP bands per individual was lower for basic taxa than for the group of secondary species. This result indicates that secondary species are much more heterozygous than the basic ones for these genes, which is logical if we assume that the secondary species arise from hybridizations between the three basic taxa. Moreover *C. medica* appears to be the least heterozygous taxon for RFLP around the genes of the carotenoid biosynthetic pathway, as already shown with isozymes (*19*), RAPD (*20*), and SSR (*21*) markers.

The four sweet oranges analyzed displayed the same profiles for all genes as did the three representatives of *C. aurantium* and the three grapefruit. In the following analysis, each of these three secondary species is represented by only one individual. The organization of genetic diversity displayed on the neighbor-joining tree based on the Dice dissimilarity index from the presence or absence of bands observed for all RFLP markers is given in **Figure 2**. Eighteen different profiles were differentiated. Three main clusters were identified; the first grouped mandarins and sweet oranges, the second pummelos and grapefruit, and the third citrons and most of the acid *Citrus*. The two lemons were close to the acid *Citrus* cluster and the three sour oranges

Table 3. Primers Used for SSR Analysis

primer	putative function	primer sequence	observed product size (bp)
1210 forward	lycopene β-cyclase	5′-GCCAAAATGCATGTTCAAGA-3′	174–183
1210 reverse		5′-GTGCCAATGATGATCACGTC-3′	
1388 forward	β-carotene hydroxylase	5′-AAAACAAAGCACCCAGATCG-3′	133–142
1388 reverse		5′-ACGGCAGCAACGAGATAAGT-3′	

Table 4. Global Description of RFLP Profiles Observed Using cDNA of Seven Genes of the Carotenoid Biosynthetic Pathway as Probes

gene	restriction enzyme	total no. of fragments	no. of monomorphic fragments	minimum no. of fragments/ individual	maximum no. of fragments/ individual	no. of fragments/ *C. reticulata*	no. of fragments/ *C. maxima*	no. of fragments/ *C. medica*	no. of fragments not observed in basic taxa	no. of profiles	no. of profiles/ gene
Psy	*Eco*RV	5	1	2	3	3	3	2	0	6	10
	*Bam*HI	5	1	2	3	3	4	2	0	4	
Pds	*Eco*RV	6	1	2	3	2	2	2	2	8	8
Zds	*Eco*RV	9	1	2	6	4	5	2	2	11	13
	*Bam*HI	5	1	3	5	3	4	3	0	7	
Lcy-b	*Eco*RV	4	0	1	2	1	2	1	1	7	7
Hy-b	*Eco*RV	4	0	1	2	2	2	1	0	6	8
	*Bam*HI	2	1	1	2	1	2	1	0	2	
Zep	*Eco*RV	7	0	2	4	4	3	2	1	10	10
Lcy-e	*Eco*RV	5	0	1	3	2	1	1	1	3	10
	*Hin*dIII	6	0	1	5	3	1	1	1	9	

close to the mandarins/sweet oranges cluster. This organization of genetic diversity based on the RFLP profiles obtained with seven genes of the carotenoid pathway is very similar to that previously obtained with neutral molecular markers such as genomic SSR (*21*) as well as the organization obtained with qualitative carotenoid compositions (*18*). All these results suggest that the observed RFLP and SSR fragments are good phylogenetic markers. It seems consistent with our basic hypothesis that major differentiation in the genes involved in the carotenoid biosynthetic pathway preceded the creation of the secondary hybrid species and thus that the allelic structure of these hybrid species can be reconstructed from alleles observed in the three basic taxa.

Gene by Gene Analysis: The *Psy* Gene. For the *Psy* probe combined with *Eco*RV or *Bam*HI restriction enzymes, five bands were identified for the two enzymes, and two to three bands

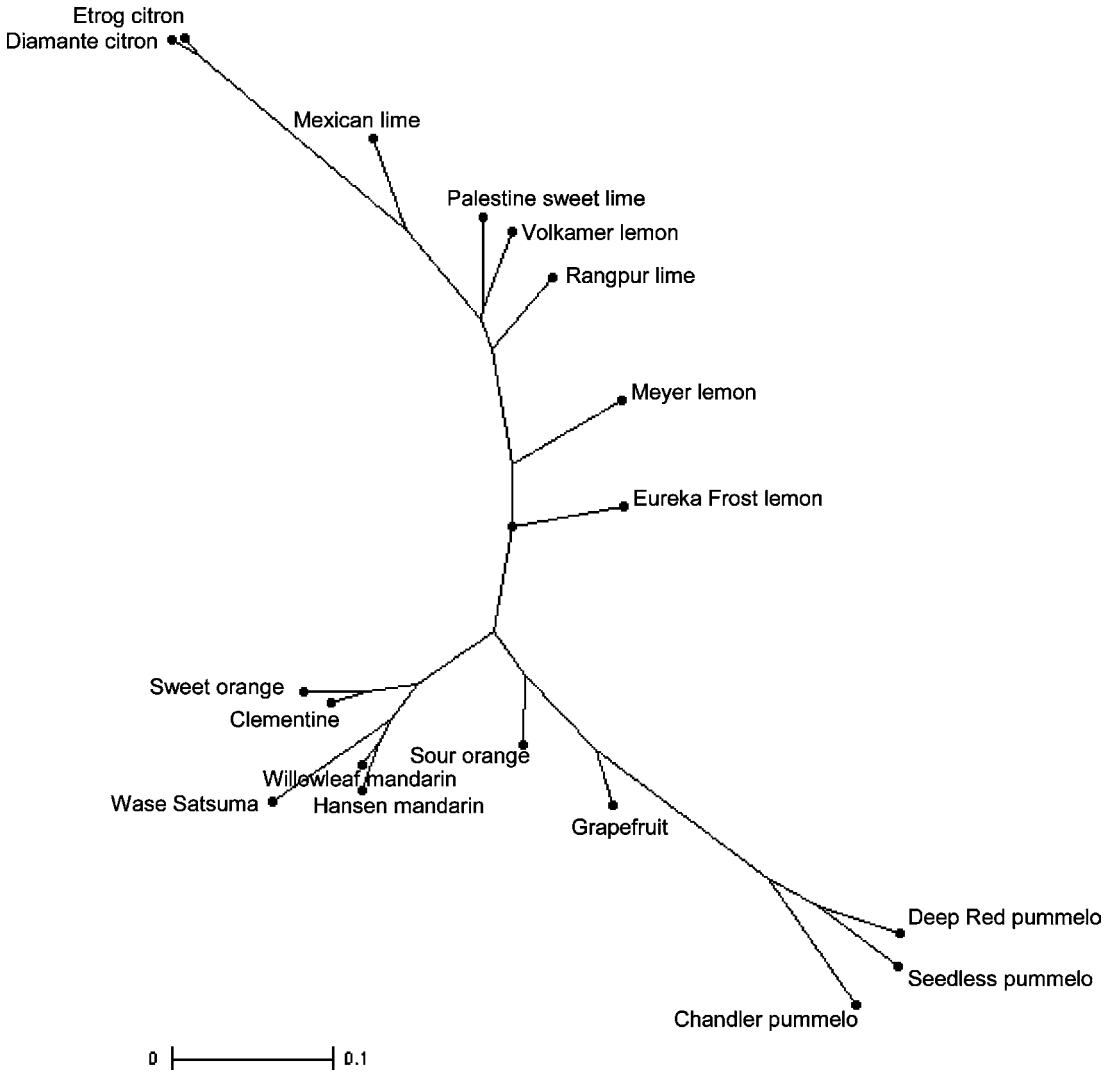

Figure 2. Organization of genetic diversity based on RFLP markers defined from expressed sequences of seven carotenoid biosynthetic genes. The tree was constructed according to the neighbor-joining method using a Dice matrix of dissimilarity.

