Theor Appl Genet (2009) 118:653–669
DOI 10.1007/s00122-008-0927-8

ORIGINAL PAPER

The 1-deoxy-D-xylulose 5-phosphate synthase gene co-localizes with a major QTL affecting monoterpene content in grapevine

Juri Battilana · Laura Costantini · Francesco Emanuelli · Federica Sevini · Cinzia Segala · Sergio Moser · Riccardo Velasco · Giuseppe Versini · M. Stella Grando

Received: 11 July 2008 / Accepted: 24 October 2008 / Published online: 27 November 2008

Abstract Muscat flavor is a relevant trait both in winemaking and in fresh grape consumption. From a chemical point of view, it is strongly related to the accumulation of monoterpenes in berries. However, knowledge of the genetic mechanisms underlying its regulation is still limited. The objective of this study was to dissect the genetic determinism of aroma in grapevine by applying the analysis of quantitative trait loci (QTL) and the candidate gene (CG) approach. Two F_1 segregating progenies were evaluated through high-resolution gas chromatography–mass spectrometry (HRGC–MS) for the amounts of individual monoterpenes over 3 and 2 years. In the Italia × Big Perlon cross 34 CGs, chosen according to gene ontology (GO) terms, were placed on a complete map and tested for linkage with QTLs for linalool, nerol and geraniol levels. Two CGs mapped within a QTL for linalool content on LG 10. A third one co-localized with a major QTL for the level of the three monoterpenes on LG 5; this gene encodes 1-deoxy-D-xylulose 5-phosphate synthase (DXS), which is the first enzyme in the plastidial pathway of terpene biosynthesis. Depending on these findings, we report the first in silico analysis of grapevine *DXS* genes based on the whole genome sequence. Further research on the functional significance of these associations might help to understand the genetic control of Muscat flavor.

J. Battilana and L. Costantini equally contributed to the work.

Communicated by H. Nybom.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0927-8) contains supplementary material, which is available to authorized users.

J. Battilana · L. Costantini (✉) · F. Emanuelli · F. Sevini · C. Segala · R. Velasco · M. S. Grando
Department of Genetics and Molecular Biology,
IASMA Research Center, Fondazione Edmund Mach,
38010 San Michele all'Adige, Italy
e-mail: laura.costantini@iasma.it

S. Moser · G. Versini
Department of Agrifood Quality,
IASMA Research Center, Fondazione Edmund Mach,
38010 San Michele all'Adige, Italy

Introduction

Aroma plays an essential role in high-quality winemaking and is greatly appreciated for fresh grape consumption as well. Different combinations and concentrations of several fruit compounds define the so-called 'varietal aroma', which in turn affects wine 'character' (Ribéreau-Gayon et al. 2000). Numerous studies have revealed that the typical flavor of Muscat grape varieties is closely related to the presence of monoterpenes. Each molecule has distinct organoleptic features and interacts with the others, thus contributing to shape the final aroma of the mixture. In particular, linalool, geraniol, nerol, citronellol and α-terpineol are often described as the major aromatic determinants based on their high concentration in Muscat cultivars and their low olfactory perception threshold (Mateo and Jiménez 2000; Ribéreau-Gayon et al. 1975). Monoterpenes have been found both in grape leaves and berries as free volatiles, free polyhydroxylated molecules (polyols) and glycosidic derivatives of the two former types. Only the free volatile compounds make a direct contribution to the aroma, whereas free polyols and glycosidic derivatives constitute a reserve of odorless precursors, which generate flavor upon hydrolysis. The distribution of monoterpenes within the berry is not uniform, free geraniol and nerol are concentrated in the skin, whereas free linalool is more

evenly distributed, like the glycosidic forms (Strauss et al. 1986).

Monoterpenes belong to the terpenoid family, which is the largest and most diverse group of natural substances including both primary and secondary metabolites with a great variety of biological functions. Monoterpenoids (C10), sesquiterpenoids (C15), diterpenoids (C20) and triterpenoids (C30) are considered secondary metabolites of ecological significance since many of them mediate plant–environment interactions (Mahmoud and Croteau 2002). Plants synthesize the five-carbon universal building block of terpenoids, isopentenyl diphosphate (IPP), by two independent pathways having as intermediates mevalonic acid (MVA) in cytoplasm and 1-deoxy-D-xylulose 5-phosphate/2C-methyl-D-erythritol 4-phosphate (DOXP/MEP) in plastids (Lichtenthaler 1999). The biosynthesis of monoterpenes via the plastidial pathway was demonstrated both in grape leaves and berries (Luan and Wüst 2002). However, only a few genes and enzymes implicated in terpene biosynthesis have been identified and characterized so far in *Vitis vinifera*: a geranyl diphosphate synthase (Clastre et al. 1993), two sesquiterpene synthases (Lücker et al. 2004), a (−)-α terpineol synthase (Martin and Bohlmann 2004) and a potential carotenoid cleavage dioxygenase (Mathieu et al. 2005). On the other hand, having the complete sequence of the grape genome, several terpene synthases genes were recently predicted (Jaillon et al. 2007; Velasco et al. 2007).

Based on the analysis of several crosses, at least five complementary dominant genes plus a modifier gene have been proposed to be involved in the regulation of Muscat flavor in grapevine (Wagner 1967). With the exception of the preliminary experiment performed by Sevini et al. (2004), the only exhaustive investigation of the genetic basis of this trait was reported so far by Doligez et al. (2006). Based on 2 years of measures in a F_1 cross between a non-aromatic variety and Muscat of Hamburg, they identified major effect QTLs for the content of linalool, nerol and geraniol co-localizing on LG 5, as well as minor QTLs controlling the accumulation of linalool on LG 2 and of nerol and geraniol on LG 13. A QTL for Muscat score established by simple tasting was also located on LG 5.

Aiming to improve our knowledge about the genetic determinism of Muscat flavor in grape, we expanded the study of Sevini et al. (2004) and carried out QTL analysis on two F_1 mapping populations, which were analyzed for the content of linalool, nerol and geraniol during more than one season. The genomic regions controlling the phenotypic variability under study were further characterized by applying the candidate gene (CG) approach (Pflieger et al. 2001). This method has been recognized, especially in plants with large genomes and long generation times, as a promising alternative to the money and labor-consuming positional cloning (reviewed in Paran and Zamir 2003; Remington et al. 2001) in order to identify and isolate genes governing important traits (Morgante and Salamini 2003; Salvi and Tuberosa 2005). In the most widespread version the CG approach attempts to link, through mapping analysis, QTLs that are responsible for the studied variation with sequences that play a potential role in the measured phenotype or have a structural similarity to known genes, i.e. R-gene analogs. The availability of whole genome sequences and expressed sequence tag (EST) databases for important crops is accelerating the process of gene discovery. Grape can be placed among the best characterized plant species with respect to ESTs (da Silva et al. 2005) and the genomic sequence of *V. vinifera* has been recently made available (Jaillon et al. 2007; Velasco et al. 2007). In three cases we observed co-segregation between CG loci and QTLs for the content of the main aromatic monoterpene compounds. These results represent an important advancement towards marker-assisted selection for crop improvement.

Materials and methods

Plant material

Two mapping populations were considered in this study: Pop1 (163 F_1 individuals, intraspecific) and Pop2 (174 F_1 individuals, interspecific), which segregate for Muscat flavor and other traits. They derived respectively from the crosses Italia (*V. vinifera*) × Big Perlon (*V. vinifera*) and Moscato Bianco (*V. vinifera*) × *V. riparia* (accession Wr 63 from the IASMA Ampelographic Collection). Pop1 and Pop2 have been grown at the Experimental Station of the University of Bari and of IASMA (Italy), respectively.

Good quality DNA was extracted from young leaves following the protocol described in Grando et al. (2003).

Candidate gene selection

CG choice was carried out in two steps. A first set of 19 *V. vinifera* ESTs was directly selected from the public database TGI (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=grape, Release 3.1) by using keywords related to terpenoid metabolism (mevalonate and non-mevalonate pathways of IPP biosynthesis, monoterpenoid and diterpenoid metabolism). A second set of 51846 ESTs was extracted from the TGI database (Release 4.0) based on 650 gene ontology (GO) terms associated with berry composition and ripening. ESTs were checked for quality and assembled with cd-hit software (Li and Godzik 2006) in order to remove redundant sequences (identity >80%). Functional characterization was available for 6637 ESTs after BLASTX alignment (E value $\leq 1 \times 10^{-6}$) against a

UniProt (http://www.expasy.uniprot.org/) partition containing only proteins annotated by one of the selected GO terms. Ninety-three sequences with a potential role in berry flavor (106 GO terms) were extracted from this characterized pool, clustered with CAP3 software (Huang and Madan 1999; terminal alignment ≥40 bases, identity ≥90%) and checked for belonging to TC (tentative consensus) sequences in TGI database. This procedure finally resulted in the selection of 53 putative CGs.

Amplification

Specific primers were designed for each EST by using the software Primer Express (Applied Biosystems, Foster City, CA, USA). In most cases, the polymerase chain reaction (PCR) mixture (12.5 µl) contained 5–10 ng of genomic DNA, 1.25 µl of 10× PCR buffer (Qiagen, Valencia, CA, USA; 1.5 mM of $MgCl_2$), 40 µM of each dNTP, 0.6 µM of each primer and 0.5 unit of HotStarTaq polymerase (Qiagen). Amplification was carried out by using a GeneAmp PCR System 9700 (Perkin–Elmer, Norwalk, CT, USA) and a touchdown protocol (Don et al. 1991). Primer sequences and detailed amplification conditions are reported in Table 1SA.

Sequencing

Two to four nanograms of amplified DNA were employed for every 100 bp to be sequenced in both directions. PCR products were purified with ExoSapIT (Amersham Pharmacia Biotech, Uppsala, Sweden) and sequenced with the Big Dye Terminator v 3.1 Cycle Sequencing Kit (Applied Biosystems) in a GeneAmp PCR System 9700. After precipitation, the sequencing products were mixed with 15 µl of HiDi™ formamide and subjected to capillary electrophoresis in an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems). The resulting data were analyzed with the softwares Sequencing analysis v 3.7 (Applied Biosystems) and ChromasPro v 1.3 (http://www.technelysium.com.au).

Marker development and analysis

Markers for the CG sequences were generated by SSCP (single-strand conformational polymorphism) and minisequencing methods.

SSCP

The gel was prepared between two glasses by mixing 7.5 ml of MDE gel solution (FCM BioProducts, Rockland, ME, USA), 17.7 ml of H_2O, 3 ml of glycerol 50% v/v, 8 ml of TBE 10×, 18.8 µl of TEMED and 150 µl of APS 10% w/v; 5 µl of PCR products were added with 7 µl of SSCP loading buffer and denatured for 3 min at 95°C. The electrophoresis run was performed for 15–16 h at a constant voltage of 135 V. The visualization of SSCP polymorphisms followed Sanguinetti et al. (1994).

Minisequencing with SnaPshot technique

The primers (Table 1SB) were designed with the software Primer Express (Applied Biosystems) and added with a variable length tail in order to multiplex the minisequencing products. PCR fragments were purified with ExoSapIT. The minisequencing reaction was performed in a GeneAmp PCR System 9700 following the recommendations of the manufacturer (Applied Biosystems). The minisequencing products (0.5 µl) were mixed with 9.45 µl of HiDi™ formamide and 0.05 µl of GeneScan-120 LIZ size standard (Applied Biosystems) and run in an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems). The resulting data were analyzed with the software GeneScanTM v 3.7 (Applied Biosystems).

Map construction

Parental and consensus linkage maps were constructed in Pop1 as described by Costantini et al. (2008). The same method was applied to Pop2, for which linkage maps were previously reported by Grando et al. (2003). An updated mapping set was used here based on 174 F_1 individuals genotyped at 119 loci with SSR (Simple Sequence Repeat), EST and RGA (Resistance Gene Analog) markers.

Metabolite profiling and quantification

Berries were collected from each genotype when their sugar content was approximately 16°Brix. This value was established in the framework of the EU-Project MASTER (Marker Assisted Selection for TablE gRape) in order to compare QTLs detected for the same trait in different genetic backgrounds, including the progeny described by Doligez et al. (2006). With the aim of minimizing the great variability among the different berries of the same cluster as well as among the berries of different clusters, sugar concentrations from 3 randomly taken berries per cluster and 2–3 representative clusters per genotype were averaged.

Pop1

The content of free linalool, geraniol and nerol was evaluated in three consecutive years (2002, 2003, 2004) in Pop1 and its parental varieties by using SPME (solid phase micro-extraction) method according to Doligez et al.

(2006). One hundred grams of frozen berries were put in a glass jar (500 ml) to which were added 60 g of $(NH_4)_2$ SO_4 in 50 ml of MilliQ water (Millipore, Bedford, MD, USA) and deuterated standards: 115 μg kg^{-1} of linalool-d5; 93.7 μg kg^{-1} of geraniol-d2 and 65.6 μg kg^{-1} of nerol-d2, these latter compounds as a mixture. All internal standards were supplied by INRA of Montpellier (France) and were synthesized as described in Doligez et al. (2006). After homogenization, 40 ml of this solution and a magnetic stir were placed in a 50 ml vial, sealed and equilibrated for 10 min at 30°C. A polyacrylate (PA)-coated fiber (85 μm; PA Codex: 57304, Supelco, Bellefonte, USA) was exposed in the headspace for 30 min and then desorbed at 250°C for 5 min in a GC-injector by working in splitless mode. High-resolution gas chromatography-mass spectrometry was performed using a PerkinElmer gas chromatograph with a TurbomassGold Mass Spectrometer equipped with a DBWax fused silica column (60 m × 0.32 mm I.D., 0.5 μm film thickness). Helium with a flow rate of 4 ml min^{-1} was used as carrier gas. The oven program was as follows: 50°C, 4°C min^{-1} to 190°C, 10°C min^{-1} to 220°C, 220°C for 10 min. Detector temperature was set to 220°C. In 2002 a MS detector was employed; both total ion chromatograms (TIC) and selection ion recording (SIR) profiles were obtained. In 2003 and 2004 monoterpenic compounds, that were well detectable also by a flame ionization detector (FID), were quantified by GC-FID. Calibration curve was achieved using a neutral grape (Alphonse Lavalle); berries of similar size and color were added with linalool (22.4 μg ml^{-1}), nerol (22.0 μg ml^{-1}) and geraniol (22.2 μg ml^{-1}) (Aldrich, >97%; checked as for the purity of each compound in respect to the other two) dissolved in MilliQ water and ethanol (1:1, v/v) to get a final concentration in the range of 0 to about 300 μg kg^{-1} (7 scores) but also to control the linearity of the curve until 1.5 mg kg^{-1} (1 score).