were observed for each genotype. One of these bands was present in all individuals. There was no restriction site in the probe sequence. These results lead us to believe that *Psy* is present at two loci, one where no polymorphism was found with the restriction enzymes used, and one that displayed polymorphism. The number of different profiles observed was six and four with *Eco*RV and *Bam*HI, respectively, for a total of 10 different profiles among the 25 individuals (**Table 4**). Two *Psy* genes have also been found in tomato (*Lycopersicon esculentum*), tobacco (*Nicotiania tabacum*), maize (*Zea mays*), and rice (*Oryza sativa*) (*27*−*29*). Conversely, only one *Psy* gene has been found in *Arabidopsis thaliana* and in pepper (*Capsicum annuum*), which also accumulates carotenoids in fruit (*30*, *31*). According to Bartley and Scolnik (*27*), *Psy1* was expressed in tomato fruit chromoplasts, while *Psy2* was specific to leaf tissue. In the same way, in Poaceae (maize, rice), Gallagher et al. (*29*) found that *Psy* gene was duplicated and that *Psy1* and not *Psy2* transcripts in endosperm correlated with endosperm carotenoid accumulation. These results underline the role of gene duplication and the importance of tissue-specific phytoene synthase in the regulation of carotenoid accumulation.

All the polymorphic bands were present in the sample of the basic taxon genomes (**Table 5**). Assuming the hypothesis that all these bands describe the polymorphism at the same locus for the *Psy* gene, we can conclude that we found allelic differentiation between the three basic taxa with three alleles for *C. reticulata*, four for *C. maxima*, and one for *C. medica* (**Table 6**).

The alleles observed for the basic taxa then enabled us to determine the genotypes of all the other species. The presumed genotypes for the *Psy* polymorphic locus are given in **Table 7**. Sweet oranges and grapefruit were heterozygous (*Psy1_1*/*Psy1_5*) with one mandarin and one pummelo allele. Sour oranges were heterozygous (*Psy1_1*/*Psy1_4*); they shared the same mandarin allele with sweet oranges but had a different pummelo allele. Clementine was heterozygous (*Psy1_1*/*Psy1_2*) with two mandarin alleles; one shared with sweet oranges (*Psy1_1*) and one with "Willow leaf" mandarin. "Meyer" lemon was heterozygous (*Psy1_1*/*Psy1_8*), with the mandarin allele also found in sweet oranges, and the citron allele. "Eureka" lemon was also heterozygous with the same pummelo allele as sour oranges (*Psy1_4*) and the citron allele (*Psy1_8*). The other acid *Citrus* were homozygous for the citron allele.

The *Pds* Gene. For the *Pds* probe combined with *Eco*RV, six different fragments were observed. One was common to all individuals. The number of fragments per individual was two or three. Results for *Pds* led us to believe that this gene is present at two loci, one where no polymorphism was found with *Eco*RV restriction, and one displaying polymorphism. Conversely, studies on *Arabidopsis*, tomato, maize, and rice showed that *Pds* was a single copy gene (*30*, *32*). However, a previous study on *Citrus* (*10*) suggests that *Pds* is present as a low-copy gene family in the *Citrus* genome, which is in agreement with our findings.

Assuming that the polymorphic bands are related to the same locus, we observed a strong differentiation between the three basic taxa with one specific allele for each taxon (**Table 6**). All the genotypes of the basic taxa were homozygous and no polymorphism was found at intraspecific level. Two additional alleles were observed in "Volkamer" lemon (*Pds1_4*: E2) and "Mexican" lime (*Pds1_5*: E4). Genotypes of secondary species are given in **Table 7**. Sweet oranges, clementine, and sour oranges were heterozygous (*Pds1_1*/*Pds1_2*) with the mandarin and the pummelo allele. Grapefruit were homozygous (*Pds1_2*/*Pds1_2*) for the pummelo allele. "Eureka" lemon, "Rangpur" lime, and "Palestine sweet" lime were heterozygous (*Pds1_1*/*Pds1_3*) with the mandarin and the citron allele. "Meyer" lemon was heterozygous (*Pds1_2*/*Pds1_3*) with the pummelo and the citron allele. "Volkamer" lemon and "Mexican" lime were also heterozygous with the citron allele and their specific alleles.

The *Zds* Gene. The *Zds* profiles were complex. Nine and five fragments were observed with *Eco*RV and *Bam*HI restriction, respectively. For both enzymes, one fragment was common to all individuals. The number of fragments per individual ranged from two to six for *Eco*RV and three to five for *Bam*HI. There was no restriction site in the probe sequence. It can be assumed that several copies (at least three) of the *Zds* gene are present in the *Citrus* genome with polymorphism for at least two of them. In *Arabidopsis*, maize, and rice, like *Pds*, *Zds* was a single-copy gene (*30*, *32*).

In these conditions and in the absence of analysis of controlled progenies, we are unable to conduct genetic analysis of profiles. However it appears that some bands differentiated the basic taxa: one for mandarins, one for pummelos, and one for citrons with *Eco*RV restriction and one for pummelos and one for citrons with *Bam*HI restriction (**Table 5**). Two bands out of the nine obtained with *Eco*RV were not observed in the samples of basic taxa. One was rare and only observed in "Rangpur" lime. The other was found in sour oranges, "Volkamer" lemon, and "Palestine sweet" lime suggesting a common ancestor for these three genotypes.

This is in agreement with the assumption of Nicolosi et al. (*20*) that "Volkamer" lemon results from a complex hybrid combination with *C. aurantium* as one parent. It will be necessary to extend the analysis of the basic taxa to conclude whether these specific bands are present in the diversity of these taxa or result from mutations after the formation of the secondary species.

The *Lcy-b* Gene with RFLP Analysis. After restriction with *Eco*RV and hybridization with the *Lcy-b* probe, we obtained simple profiles with a total of four fragments. One to two fragments were observed for each individual, and seven profiles were differentiated among the 25 genotypes. These results provide evidence that *Lcy-b* is present at a single locus in the haploid *Citrus* genome. Two lycopene β-cyclases encoded by two genes have been identified in tomato (*33*). The B gene encoded a novel type of lycopene β-cyclase whose sequence was similar to capsanthin−capsorubin synthase. The B gene expressed at a high level in β mutants was responsible for strong accumulation of β-carotene in fruit, while in wild-type tomatoes, B was expressed at a low level.

The *Lcy-b* Gene with SSR Analysis. Four bands were detected at locus 1210 (*Lcy-b* gene) (**Figure 3**). One or two bands were detected per variety confirming that this gene is monolocus. Six different profiles were observed among the 25 genotypes. As with RFLP analysis, no intrataxon molecular polymorphism was found within *C. paradisi*, *C. sinensis*, and *C. aurantium*.