Pop2

In Pop2 and its female parent (Moscato Bianco) linalool, nerol and geraniol were quantified in 2 years (2002 and 2004) in their free and bound forms by using SPE (solid phase extraction) method. According to Versini et al. (1988) and Günata et al. (1989), the juice obtained by crushing 100 g of frozen berries was submitted to a XAD-2 resin (particle size 0.2–0.25 mm). The two adsorbed forms were separately eluted, the bound ones were hydrolyzed as aglycons after reaction with AR 2000 enzyme (Gist Brocades) and both fractions were enriched in an organic solvent solution. The analytical methodology was the same as in the cited papers; the quantification by HRGC–MS analysis was referred to the internal standard 1-heptanol with response factor (RF) = 1.

Data and QTL analysis

Metabolic data were tested for normality with the Kolmorogov–Smirnov test. When not normally distributed, they were ln-transformed. Correlations between years within traits and between traits within years were determined with the non-parametric Spearman rank correlation test. Both tests were implemented in the software SPSS v 11.0.

QTL detection was carried out with the software MapQTL v 4.0 (Van Ooijen et al. 2002) on the four parental varieties by separately analyzing every year of phenotypic evaluation. Non-parametric Kruskal–Wallis (KW) rank-sum test and interval mapping (Lander and Botstein 1989) methods were applied. In multiple QTL mapping (MQM) the markers closest to the QTLs found with either simple interval mapping (SIM) or KW were employed as cofactors. For SIM LOD thresholds at 0.95 significance were determined through 1,000 permutations at both chromosome-wide and genome-wide level (Churchill and Doerge 1994); for MQM the same values as for SIM were used, since this function is not available in MapQTL. One-LOD support interval was adopted for the confidence interval.

In silico analysis

The total number of genes underlying QTLs on LGs 5 and 10 was estimated through the procedure described in Velasco et al. (2007). Based on the results obtained from QTL analysis in Pop1, gene prediction was performed at three different magnification levels on the genomic sequence of Pinot Noir clone ENTAV115: (1) within the region encompassed by the two closest SSR markers to the LOD peak (on LG 10) or between the only microsatellite mapped on one side of the LOD peak and the end of the genomic metacontig on the other side (on LG 5), (2) within the region where the LOD values were above the genome-wide LOD threshold ($P < 0.05$), and (3) in the 1-LOD interval. Mean values across traits and years were calculated for the length of these regions.

In order to investigate *DXS* gene copy number and organization, the genomic sequence of grapevine (Jaillon et al. 2007; Velasco et al. 2007) was screened with BLASTN algorithm by using as query sequence the TGI tentative consensus on which marker DXS was developed (TC56417). Nucleotide and amino acid sequence analysis of the aligning genomic contigs was performed with online bioinformatic tools: the softwares FGENESH v 2.5 (Salamov and Solovyev 2000) for prediction of potential genes, FGENESH+ (Salamov and Solovyev 2000) and GeneWise v 2.0 (Birney et al. 2004) for analysis of gene structure, Predotar v 1.03 (Small et al. 2004) and SignalP v 3.0 (Bendtsen et al. 2004) for prediction of protein subcellular localization. Protein organization into functional domains

was derived from the conserved domain database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml).

Results

Marker development and analysis

Specific primer pairs were designed for 53 CG sequences and tested on the two parents of Pop1, Italia and Big Perlon; 46 of them produced a unique and reproducible band. They were initially assessed for single-strand conformational polymorphisms in a six individual-subset of the Italia × Big Perlon progeny. Segregating polymorphisms were detected in 16 cases. The remaining 30 sequences produced monomorphic profiles and were thus sequenced in both parents in order to identify single nucleotide polymorphisms (SNPs). All of them were confirmed to correspond to the originally selected genes through BLASTN alignment against the database TGI; 21 sequences contained SNPs, which were successfully detected in 18 cases with minisequencing. Altogether, 34 EST-derived markers were developed (Table 1).

Map construction

The dataset of CG markers was used to genotype Pop1, along with the other marker types described in Costantini et al. (2008). By applying a similar procedure to that described in Costantini et al. (2008), linkage maps were also obtained for Moscato Bianco and *V. riparia* with the main goal to validate QTL results achieved in Pop1.

Four CG markers, which were monomorphic in Pop1 (ACTRANS, GAI, G10H and PMVAK), were analyzed in Pop2 and located for synteny in Italia x Big Perlon maps. The DXS marker was mapped in both populations.

Metabolite profiling and quantification

Linalool, nerol and geraniol content showed a continuous variation, which is typical of quantitative traits, and a transgressive segregation in both progenies (Fig. 1).

Data and QTL analysis

The one-sample Kolmogorov–Smirnov test indicated departures from normality ($P < 0.05$) for the content of linalool, nerol and geraniol, which was highly skewed towards low values in both populations. Data were ln-transformed in the attempt to achieve normality, which is preferred for the QTL analysis based on interval mapping. However, some ln-transformed distributions were still significantly different from normal at $P < 0.05$ (Pop1: ln-linalool 2002, ln-nerol 2003 and 2004, ln-geraniol 2004; Pop2: free ln-nerol 2004, bound ln-nerol 2002 and 2004, free and bound ln-geraniol 2002).

In all years and both populations significant ($P < 0.01$) Spearman rank-order correlations were found among the three monoterpenes (except for bound geraniol in Pop2, 2004), and was always stronger between nerol and geraniol (Tables 2S, 3S and 4S). In Pop1 they were significantly ($P < 0.01$) correlated only between 2002 and 2004, probably due to the uncommonly high temperatures measured during summer 2003; in Pop2 correlation between years was significant ($P < 0.05$) only for free nerol content (Tables 2S, 3S and 4S).

Several QTLs were suggested by simple interval mapping (data not shown), but only some of them were confirmed with MQM and were significant at both the chromosome-wide and the genome-wide thresholds (Table 2). In the aromatic parent of Pop1 [Table 2(A)] a major QTL for the amount of linalool, nerol and geraniol was detected on LG 5 in 3 years. This QTL explained 26–36% of the total variance of linalool, 69–83% of nerol, and 69–84% of geraniol content. One additional QTL for linalool content (11–36% of explained variance) was found on LG 10 in 3 years. Because ratios of compound concentrations are more robust than individual metabolite levels (Morreel et al. 2006), we also calculated the linalool/nerol, linalool/geraniol and nerol/geraniol ratios. The QTLs explaining a high percentage of the total variance for linalool/nerol (44–57%) and linalool/geraniol (23–52%) ratios were located on LG 10 in 3 years in the same position where the QTL for linalool content was identified [Table 3(A)]. A QTL for nerol/geraniol ratio was detected on LG 7 in 2 years, which explained 8–20% of the total variance [Table 3(A)]. The QTLs on LGs 5 [Table 2(A)] and 10 [Tables 2(A), 3(A)] were unambiguously confirmed by KW analysis and were recognized also in Big Perlon, the non-aromatic parent of Pop1 (data not shown).