Taken together, the information obtained from RFLP and SSR analyses enabled us to identify a complete differentiation among the three basic taxon samples. Each of these taxons displayed two alleles for the analyzed sample. An additional allele (*Lcy-b_7*) was identified for "Mexican" lime (**Table 7**). The profiles for all secondary species can be reconstructed from these alleles. Deduced genetic structure is given in **Table 7**. Sweet oranges and clementine were heterozygous (*Lcy-b_1*/*Lcy-b_3*) with one mandarin and one pummelo allele. Sour oranges were also heterozygous (*Lcy-b_1*/*Lcy-b_4*) sharing the same mandarin

Table 5. RFLP and SSR Profiles of the Genotypes Analyzed (Only Polymorphic Fragments Are Presented)

	Psy gene		*Pds* gene	*Zds* gene		*Lcy-b* gene	
	*Eco*RV	*Bam*H1	*Eco*RV	*Eco*RV	*Bam*H1	*Eco*RV	SSR
C. reticulata cv. Willow leaf	E1[a]E2	B1[b]	E1	E2E3E6	B1B3	E1	S3[c]
C. reticulata cv. Satsuma	E1	B3	E1	E2E3E6	B1B3	E1	S3S4
C. reticulata cv. Hansen	E1	B1	E1	E2E3E6	B1B3B4	E1	S3
C. maxima cv. Seedless	E1E2	B2B4	E3	E1E2E6	B1B2B3	E2	S3
C. maxima cv. Deep red	E2	B4	E3	E1E2E6E8	B1B2B3	E2	S3
C. maxima cv. Chandler	E4	B1	E3	E1E2	B1B2B3	E4	S3
C. medica cv. Etrog	E3	B1	E5	E7	B3B4	E4	S2
C. medica cv. Diamante	E3	B1	E5	E7	B3B4	E4	S1S2
C. sinensis[d]	E1E2	B1B4	E1E3	E1E2E3E6	B1B3	E1E2	S3
C. aurantium[d]	E1	B1B4	E1E3	E1E2E3E5E6	B1B2B3B4	E1E4	S3
C. paradisi[d]	E1E2	B1B4	E3	E1E2E3E6	B1B2B3	E2E4	S3
C. limon cv. Eureka	E1E3	B1B4	E1E5	E1E2E6	B2B3B4	E1E4	S2S3
C. limonia cv. Volkamer	E3	B1	E2E5	E5E7	B1B3B4	E1E4	S1S3
C. meyeri cv. Meyer	E1E3	B1	E3E5	E2E3E6E7	B1B3B4	E1E4	S2S3
C. limonia cv. Rangpur	E3	B1	E1E5	E4E5E7	B1B2B4	E1E4	S2S3
C. aurantifolia cv. Mexican	E3	B1	E4E5	E6E7	B1B2B4	E3E4	S1S3
C. limettioides cv. Palestine	E3	B1	E1E5	E5E7	B1B3B4	E1E4	S2S3
C. clementina	E1E2	B1	E1E3	E2E3E6	B1B3	E1E2	S3
C. reticulata cv. Willow leaf	E3	*e*	S2	E1E6	E4E5	H1[f]H4H5	
C. reticulata cv. Satsuma	E3E4		S2	E1E2E6E7	E4E5	H1H4H5	
C. reticulata cv. Hansen	E3		S2	E1E6	E4E5	H1H4H5	
C. maxima cv. Seedless	E2E3	B1	S1S2	E1E2E4E7	E2	H3	
C. maxima cv. Deep red	E2E3	B1	S2	E2E7	E2	H3	
C. maxima cv. Chandler	E2E3	B1	S2	E2E7	E2	H3	
C. medica cv. Etrog	E1		S1	E2E5	E1	H3H6	
C. medica cv. Diamante	E1		S1	E2E5	E1	H6	
C. sinensis[d]	E3		S2	E1E2E4E6	E4E5	H1H2H4H5	
C. aurantium[d]	E2E3		S2	E1E2E6E7	E2E4E5	H1H2H3H4H5	
C. paradisi[d]	E3	B1	S2	E1E2E3E6	E2E4E5	H1H2H3H5	
C. limon cv. Eureka	E2E3		S1S2	E1E2E5E6	E1E4E5	H1H4H5H6	
C. limonia cv. Volkamer	E1E3		S1S2	E1E2E5E6	E1E4E5	H2H4H5H6	
C. meyeri cv. Meyer	E1E3		S1S2	E2E5E7	E1E4E5	H1H4H5H6	
C. limonia cv. Rangpur	E1E3		S1S2	E1E2E5E6	E1E4E5	H1H4H5H6	
C. aurantifolia cv. Mexican			S1S3	E2E5E6	E1E3	H3H5H6	
C. limettioides cv. Palestine	E2		S1S2	E2E5E6	E1E4E5	H2H4H5H6	
C. clementina	E3		S2	E1E6	E4E5	H1H2H4H5	

[a] E1, polymorphic band 1 obtained from RFLP analysis with *Eco*RV restriction. [b] B1, polymorphic band 1 obtained from RFLP analysis with *Bam*H1 restriction. [c] S3, polymorphic band 3 obtained from SSR analysis. [d] Each of these secondary species is represented by only one individual. [e] There is no polymorphic band for this combination: *Bam*H1 restriction enzyme and the *Hy-b* probe in this genotype. [f] H1, polymorphic band 1 obtained from RFLP analysis with *Hind*III restriction.

Table 6. Alleles Identified in the Three Basic Species[a]

	alleles of the *Psy* gene	alleles of the *Pds* gene	alleles of the *Lcy-b* gene	alleles of the *Lcy-e* gene
C. reticulata	*Psy1_1*, E1B1 *Psy1_2*, E2B1 *Psy1_3*, E1B3	*Pds1_1*, E1	*Lcy-b_1*, E1S3 *Lcy-b_2*, E1S4	*Lcy-e1_4*, E4E5
C. maxima	*Psy1_4*, E1B4 *Psy1_5*, E2B4 *Psy1_6*, E2B2 *Psy1_7*, E4B1	*Pds1_2*, E3	*Lcy-b_3*, E2S3 *Lcy-b_4*, E4S3	*Lcy-e1_2*, E2
C. medica	*Psy1_8*, E3B1	*Pds1_3*, E5	*Lcy-b_5*, E4S1 *Lcy-b_6*, E4S2	*Lcy-e1_1*, E1

[a] Alleles are named by the following convention: *Psy1_1* denotes allele 1 of locus 1 of the *Psy* gene.

Table 7. Proposed Genotypes and Proposed Phylogenetic Origin of Carotenoid Biosynthetic Genes of Hybrid *Citrus* Species