In the aromatic parent of Pop2 [Table 2(B)] a major QTL for the content of the three monoterpenes, in their free and bound forms, was detected in both years on LG 5, as in Pop1. This QTL explained 64–72% of the total variance of linalool, 84–93% of nerol and 83–90% of geraniol content. The major QTL on LG 5 was confirmed by KW analysis and was found also in *V. riparia*, the non-aromatic parent of Pop2 (data not shown). One additional QTL for the content of free linalool (32% of explained variance) was identified on LG 2 of Moscato Bianco map in 2002. This QTL was not supported by KW analysis [Table 2(B)]. The ratio between free linalool and nerol or geraniol, and the ratio between free nerol and geraniol turned out to be regulated in both years, respectively, by the major QTL on LG 5 (35–55% of explained variance) and by a QTL on LG 6 (28–53% of explained variance).

Table 1 List of the CG markers positioned onto Italia and Big Perlon maps

Marker	LG	cM	TC	Annotation	Method
(A) Terpenoid metabolism					
G10H[a]	2	47.6	TC9597[b]	Geraniol 10-hydroxylase	SSCP
PMVAK[a]	2	50.6	TC34499	Phosphomevalonate kinase	SSCP
YGBB	2	92.0	TC58474	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, chloroplast precursor	SNP
ISPH	3	20.9	TC14300[b]	LYTB-like protein precursor	SSCP
GGPP-S	4	46.7	TC51973	Geranylgeranyl pyrophosphate synthetase, chloroplast precursor	SSCP
DXS	5	2.7	TC56417[c]	1-Deoxy-D-xylulose 5-phosphate synthase	SNP
CDP-ME	6	14.4	TC14915[b]	4-Diphosphocytidyl-2-C-methyl-D-erythritol kinase, chloroplast precursor	SSCP
CRTISO-sscp[d]	8	73.2	TC62553	Carotenoid isomerase, chloroplast precursor	SSCP
ACTRANS[a]	12	34.4	TC52127	Acetoacetyl-CoA thiolase	SSCP
B-diox-II-sscp	13	0.0	TC62780	9,10[9′, 10′]Carotenoid cleavage dioxygenase	SSCP
IPPISOM	14	6.2	TC32368[b]	Isopentenyl-diphosphate delta-isomerase 1	SSCP
HMGS	14	26.2	TC68763	Hydroxymethylglutaryl coenzymeA synthase	SSCP
Gib20ox	16	22.6	TC11917[b]	Gibberellin 20-oxidase	SSCP
DXR	17	39.9	TC64939	1-Deoxy-D-xylulose 5-phosphate reductoisomerase precursor	SSCP
Gib2ox	19	33.8	TC60763	Gibberellin 2-oxidase	SSCP
(B) Aromatic amino acid metabolism					
DHAP-S	2	56.7	TC59396	3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase	SNP
DHAP-S-p	7	54.3	TC57642	3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase precursor	SNP
FAH1	10	91.3	TC53032	Fumarylacetoacetase	SNP
FAH	10	93.5	TC56398	Fumarylacetoacetase	SNP
HPD1-sscp[d]	12	1.2	TC58798[e]	4-Hydroxyphenylpyruvate dioxygenase	SSCP
HPD	12	5.4	TC58798[e]	4-Hydroxyphenylpyruvate dioxygenase	SNP
PHEA-sscp[d]	12	23.1	TC55648	Putative P-protein: chorismate mutase prephenate dehydratase	SSCP
IGPS	12	35.4	TC64485	Indole-3-glycerol phosphate synthase	SNP
PAL	13	43.0	TC60180	Phenylalanine ammonia lyase	SNP
PAI1	14	33.7	TC60070	NADH-ubiquinone oxidoreductase	SNP
trpB	19	60.1	TC62261	Tryptophan synthase beta subunit	SNP
TAT	19	75.0	TC53133	Tyrosine aminotransferase	SNP
HGOb-sscp	19	79.2	TC66094	Homogentisate 1,2-dioxygenase	SSCP
HGOa	19	81.6	CF518271	Homogentisate 1,2-dioxygenase	SNP
(C) Stress and gene regulation					
AIP	3	0.0	TC53877	Putative auxin-regulated protein	SNP
pepA1	8	33.5	TC66643	Putative CND41, chloroplast nucleoid DNA binding protein Pepsin A	SNP
cnd41	10	84.0	TC57402	Putative chloroplast nucleoid DNA-binding protein cnd41	SNP
conG-p	14	3.6	TC51730	Putative CND41 Conglutin gamma precursor	SNP
pDNAbP	15	47.9	TC67994	Chloroplast nucleoid DNA binding protein putative	SNP

LG linkage group, *cM* marker position on the Italia × Big Perlon consensus map (Costantini et al. 2008), *TC* tentative consensus, *SSCP* single strand conformational polymorphism, *SNP* single nucleotide polymorphism

[a] Monomorphic markers in Pop1 were mapped for synteny after being analyzed in Pop2

[b] These TCs were split into two or more tentative consensus in TGI Database (Release 5.0, 21 June 2006, http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=grape)

[c] DXS was mapped also in Pop2 as SSCP marker

[d] A SNP marker (SnaPshot technique) within the same TC was additionally mapped in Costantini et al. (2008)

[e] These molecular markers were developed from two distinct ESTs belonging to the same TC

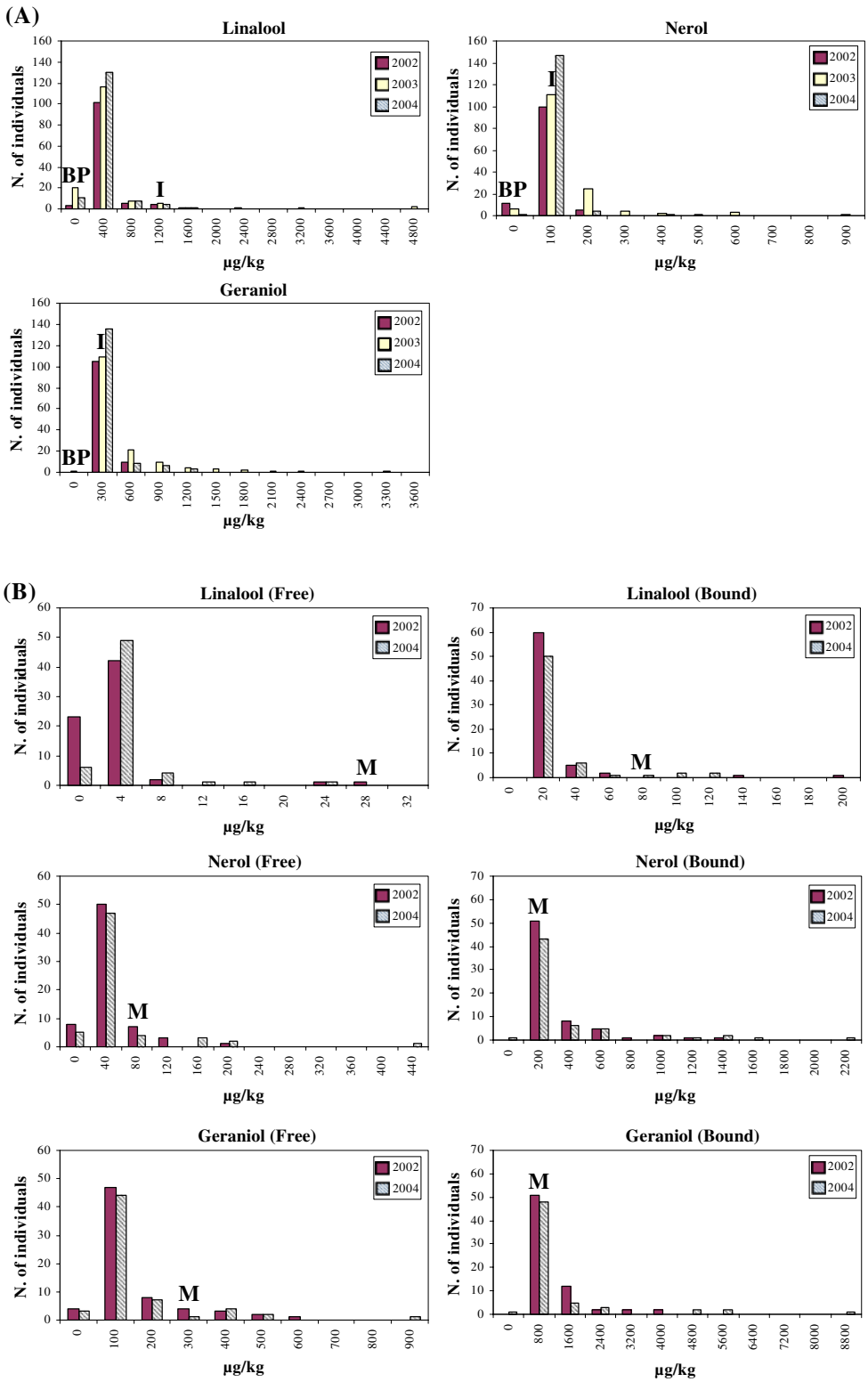


Fig. 1 Distribution of monoterpene content in Pop1 (**a**) and Pop2 (**b**). *I* Italia, *BP* Big Perlon, *M* Moscato Bianco

Both QTLs were confirmed by KW analysis [Table 3(B)]. Finally, a QTL controlling the ratio between free linalool and nerol (35% of explained variance) was detected on LG 10 in 2002 [Table 3(B)].

CG markers co-localizing with QTLs

In Pop1 interval mapping and KW analysis revealed a significant association between three CG markers and QTLs

Table 2 QTLs identified in Italia (A) and Moscato Bianco (B) for the content of linalool, nerol and geraniol

Year	Compound	LG	Marker	Position (cM)	1-LOD interval (cM)	LOD peak	LOD threshold 0.95		% Var expl	KW
							Chromosome-wide	Genome-wide		
(A)										
2002	Free linalool	5	DXS	12.8	12.3–13.3	10.2	3.1	8.7	26.3	*******
	Free linalool	10	FAH1	86.1/88.2	81.0–89.0	13.1	6.7	8.7	36.4	*****
	Free nerol	5	DXS	12.8	12.2–12.9	28.1	2.9	5.9	68.7	*******
	Free geraniol	5	DXS	12.8	12.5–13.0	33.3	3.0	9.2	75.5	*******
2003	Free linalool	5	DXS	12.8	11.0–14.0	13.0	2.9	8.0	31.1	*******
	Free linalool	10	cnd41	81.1	81.0–87.2	15.3	6.0	8.0	34.8	*******
	Free nerol	5	mCAGeATG16	10.9	11.0–14.0	33.3	3.0	12.8	68.8	*******
	Free geraniol	5	DXS	12.8	11.0–13.0	37.7	3.0	10.8	68.9	*******
2004	Free linalool	5	DXS	12.8	12.1–13.0	15.5	3.1	5.5	36.4	*******
	Free linalool	10	FAH1	86.1/88.2	74.5–86.5	18.1	4.0	5.5	10.6	*******
	Free nerol	5	DXS	12.8	12.5–12.9	41.8	3.4	12.6	82.9	*******
	Free geraniol	5	DXS	12.8	12.4–12.9	42.8	3.1	12.6	83.7	*******
(B)										
2002	Free linalool	2	VMC3B10	19.3	17.8–19.8	5.8	2.7	4.7	31.6	–
	Free linalool	5	VrZAG47	73.7/68.7	69.0–79.0	10.0	3.0	4.7	63.6	*******
	Free nerol	5	VrZAG47	73.7/68.7	72.0–83.0	17.7	5.6	8.2	84.1	*******
	Free geraniol	5	VrZAG47	73.7/68.7	72.0–79.0	25.3	11.4	14.4	90.4	*******
2004	Free linalool	5	VrZAG47	78.7/68.7	73.7–83.7	10.2	3.2	5.0	66.8	*******
	Free nerol	5	VrZAG47	78.7/68.7	73.7–83.7	22.3	7.9	11.2	93.0	*******
	Free geraniol	5	VrZAG47	73.7/68.7	71.0–80.0	18.0	5.4	7.9	84.3	*******
2002	Bound linalool	5	VrZAG47	78.7/68.7	73.0–83.7	14.7	3.5	4.9	72.4	*******
	Bound nerol	5	VrZAG47	73.7/68.7	73.0–81.0	27.3	14.3	15.7	91.0	*******
	Bound geraniol	5	VrZAG47	73.7/68.7	73.4–81.3	24.6	12.2	12.8	89.0	*******
2004	Bound linalool	5	VrZAG47	78.7/68.7	73.7–83.7	12.2	3.3	4.7	72.2	*******
	Bound nerol	5	VrZAG47	78.7/68.7	73.7–83.7	26.3	13.0	14.5	92.2	*******
	Bound geraniol	5	VrZAG47	78.7/68.7	73.7–83.7	17.5	5.4	6.8	82.6	*******

LG linkage group, *Marker* marker nearest to the QTL position, *Position* QTL position (two values were reported when there was no coincidence between the LOD peak and the marker; in this case the first value refers to the LOD peak and the second one to the marker), *LOD peak* LOD (log of odds) value at QTL position, *LOD threshold* chromosome-wide and genome-wide LOD threshold ($P < 0.05$), *% Var expl* proportion of the total phenotypic variance explained by the QTL, *KW* = Kruskal–Wallis significance level, given by the *P* value (* 0.1, ** 0.05, *** 0.01; **** 0.005; ***** 0.001; ****** 0.0005; ******* 0.0001)

for monoterpene concentration [Table 2(A) and Fig. 2a). A DXS marker (at 12.8 cM on LG 5 of Italia) was found to be linked to the content of linalool, nerol and geraniol. Another relevant association emerged between the markers cnd41/FAH1 (respectively at 81.1 and 88.2 cM on LG 10 of Italia) and linalool content. The cnd41 marker was also related to linalool/nerol and linalool/geraniol ratios [Table 3(A)]. A linkage between DXS and monoterpene content was observed in Pop2 as well. This marker turned out homozygous in Moscato Bianco and heterozygous in *V. riparia*, as a consequence it could be mapped only in the non-aromatic parent of Pop2 (at 64.4 cM on LG 5 of *V. riparia*). Monoterpene content was found to be significantly associated with DXS in *V. riparia* (data not shown) and with VrZAG47 in Moscato Bianco [Tables 2(B), 3(B)]. In this last case the LOD peak of the QTL for free and bound compounds did not coincide with VrZAG47 (at 68.7 cM on LG 5 of Moscato Bianco), and it was detected at a position (between 73.7 and 78.7 cM) where DXS could be expected based on the Moscato Bianco × *V. riparia* and Italia × Big Perlon consensus maps (Fig. 2b).