gene:	*Psy1*		*Pds1*		*Lcy-b*		*Lcy-e1*	
data used:	*Eco*RV + *Bam*HI		*Eco*RV		*Eco*RV + SSR		*Eco*RV	
	genotype[a]	phylogenetic origin[b]	genotype[a]	phylogenetic origin[b]	genotype[a]	phylogenetic origin[b]	genotype[a]	phylogenetic origin[b]
C. reticulata cv. Willow leaf	*Psy1_1*/*Psy1_2*	M/M	*Pds1_1*/*Pds1_1*	M/M	*Lcy-b_1*/*Lcy-b_1*	M/M	*Lcy-e1_4*/*Lcy-e1_4*	M/M
C. reticulata cv. Satsuma	*Psy1_3*/*Psy1_3*	M/M	*Pds1_1*/*Pds1_1*	M/M	*Lcy-b_1*/*Lcy-b_2*	M/M	*Lcy-e1_4*/*Lcy-e1_4*	M/M
C. reticulata cv. Hansen	*Psy1_1*/*Psy1_1*	M/M	*Pds1_1*/*Pds1_1*	M/M	*Lcy-b_1*/*Lcy-b_1*	M/M	*Lcy-e1_4*/*Lcy-e1_4*	M/M
C. maxima cv. seedless	*Psy1_4*/*Psy1_6*	P/P	*Pds1_2*/*Pds1_2*	P/P	*Lcy-b_3*/*Lcy-b_3*	P/P	*Lcy-e1_2*/*Lcy-e1_2*	P/P
C. maxima cv. Deep red	*Psy1_5*/*Psy1_5*	P/P	*Pds1_2*/*Pds1_2*	P/P	*Lcy-b_3*/*Lcy-b_3*	P/P	*Lcy-e1_2*/*Lcy-e1_2*	P/P
C. maxima cv. Chandler	*Psy1_7*/*Psy1_7*	P/P	*Pds1_2*/*Pds1_2*	P/P	*Lcy-b_4*/*Lcy-b_4*	P/P	*Lcy-e1_2*/*Lcy-e1_2*	P/P
C. medica cv. Etrog	*Psy1_8*/*Psy1_8*	C/C	*Pds1_3*/*Pds1_3*	C/C	*Lcy-b_6*/*Lcy-b_6*	C/C	*Lcy-e1_1*/*Lcy-e1_1*	C/C
C. medica cv. Diamante	*Psy1_8*/*Psy1_8*	C/C	*Pds1_3*/*Pds1_3*	C/C	*Lcy-b_5*/*Lcy-b_6*	C/C	*Lcy-e1_1*/*Lcy-e1_1*	C/C
C. sinensis	*Psy1_1*/*Psy1_5*	M/P	*Pds1_1*/*Pds1_2*	M/P	*Lcy-b_1*/*Lcy-b_3*	M/P	*Lcy-e1_4*/*Lcy-e1_4*	M/M
C. aurantium	*Psy1_1*/*Psy1_4*	M/P	*Pds1_1*/*Pds1_2*	M/P	*Lcy-b_1*/*Lcy-b_4*	M/P	*Lcy-e1_4*/*Lcy-e1_2*	M/P
C. paradisi	*Psy1_1*/*Psy1_5*	M/P	*Pds1_2*/*Pds1_2*	P/P	*Lcy-b_3*/*Lcy-b_4*	P/P	*Lcy-e1_4*/*Lcy-e1_2*	M/P
C. limon	*Psy1_4*/*Psy1_8*	P/C	*Pds1_1*/*Pds1_3*	M/C	*Lcy-b_1*/*Lcy-b_6*	M/C	*Lcy-e1_4*/*Lcy-e1_1*	M/C
C. limonia	*Psy1_8*/*Psy1_8*	C/C	*Pds1_3*/*Pds1_4*	C/?	*Lcy-b_1*/*Lcy-b_5*	M/C	*Lcy-e1_4*/*Lcy-e1_1*	M/C
C. meyeri	*Psy1_1*/*Psy1_8*	M/C	*Pds1_2*/*Pds1_3*	P/C	*Lcy-b_1*/*Lcy-b_6*	M/C	*Lcy-e1_4*/*Lcy-e1_1*	M/C
C. limonia	*Psy1_8*/*Psy1_8*	C/C	*Pds1_1*/*Pds1_3*	M/C	*Lcy-b_1*/*Lcy-b_6*	M/C	*Lcy-e1_4*/*Lcy-e1_1*	M/C
C. aurantifolia	*Psy1_8*/*Psy1_8*	C/C	*Pds1_3*/*Pds1_5*	C/?	*Lcy-b_7*/*Lcy-b_5*	?/C	*Lcy-e1_3*/*Lcy-e1_1*	?/C
C. limettioides	*Psy1_8*/*Psy1_8*	C/C	*Pds1_1*/*Pds1_3*	M/C	*Lcy-b_1*/*Lcy-b_6*	M/C	*Lcy-e1_4*/*Lcy-e1_1*	M/C
C. clementina	*Psy1_1*/*Psy1_2*	M/M	*Pds1_1*/*Pds1_2*	M/P	*Lcy-b_1*/*Lcy-b_3*	M/P	*Lcy-e1_4*/*Lcy-e1_4*	M/M

[a] For each carotenoid biosynthetic gene, loci and alleles are identified by the following convention: *Psy1_1* for allele 1 of locus 1 of the *Psy* gene. [b] M, mandarin (*C. reticulata*); P, pummelo (*C. maxima*); C, citron (*C. medica*); ?, specific allele not observed in the three basic taxa.

allele as sweet oranges but with another pummelo allele. Grapefruit were heterozygous (*Lcy-b_3*/*Lcy-b_4*) with two pummelo alleles. All the acid secondary species were heterozygous, having one allele from citrons and the other one from mandarins (*Lcy-b_1*) except for "Mexican" lime, which had a specific allele.

The *Hy-b* Gene with RFLP Analysis. Four and two bands were clearly differentiated with *Eco*RV and *Bam*HI restriction enzymes, respectively. One to two bands were observed for each genotype with each enzyme. This would appear to be coherent with the monolocus pattern of *Hy-b*. However the fact that the three pummelos presented the same two band profiles with the two enzymes make this hypothesis questionable. Indeed the probability of these three genotypes being heterozygous for the same alleles is rather low, and an alternative hypothesis should be that pummelos have two copies of the *Hy-b* gene per haploid genome. It prevented us from proposing a genetic interpretation for the concerned profiles. Moreover, two *Hy-b* genes were found in the *Arabidopsis thaliana* genome and in tomato (*30*, *34*), while the Southern blot pattern obtained from "Satsuma" mandarin (*11*) with three restriction enzymes showed one to three bands, which is more consistent with the hypothesis of one locus. Additional analyses with more restriction enzymes and the study of haploid genotypes of the three basic taxa will probably enable us to determine the number of loci of *Hy-b* in the *Citrus* genome.

All bands were observed in the basic taxon samples with clear differentiation among them (**Table 5**). With *Eco*RV, one band (E1) was shared by the two citrons, "Volkamer" and "Meyer" lemons, and "Rangpur" lime; one (E2) was shared by the three pummelos, the three *C. aurantium*, and "Palestine sweet" lime. The E3 band was present for all samples except citrons, and one band (E4) was specific to Satsuma. With *Bam*HI, all

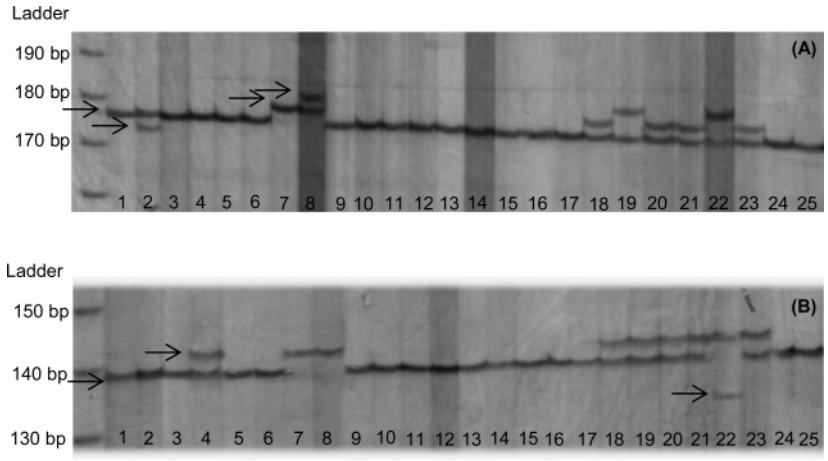

Figure 3. Silver-stained 5% acrylamide gel showing polymorphism at locus 1210 (A, *Lcy-b* gene) and 1388 (B, *Hy-b* gene) in 25 *Citrus* genotypes. Numbers represent genotypes described in **Table 1**. The arrows indicate different fragments (174, 177, 180, and 183 bp for the gel A and 133, 139, and 142 bp for the gel B).

individuals displayed a common band, while a second band was only present for the three pummelos and the three grapefruits.