In silico analysis

The in silico analysis of Pinot Noir genomic sequence (Velasco et al. 2007) allowed us to estimate the physical length of the QTL intervals on LGs 5 and 10 of the Italia map and to predict the number of genes that are within. The molecular markers reported in Table 4 represent the closest loci to these QTLs. They include the three CGs underlying

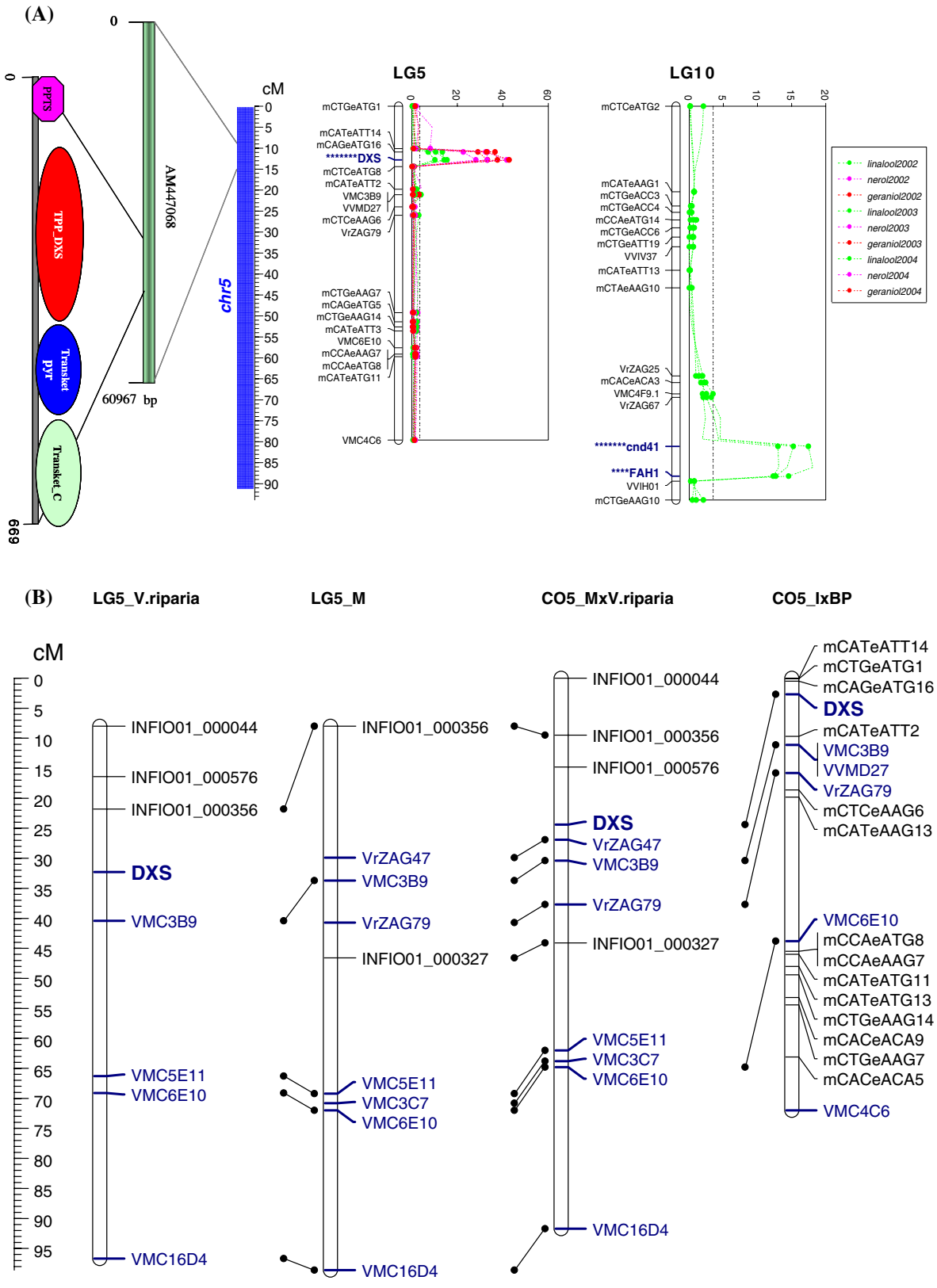
Table 3 QTLs identified in Italia (A) and Moscato Bianco (B) for the ratio between free linalool, nerol and geraniol

Year	Compound	LG	Marker	Position (cM)	1-LOD interval (cM)	LOD peak	LOD threshold 0.95		% Var expl	KW
							Chromosome-wide	Genome-wide		
(A)										
2002	Linalool/Nerol	2	VMC7G3	8.7/3.7	0.0–18.8	6.1	2.7	4.6	15.6	**
	Linalool/Nerol	2	VVIB01	44.5/46.8	39.0–46.8	4.8	2.7	4.6	11.2	–
	Linalool/Nerol	10	cnd41	79.4/81.1	74.0–82.0	16.4	3.2	4.6	43.8	*******
	Linalool/Geraniol	10	cnd41	79.4/81.1	72.0–82.0	15.4	3.0	4.4	47.0	*******
	Nerol/Geraniol	1	VVIS21	49.7/44.7	41.8–55.0	4.6	2.9	4.4	11.6	–
	Nerol/Geraniol	6	VMC4H5	19.6	19.0–20.0	6.0	2.6	4.4	15.8	******
	Nerol/Geraniol	7	VMC1A12	61.3/63.3	56.3–69.0	6.8	2.7	4.4	20.2	*
2003	Linalool/Nerol	10	cnd41	79.4/81.1	74.4–80.0	18.1	3.6	5.5	56.9	*******
	Linalool/Geraniol	10	cnd41	79.4/81.1	72.0–81.0	16.9	5.2	7.2	52.4	*******
2004	Linalool/Nerol	2	mCTCeATG3	0.0	0.0–2.0	5.2	2.7	4.3	12.0	**
	Linalool/Nerol	10	cnd41	79.4/81.1	75.5–86.5	17.6	3.0	4.3	46.9	*******
	Linalool/Geraniol	10	cnd41	81.1	75.0–86.5	16.5	3.1	4.5	23.4	*******
	Nerol/Geraniol	2	VMC7G3	13.7/3.7	0.0–18.8	6.3	2.6	4.5	15.4	******
	Nerol/Geraniol	7	VMC1A12	63.3	58.0–80.0	6.1	3.0	4.5	8.0	***
	Nerol/Geraniol	10	mCTCeATG2	0.0	0.0–5.3	5.2	3.0	4.5	48.7	–
(B)										
2002	Linalool/Nerol	5	VrZAG47	73.7/68.7	71.0–84.0	5.5	3.3	3.7	35.2	****
	Linalool/Nerol	10	VrZAG67	28.1/32.9	14.0–32.9	4.7	2.5	3.7	35.4	–
	Linalool/Geraniol	5	VrZAG47	68.7	67.0–79.0	6.3	3.2	4.5	50.0	*******
	Linalool/Geraniol	12	VMC8G6	40.3/47.5	31.0–47.5	5.5	2.8	4.5	34.1	****
	Nerol/Geraniol	6	VMC5C5	30.6	26.5–31.3	6.5	2.6	4.1	27.8	******
	Nerol/Geraniol	15	4H04	51.5	45.2–51.5	5.5	2.5	4.1	33.4	–
2004	Linalool/Nerol	5	VrZAG47	68.7	67.5–79.5	8.5	3.2	5.2	55.1	*******
	Linalool/Geraniol	5	VrZAG47	68.7	64.9–79.5	7.9	3.0	4.4	50.6	*******
	Nerol/Geraniol	6	VMC5C5	28.2/30.6	26.2–29.9	7.0	2.6	4.0	52.7	*******

LG linkage group, *Marker* marker nearest to the QTL position, *Position* QTL position (two values were reported when there was no coincidence between the LOD peak and the marker; in this case the first value refers to the LOD peak and the second one to the marker), *LOD peak* LOD (log of odds) value at QTL position, *LOD threshold* chromosome-wide and genome-wide LOD threshold ($P < 0.05$), *% Var expl* proportion of the total phenotypic variance explained by the QTL, *KW* Kruskal–Wallis significance level, given by the *P* value (* 0.1, ** 0.05, *** 0.01; **** 0.005; ***** 0.001; ****** 0.0005; ******* 0.0001)

the LOD peak and one microsatellite on each side of the LOD peak (only one side of the QTL on LG 5 was covered by microsatellites). On LG 10 two distinct metacontigs were found to house the four indicated markers. On LG 5 the average 1.3 cM-long 1-LOD interval corresponded to 267 kb, which was predicted to contain 27 genes. On LG 10 the average 8.7 cM-long 1-LOD interval was estimated to correspond to ca. 1,540 kb spanning two distinct metacontigs (Table 4). This estimation was made from the genetic distance according to Velasco et al. (2007) and was unavoidable due to the presence of a gap between the two metacontigs containing VVIH01 and VrZAG67. Around 200 genes might exist in the 1-LOD interval on LG 10.

The expressed sequence on which the DXS marker was developed (TC56417) was blasted against the nucleotide collection of the NCBI database (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi). This sequence aligned completely with locus AM447068 and partially with nine additional loci (AM431242, AM431698, AM441419, AM449097, AM461604, AM471792, AM480310, AM487977, and AM488379). Contig AM447068 was 60967 bp long. The software FGENESH predicted the existence of a gene containing 10 exons. The deduced grape protein consisted of 669 amino acids and showed high identity with the 1-deoxy-D-xylulose 5-phosphate synthase (DXS) of other plant species (the highest identity was 86% with *Pueraria montana*). Consultation of the conserved domain database showed that proteins of the DXS family contain three functional domains: a thiamine pyrophosphate (TPP)-binding module, a transketolase, pyridine binding domain and the C-terminal domain of transketolase, which has been proposed as a regulatory molecule binding site (Fig. 2a). The nine contigs that partially



aligned to TC58417 strongly matched with bases 2,705–2,825 and 3,290–4,569 of the *DXS* sequence. In four cases (AM431242, AM441419, AM461604 and AM488379) the protein encoded by the predicted gene showed high identity with only two of the three DXS functional domains. These contigs were no longer considered as part of our study

◀ **Fig. 2** CG markers co-localizing with QTLs for monoterpene content. **a** LGs 5 and 10 refer to the Italia map; AM447068 is the Pinot Noir genomic contig aligning with the DXS marker; the domain architecture of the predicted DXS1 protein is shown on the left. *PPTS* putative plastid target sequence, *TPP_DXS* thiamine pyrophosphate (TPP) family, DXS subfamily, TPP-binding module, *Transket_pyr* transketolase, pyridine binding domain, *Transket_C* transketolase C-terminal domain. **b** Representation of LG 5 from the maps of *V. riparia*, Moscato Bianco (M), Moscato Bianco × *V. riparia* and Italia (I) × Big Perlon (BP). INFIO01_000327, INFIO01_000356 and INFIO01_000044 are SSCP markers which were developed on EST sequences (http://www.ncbi.nlm.nih.gov), whereas INFIO01_000576 corresponds to a SSR within a EST

because they did not contain a putative *DXS* sequence. The DXS sequence from contig AM447068 proved to be similar to DXS class 1 (DXS1), while putative DXS proteins from contigs AM449097, AM471792, AM480310 and AM487977 were highly similar to DXS class 2. Finally, the DXS sequence obtained from contig AM431698 aligned with At5DXS of *Arabidopsis thaliana* (NP196699) and DXS of *Medicago truncatula* (ABE78977), which belong to class 3 (Table 5). Figure 3 shows the position of exons and introns in the six *DXS* sequences, as predicted by the softwares FGENESH+ and Genewise. A plastid targeting sequence is suggested for DXS1, DXS2 B, DXS2 C and DXS2 D (Table 5S).

Discussion

In this work we applied the CG method, combined with QTL analysis, to the identification of gene sequences putatively involved in the regulation of Muscat flavor in grapevine.

Table 5 Description of *DXS* gene family in *Vitis*

Locus name[a]		Gene	LG
Velasco et al. (2007)	Jaillon et al. (2007)		
AM447068	CAAP02001192	*DXS1*	5
AM471792	CAAP02003958	*DXS2 A*	7
AM487977	CAAP02003415	*DXS2 B*	15
AM480310	CAAP02001170	*DXS2 C*	11
AM449097	CAAP02004722	*DXS2 D*	15
AM431698	CAAP02001517	*DXS3*	4

LG linkage group

[a] From the Whole-Genome Shotgun reads (WGS) database (http://www.ncbi.nlm.nih.gov)

Candidate gene mapping

The choice of ESTs for CGs was based on keywords and GO terms related to Muscat aroma. This strategy, followed by EST quality checking, functional characterization and clustering, proved to be suitable for picking out a manageable number of gene sequences. The 53 CGs selected are involved in a few biological processes: 26 in the metabolism of terpenoids, 20 in the metabolism of aromatic amino acids and seven in stress or regulation of chloroplast gene expression; 87% (46/53) of these sequences could be amplified. The CG marker development was based on the detection of molecular polymorphisms through SSCP and minisequencing methods. Both techniques were successful, but the second one turned out to be more powerful in terms of yield and reproducibility.

Table 4 Genetical and physical distances of QTL intervals, number of predicted genes underlying QTLs found in Pop1

Molecular markers	Total length of metacontig (kb)	Total length of metacontig under study (kb)	Gene predictions
Chromosome 5			
DXS1	5,200	1,087	120
VMC3B9			
QTL interval	Interval above genome wide LOD threshold (cM/kb)	2.2/413	37
	1-LOD interval (cM/kb)	1.3/267	27
Chromosome 10			
VVIH01	3,020	3,020	388
FAH1			
cnd41			
VrZAG67	3,311	970	98
QTL interval	Interval above genome wide LOD threshold (cM/kb)	11/2000	250
	1-LOD interval (cM/kb)	8.7/1540	200

Molecular markers represent the closest loci to the QTLs detected on LGs 5 and 10. Total length of metacontig corresponds to the entire metacontig containing these loci, whereas total length of metacontig under study refers to the region encompassed by the two microsatellites surrounding the QTL. Additional analyzed regions are the QTL interval where the LOD values are above the 0.95 genome-wide LOD threshold and the 1-LOD interval