The *Hy-b* Gene with SSR Analysis. Three alleles were detected at locus 1388 (*Hy-b* gene, **Figure 3**). One or two alleles were detected for each variety. No intrataxon molecular polymorphism was found within grapefruit, sweet oranges, and sour oranges. One band (S1) was shared by "Seedless" pummelo, the two citrons, and all acid *Citrus*. The S2 band was observed for all *Citrus* except citrons and "Mexican" lime. This latter variety displayed a specific S3 band. All acid varieties were heterozygous and possessed the citron allele. It should be noted that with both RFLP and SSR analyses, except for S3, all fragments were observed in the basic taxon samples.

The *Zep* Gene. Seven fragments were observed with the *Eco*RV restriction enzyme. Ten different profiles were identified among the 25 varieties. No *Eco*RV restriction site was found in the probe sequence. Each variety displayed two to four bands, which suggests that this gene is present at two different loci in the *Citrus* genome and that we revealed polymorphism for the two loci. Only one copy of the *Zep* gene was found in *Arabidopsis* genome (*30*). No previous data were available on the number of copies of *Zep* in *Citrus* genome.

All bands, except one only found in grapefruit (E3), were observed in the basic taxon samples (**Table 5**). Band E5 was specific to citrons among the basic taxon samples and shared by all acid *Citrus*. Band E6 was specific to mandarins among the basic taxon samples and shared in heterozygous status by all grapefruit, sweet oranges, and sour oranges, as well as clementine and all acid *Citrus* except "Meyer" lemon. Except for "Seedless" pummelo and "Satsuma" mandarin, all other members of the basic taxa presented only two bands and were probably homozygous at the two loci. Moreover these homozygous individuals did not display intrataxon diversity. Most of the secondary species were heterozygous for at least one of the loci (three bands or more) except for clementine, which was identical to "Willow leaf" mandarin with only two bands.

The *Lcy-e* Gene. Five RFLP fragments were observed with *Eco*RV and five with *Hin*dIII restriction enzymes. With the two restriction enzymes, we identified 11 different profiles. The number of bands per individual ranged from one to three with *Eco*RV and one to five with *Hin*dIII. The individuals of two basic taxa, *C. medica* and *C. maxima*, displayed only one band for *Eco*RV and *Hin*dIII, while the three varieties of *C. reticulata* displayed the same profile with two bands with *Eco*RV and three with *Hin*dIII. Taking into account (i) the presence of a *Hin*dIII restriction site in the EST probe sequence and (ii) the observation of three and five bands with *Eco*RV and *Hin*dIII, respectively, for secondary species, we proposed the hypothesis that *Lcy-e* is present at only one locus in *C. maxima* and *C. medica* genomes but at two loci in mandarin and other *Citrus* genomes. The results of Ronen et al. (*35*) on tomato suggest that *Lcy-e* is a single-copy gene and that the *Del* mutation is an allele of *Lcy-e*. The tomato delta mutants were orange, and accumulated δ-carotene at the expense of lycopene. The delta mutation changed the mRNA level of the *Lcy-e* gene during fruit development.

Due to the complexity and incertitude of genetic interpretation because of the presence of a *Hin*dIII restriction site in the probe sequence, we limited our genetic interpretation to the profiles observed with *Eco*RV. Differentiation among the three basic taxa was observed (**Table 5**) with one specific band for *C. medica* (E1), one very close by for *C. maxima* (E2), and two specific bands for *C. reticulata* (E4 and E5). The two bands of *C. reticulata* obtained with *Eco*RV were also observed in all secondary species that had mandarin as one ancestral parent. In the same way, one of the bands observed in *C. medica* or *C. maxima* was observed in all secondary species with the exception of sweet oranges and clementine, which displayed the same profile as mandarins. "Mexican" lime presented a specific band (E3) and the band of citrons. Under the hypothesis that *Lcy-e* is present in two copies per haploid genome in *C. reticulata*, whereas it is present in only one copy in *C. maxima* and *C. medica*, most of the secondary species should be heterozygous for the common locus (**Table 7**). *C. paradisi* and *C. aurantium* presented the allele of mandarins and the allele of pummelos, while all acid *Citrus* except "Mexican" lime presented the citron and the mandarin allele. However, sweet oranges and clementine should be homozygous for the mandarin allele. For the second locus of *C. reticulata*, we were unable to determine whether the secondary species were homozygous or heterozygous (presence/absence).

Allelic Structures of Major Commercial *Citrus* Species Are in Agreement with Classical Hypothesis on Cultivated Citrus Evolution. The proposed genotypes of secondary species from the alleles proposed for the three basic taxa (**Table 7**) are in full agreement with the hypothesis on the origin of these secondary species (*19*−*21*). We illustrate this assumption for six major commercial *Citrus* species.

Sweet oranges and sour oranges were assumed to be complex hybrids between mandarins and pummelos, probably back-

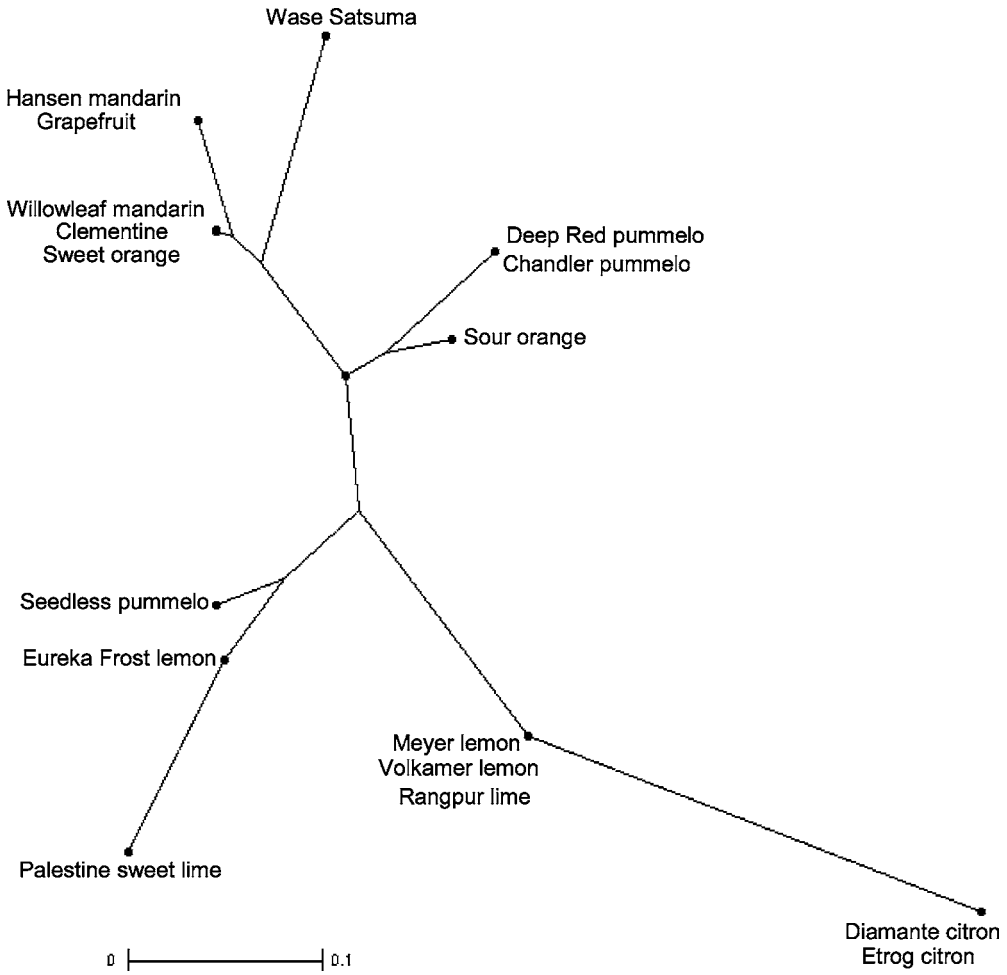

Figure 4. Representation of diversity based on RFLP and SSR analysis for the *Hy-b* gene. The tree was constructed according to the neighbor-joining method using a Dice matrix of dissimilarity.

crossed with the mandarin gene pool (*19*). For *Psy* and *Lcy-b* genes, they presented heterozygotic profiles sharing the same mandarin allele and two different pummelo alleles. For *Pds*, sweet and sour oranges shared the same heterozygotic profile with one mandarin and one pummelo allele. For *Lcy-e* analyzed after *Eco*RV restriction, we concluded that sweet oranges had two mandarin alleles, while sour oranges were heterozygous with one pummelo and one mandarin allele. For all other genes, we found a differentiation between sweet and sour oranges. Sweet oranges were generally closer to mandarins than to sour oranges.

Clementine was assumed to be a hybrid between "Willow leaf" mandarin and a sweet orange (*20*). We found the following organization for the carotenoid biosynthetic genes: Clementine was heterozygous for the *Psy* gene with two mandarin alleles, one shared with sweet oranges and one specific to "Willow leaf" mandarin. It was heterozygous for *Pds* and *Lcy-b* with, for each gene, one allele of pummelo (shared with sweet oranges) and one allele of mandarin. Moreover, clementine presented the same profiles as "Willow leaf" mandarin for *Hy-b* and *Zep* genes. Its profile was identical to sweet oranges for *Lcy-e*, and we thus concluded that they share two alleles from *C. reticulata* for the locus common to all taxa.

Grapefruit is supposed to be a hybrid between pummelos and sweet oranges that originated in the Caribbean after the introduction of *Citrus* in the New World by Christopher Columbus (*19*, *20*). For the *Psy* gene, *C. paradisi* was heterozygous and displayed an identical profile to sweet oranges with one pummelo and one mandarin allele. For *Pds* and *Lcy-b*, *C. paradisi* displayed only pummelo alleles with a homozygous status for *Pds* and heterozygous for *Lcy-b*. The pummelo alleles found in sweet oranges were present in the two genes. For *Lcy-e*, from the *Eco*RV analysis, we concluded that grapefruit were heterozygous with one mandarin and one pummelo allele. Moreover grapefruit shared the *Hin*dIII restriction bands of sweet oranges not observed in the limited basic taxon samples.

Nicolosi et al. (*20*) proposed that lemons arose from hybridization between sour oranges and citrons. For *Psy*, we found that "Eureka" lemon was heterozygous with one allele of citron and one of pummelo also found in sour oranges. For *Pds*, *Lcy-b*, and *Lcy-e*, "Eureka" lemon was heterozygous with one allele of mandarin (also observed in sour oranges) and one citron allele.

"Mexican" lime was assumed to be a hybrid between *C. medica* and *C. micrantha* (*20*). For all genes of the carotenoid biosynthetic pathway, we found that "Mexican" lime had one allele or specific bands from citrons, confirming the direct parentage of citrons. Moreover, among the eight out of 58 fragments not observed in the three basic taxa, three were specific to "Mexican" lime, as was one SSR band for *Hy-b*. It can be supposed that these four rare alleles of Mexican lime come from *C. micrantha*, which was not included in the study.

Relation between the Phylogenic Origin of Alleles and Carotenoid Content. From analysis of carotenoid contents, we previously proposed (*18*) that modification of four major steps of the biosynthetic pathway can explain the qualitative phenotypic differentiation among *Citrus* species. The variability of

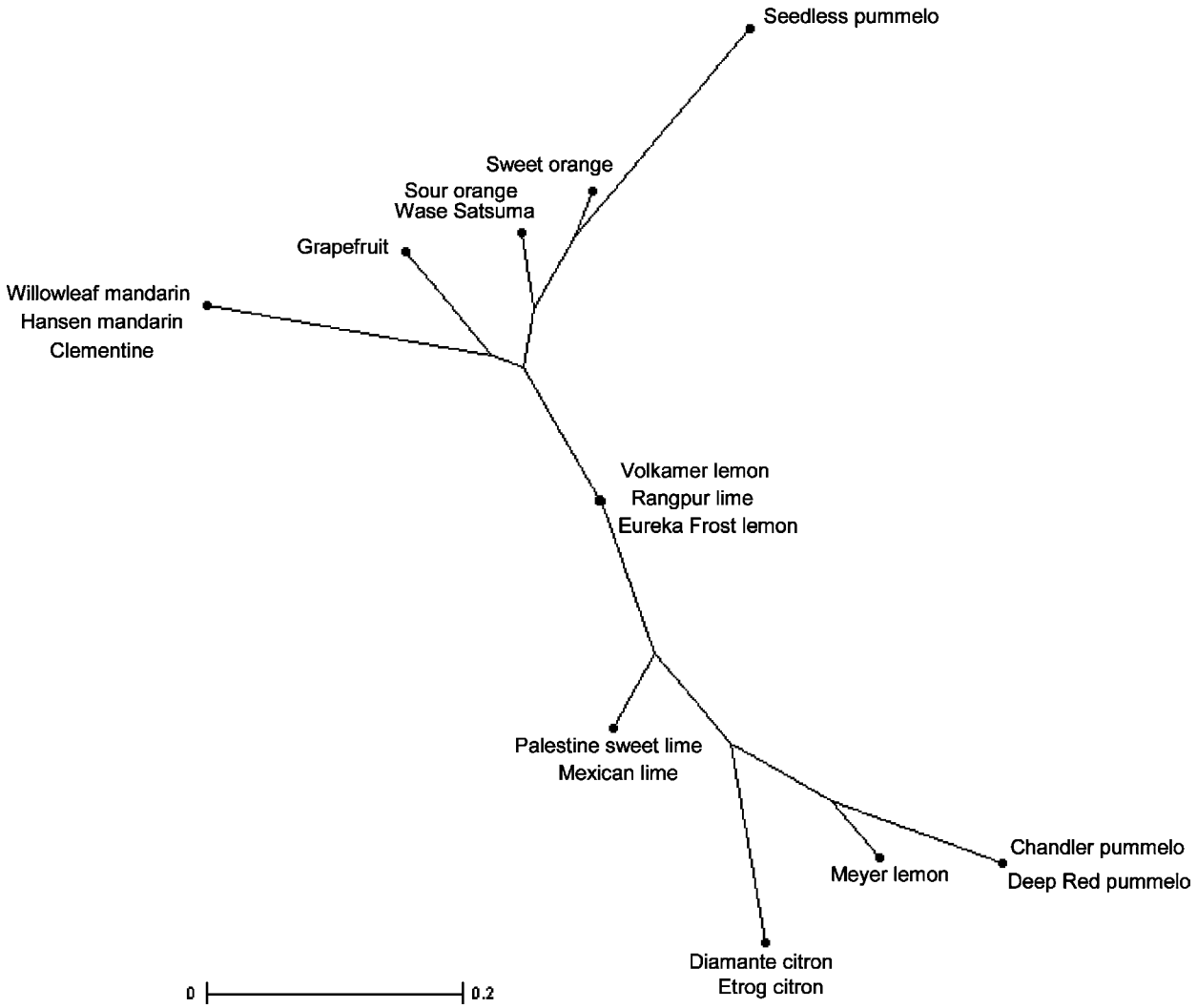

Figure 5. Representation of the diversity based on RFLP analysis around the *Zep* gene. The tree was constructed according to the neighbor-joining method using a Dice matrix of dissimilarity.

these steps is probably due to a modification in the specific activity of the key enzymes (supported by allelic variability of the corresponding genes) or by modifications in gene expression. By analyzing the relation between phenotypic variability (*18*) and the organization of genetic diversity of the key genes observed in the same genotype samples, we propose the hypothesis that phenotypic variability is linked to allelic diversity of the key gene for the four steps concerned.

We proposed (*18*) that the formation of phytoene from geranylgeranyl diphosphate catalyzed by phytoene synthase is a key step in the differentiation of citrons and several acid lemons, such as "Palestine sweet" lime and "Eureka" and "Meyer" lemons, from the other *Citrus*. In the present study, we found that for the polymorphic *Psy* locus, both acid *Citrus* producing very low levels of carotenoids (citrons, lemons) and those with higher carotenoid contents ("Rangpur" lime, "Volkamer" lemon) were homozygous for the citron alleles. It is thus clear that the phenotypic differentiation cannot be attributed to the allelic variability observed among the basic taxa for *Psy* locus. The carotenoid composition of acid *Citrus* may be due to the level of expression of carotenoid biosynthetic genes. Kato et al. (*15*) showed that the level of expression of genes that produced β,β-xanthophylls was lower in juice sacs of "Lisbon" lemon than in juice sacs of "Satsuma" mandarin. The genes responsible for carotenoid catabolism were also shown to be involved. According to Kato et al. (*36*), the expression level of carotenoid cleavage dioxygenase genes (*CitNCED* genes), which are involved in the cleavage of β,β-xanthophylls and abscissic acid synthesis, controlled the accumulation of 9-*cis*-violaxanthin in juice sacs of "Lisbon" lemon, "Satsuma" mandarin, and "Valencia" orange. Moreover the expression level of genes of the methylerythritol phosphate pathway may play an important role as is the case in tomato fruit. Indeed, studies on tomato showed that carotenoid contents were controlled by the coordinated expression of *Dxs* and *Psy* genes (*5*). Regulations at the transcriptional level appear to play a major role in acid *Citrus*.

A second key step is the cyclization of lycopene catalyzed by the lycopene β-cyclase (LCY-b). We proposed (*18*) that this step is involved in the differentiation of pummelos and grapefruit from other *Citrus*. Pummelos and grapefruit produced β-carotene but accumulated mainly lycopene. Our results showed that *Lcy-b* was a single-copy gene and that only grapefruit had two pummelo alleles for this gene, while other secondary species presented one allele from mandarin and one allele from pummelo or from citron. This suggests that the carotenoid composition of pummelos and grapefruit is due to pummelo alleles for the *Lcy-b* gene. The pummelo allele might be involved in down-regulation of the activity of LCY-b occurring at either the transcriptional or post-transcriptional level.

The two other key steps are catalyzed by β-carotene hydroxylase (HY-b) and zeaxanthin epoxidase (ZEP), which

convert β-carotene into violaxanthin. According to results of our previous study (*18*), mandarins, sweet and sour oranges, and clementine were characterized by the presence of both β-cryptoxanthin and violaxanthin in juice sacs. Pummelo and grapefruits do not produce β-cryptoxanthin. "Rangpur" lime and "Volkamer" lemon, which are hybrids between citrons and mandarins and citrons and sour oranges (probably backcrossed with mandarins (*20*)), also produced β-cryptoxanthin and violaxanthin though in smaller amounts.

A global analysis of the relationship using data from a neighbor-joining analysis of RFLP and SSR data of *Hy-b* gene is presented in **Figure 4**. The tree shows a clear separation between mandarin, clementine, sweet and sour orange, and grapefruit genotypes and acid *Citrus*. "Seedless" pummelo is in the acid *Citrus* cluster because it shares with them the S1 SSR fragment. "Deep Red" pummelo and all grapefruit that did not produce β-cryptoxanthin but did produce β-carotene and accumulated lycopene were grouped in the mandarin/orange cluster. This provides some evidence that the diversity in carotenoid composition at the interspecific level is not linked to the allelic diversity observed for the *Hy-b* gene. The level of expression of this gene in the juice sacs of pummelos and grapefruit may play a role. However, no study has yet been conducted on the expression of carotenoid biosynthetic genes in juice sacs of pummelo or grapefruit fruit.

Global genetic relationships for the *Zep* gene revealed by neighbor-joining tree analysis are presented in **Figure 5**. Grapefruit were clustered with mandarins and sweet and sour oranges and "Eureka" lemon was close to "Volkamer" lemon and "Rangpur" lime. These results suggest that carotenoid composition is not linked to the allelic diversity observed for the *Zep* gene.

Previous studies on carotenoid composition of *Citrus* juices indicated that only mandarins, clementines, and sweet oranges accumulated α-carotene, zeinoxanthin, and lutein (*14*, *16*−*18*). The genetic interpretation of *Lcy-e* RFLP profiles leads us to propose that *Lcy-e* is present in one copy in *C. maxima* and *C. medica* and in two copies in *C. reticulata* and secondary species. Under this hypothesis, only mandarins, sweet oranges, and clementine were homozygous for the mandarin allele at the common locus, suggesting that allelic variability plays a role in the synthesis of α-carotene and β,ε-xanthophylls.

ABBREVIATIONS USED

GGPP, geranylgeranyl diphosphate; HY-b, β-carotene hydroxylase; HY-e, ε-carotene hydroxylase; LCY-b, lycopene β-cyclase; LCY-e, lycopene ε-cyclase; MEP pathway, methylerythritol phosphate pathway; PDS, phytoene desaturase; PSY, phytoene synthase; RFLP, restriction fragment length polymorphism; SSR, simple sequence repeats; ZDS, ζ-carotene desaturase; ZEP, zeaxanthin epoxidase.

ACKNOWLEDGMENT

We thank C. Jacquemond and F. Curk (Unité GEQA, INRA, San Giuliano, France) for their assistance and for helpful discussions.

LITERATURE CITED

(1) Olsen, J. A. Provatim A function of carotenoid. The conversion of beta-carotene into vitamin A. *J. Nutr.* **1989**, *119*, 105−108.
(2) Krinsky, N. Y. Actions of carotenoids in biological systems. *Annu. Rev. Nutr.* **1993**, *13*, 561−587.
(3) Hirschberg, J. Carotenoid biosynthesis in flowering plants. *Curr. Opin. Plant Biol.* **2001**, *4*, 210−218.
(4) Sandmann, G. Carotenoid biosynthesis and biotechnological application. *Arch. Biochem. Biophys.* **2001**, *385*, 4−12.
(5) Rodriguez-Concepcion, M.; Boronat, A. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol.* **2002**, *130*, 1079−1089.
(6) Isaacson, T.; Ronen, G.; Zamir, D.; Hirschberg, J. Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of β-carotene and xanthophylls in plants. *Plant Cell* **2002**, *14*, 333−342.
(7) Park, H.; Kreunen, S. S.; Cutriss, A. J.; DellaPenna, D.; Pogson, B. J. Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, promellar body formation and photomorphogenesis. *Plant Cell* **2002**, *14*, 321−332.
(8) Fraser, P. D.; Bramley, P. M. The biosynthesis and nutritional uses of carotenoids. *Prog. Lipid Res.* **2004**, *43*, 228−265.
(9) Ikoma, Y.; Komatsu, A.; Kita, M.; Ogawa, K.; Omura, M.; Yano, M.; Moriguchi, T. Expression of a phytoene synthase gene and characteristic carotenoid accumulation during citrus fruit development. *Physiol. Plant.* **2001**, *111*, 232−238.
(10) Kita, M.; Komatsu, A.; Omura, M.; Yano, M.; Ikoma, Y.; Moriguchi, T. Cloning and expression of CitPDS1, a gene encoding phytoene desaturase in Citrus. *Biosci., Biotechnol., Biochem.* **2001**, *65*, 1424−1428.
(11) Kim, I.-J.; Ko, K.-C.; Kim, C.-S.; Chung, W.-I. Isolation and characterization of cDNA encoding β-carotene hydroxylase in *Citrus*. *Plant Sci.* **2001**, *161*, 1005−1010.
(12) Inoue, K.; Furbee, K. J.; Uratsu, S.; Kato, M.; Dandekar, A. M.; Ikoma, Y. Catalytic activities and chloroplast import of carotenogenic enzymes from citrus. *Physiol. Plant.* **2006**, *127* (4), 561−570.
(13) Mouly, P. P.; Gaydou, E. M.; Lapierre, L.; Corsetti, J. Differentiation of several geographical origins in single-strengh Valencia orange juices using quantitative comparison of carotenoid profiles. *J. Agric. Food Chem.* **1999**, *47*, 4038−4045.
(14) Dhuique-Mayer, C.; Caris-Veyrat, C.; Ollitrault, P.; Curk, F.; Amiot, M.-J. Varietal and interspecific influence on micronutrient contents in citrus from the mediterranean area. *J. Agric. Food Chem.* **2005**, *53*, 2140−2145.
(15) Kato, M.; Ikoma, Y.; Matsumoto, H.; Sugiura, M.; Hyodo, H.; Yano, M. Accumulation of carotenoids and expression of carotenoid biosynthetic genes during maturation in citrus fruit. *Plant Physiol.* **2004**, *134*, 1−14.
(16) Goodner, K. L.; Rouseff, R. L.; Hofsommer, H. J. Orange, mandarin, and hybrid classification using multivariate statistics based on carotenoid profiles. *J. Agric. Food Chem.* **2001**, *49*, 1146−1150.
(17) Xu, C.-J.; Fraser, P. D.; Wang, W.-J.; Bramley, P. M. Differences in the carotenoid content of ordinary citrus and lycopene-accumulating mutants. *J. Agric. Food Chem.* **2006**, *54*, 5474−5481.
(18) Fanciullino, A. L.; Dhuique-Mayer, C.; Luro, F.; Casanova, J.; Morillon, R.; Ollitrault, P. Carotenoid diversity in cultivated citrus is highly influenced by genetic factors. *J. Agric. Food Chem.* **2006**, *54*, 4397−4406.
(19) Ollitrault, P.; Dambier, D.; Froelicher, Y.; Luro, F.; Cottin, R. La diversité des agrumes: structuration et exploitation par hybridation somatique. *C. R. Acad. Agric.* **2000**, *86*, 197−221.
(20) Nicolosi, E.; Deng, Z. N.; Gentile, A.; LaMalfa, S.; Continella, G.; Tribulato, E. Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theor. Appl. Genet.* **2000**, *100*, 1155−1166.
(21) Luro, F.; Rist, D.; Ollitrault, P. Evaluation of genetic relationships in Citrus genus by means of sequence tagged microsatellites. *Acta Hortic.* **2001**, *546*, 237−242.
(22) Doyle, J. J.; Doyle, J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11−15.

(23) Manning, K. Isolation of nucleic acids from plants by differential solvent precipitation. *Anal. Biochem*. **1991**, *195*, 45−50.
(24) Sambrook, J.; Fitsh, E. F.; Maniatis, T. *Molecular cloning: a laboratory manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.
(25) Terol, J.; Conesa, A.; Colmenero, J. M.; Cercos, M.; Tadeo, F.; Agusti, J.; Alos, E.; Andres, F.; Soler, G.; Brumos, J.; Iglesias, D. J.; Götz, S.; Legaz, F.; Argout, X.; Courtois, B.; Ollitrault, P.; Dossat, C.; Wincker, P.; Morillon, R.; Talon, M. Analysis of 13000 unique citrus clusters associated with fruit quality, production and salinity tolerance. *BMC Genomics* **2007**, *25*, 8−31.
(26) Beidler, J. L.; Hilliard, P. R.; Rill, R. L. Ultrasensitive staining of nucleic acids with silver. *Anal. Biochem.* **1982**, *126*, 374−380.
(27) Bartley, G. E.; Scolnik, P. A. cDNA cloning, expression during development, and genome mapping of PSY2, a second tomato gene encoding phytoene synthase. *J. Biol. Chem.* **1993**, *268* (34), 25718−25721.
(28) Busch, M.; Seuter, A.; Hain, R. Functional analysis of the early steps of carotenoid biosynthesis in tobacco. *Plant Physiol.* **2002**, *128* (2), 439−453.
(29) Gallagher, C. E.; Matthews, P. D.; Li, F. Q.; Wurtzel, E. T. Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses. *Plant Physiol.* **2004**, *135* (3), 1776−1783.
(30) Lange, B. M.; Ghassemian, M. Genome organization in *Arabidopsis thaliana*: a survey for genes involved in isoprenoid and chlorophyll metabolism. *Plant Mol. Biol.* **2003**, *51*, 925−948.
(31) Romer, S.; Hugueney, P.; Bouvier, F.; Camara, B.; Kuntz, M. Expression of the genes encoding the early carotenoid biosynthetic enzymes in *Capsicum annuum*. *Biochem. Biophys. Res. Commun.* **1993**, *196* (3), 1414−1421.
(32) Matthews, P. D.; Luo, R. B.; Wurtzel, E. T. Maize phytoene desaturase and ζ-carotene desaturase catalyse a poly-Z desaturation pathway: implications for genetic engineering of carotenoid content among cereal crops. *J. Exp. Bot.* **2003**, *54* (391), 2215−2230.
(33) Ronen, G.; Carmel-Goren, L.; Zamir, D.; Hirschberg, J. An alternative pathway to beta-carotene formation in plant chromoplasts discovered by map-based cloning of beta and old-gold color mutations in tomato. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (20), 11102−11107.
(34) Galpaz, N.; Ronen, G.; Khalfa, Z.; Zamir, D.; Hirschberg, J. A chromoplast-specific carotenoid biosynthesis pathway is revealed by cloning of the tomato white-flower locus. *Plant Cell* **2006**, *18* (8), 1947−1960.
(35) Ronen, G.; Cohen, M.; Zamir, D.; Hirschberg, J. Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant Delta. *Plant J.* **1999**, *17*, 341−351.
(36) Kato, M.; Matsumoto, H.; Ikoma, Y.; Okuda, H.; Yano, M. The role of carotenoid cleavage dioxygenase in the regulation of carotenoid profiles during maturation in citrus fruit. *J. Exp. Bot.* **2006**, *57*, 2153−2164.

Received for review March 12, 2007. Revised manuscript received June 14, 2007. Accepted June 21, 2007. We thank the *Collectivité Territoriale de Corse* for granting this study.

JF070711H