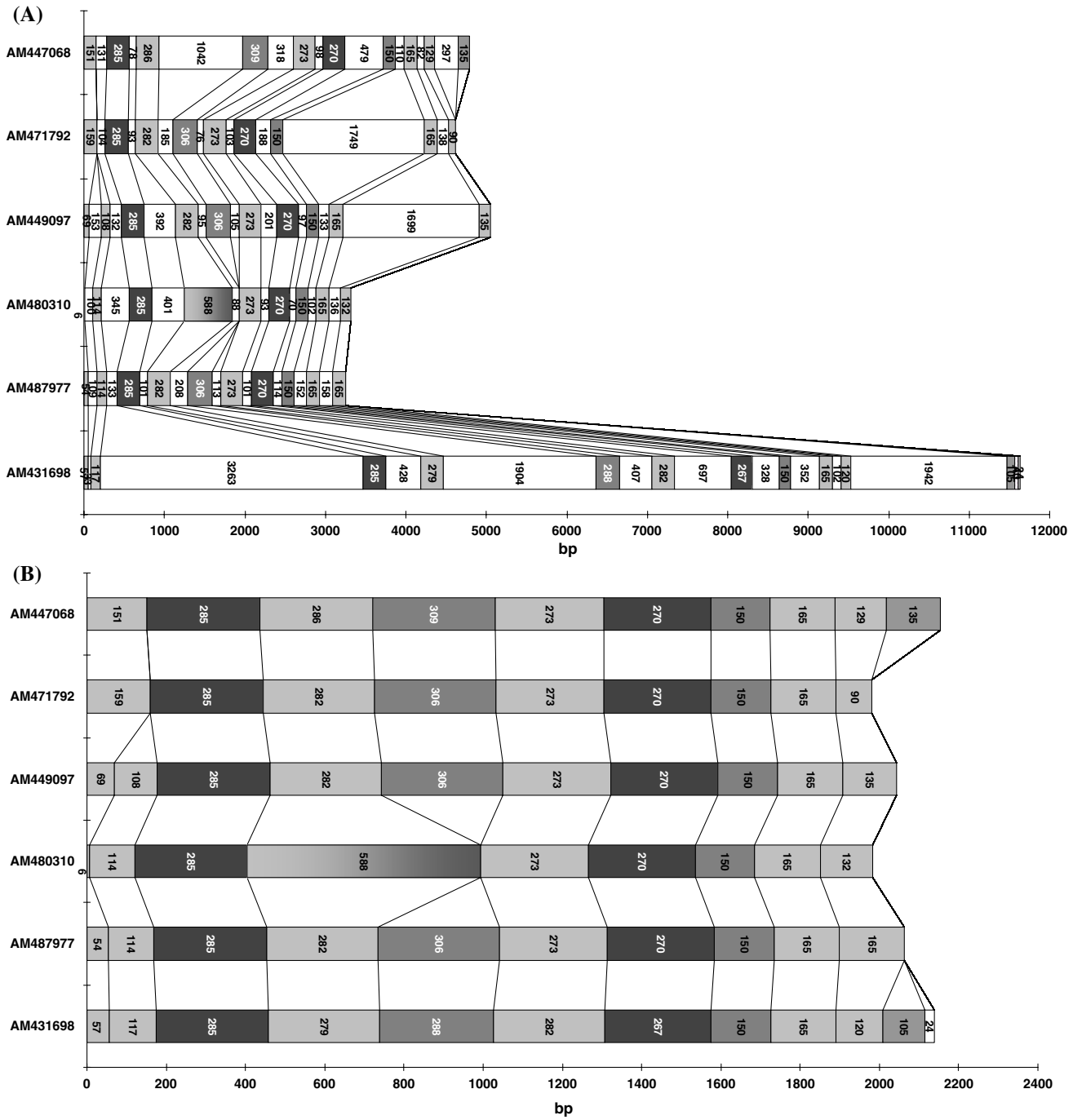


Fig. 3 Putative genomic organization of the six *DXS* sequences identified in grape (3′–5′). The complete gene structure is shown in (**a**) (introns are white and exons are grey), whereas only the coding portion is displayed in (**b**). *Numbers* represent nucleotidic sequence length in bp

QTL analysis

In this work we analyzed two segregating progenies, which were obtained from distinct aromatic varieties. We were interested in comparing the results of these experiments in order to highlight similarities and differences. Similarities could be expected given the narrow genetic basis of the Muscat family (Crespan and Milani 2001); on the other side, differences could be linked to the specificity of the genetic background. Italia, the aromatic parent of Pop1, derived from Muscat of Hamburg, which in turn is considered a progeny of Muscat of Alexandria. Moscato Bianco, the aromatic parent of Pop2, and Muscat of Alexandria were proposed as the progenitors of the two main recognized Muscat families and were found to be joined by a direct parent-offspring link (Crespan and Milani 2001).

Moreover, the two aromatic accessions used in this study showed opposite geraniol/linalool ratios (low in Italia, high in Moscato Bianco).

Analysis of the two segregating progenies demonstrated the existence of QTLs with effects either in diverse or in specific genetic backgrounds. Moreover, search for QTLs for the content of single monoterpenes rather than QTL analysis based on taste scoring provided interesting information about branch points playing an essential role in monoterpene biosynthesis.

A major QTL for the ln-transformed content of linalool, nerol and geraniol was identified for all three years on LG 5 in both the Italia × Big Perlon and Moscato Bianco × *Vitis riparia* maps (Table 2). In accordance with our results, Doligez et al. (2006) detected a major QTL for linalool, nerol and geraniol content on LG 5 in the interval VrZAG79-VVC6, which exhibited smaller effects for linalool than for nerol and geraniol. These findings suggest that a large part of the genetic variability of Muscat aroma might be due to a few genes with pleiotropic effects, but it is also possible that there are linked genes influencing the content of different aromatic compounds. These regulatory loci may occur in different genetic backgrounds.

Additional QTLs were identified for linalool content on LG 2 of the Moscato Bianco map and on LG 10 of the Italia × Big Perlon map (Table 2). The QTL on LG 2 of the Moscato Bianco map was detected in both years but in 2004 [not reported in Table 2(B)] it was significant only at the chromosome-wide level. This QTL was not sustained by KW analysis and its position needs to be refined by mapping new markers; however, its reliability is supported by the identification of a QTL for linalool content on the same chromosome of the non-Muscat parent utilized by Doligez et al. (2006). These findings suggest that, in spite of the close genetic relationship existing between many Muscat cultivars (Crespan and Milani 2001), different regulation points or allelic forms could be present and explain their aroma typicity (total terpene content and ratio between specific compounds, i.e. linalool and geraniol). However, we cannot exclude that the linalool-specific QTLs detected on LGs 2 (Pop2) and 10 (Pop1) play a regulatory role in both populations, even if it was not always revealed by QTL analysis. In order to test this possibility, as well as the interaction between loci, we investigated the association between geraniol/linalool ratio and the genotypes of the markers closest to the QTLs on LGs 5 (DXS in both populations) and 10 (cnd41 in Pop1 and VVIH01 in Pop2) by means of univariate analysis of variance (Figs. 1S, 2S). Our results suggest the existence of a locus regulating linalool accumulation on LG 10 in both populations, in addition to the locus controlling the level of all three monoterpenes on LG 5. This hypothesis is in agreement with the identification of a QTL for linalool/nerol ratio on LG 10 of Moscato Bianco in 2002 [Table 3(B)]. Analogous considerations do not seem to hold for LG 2 in Pop1, where only a small effect QTL for linalool/nerol ratio was detected through QTL analysis [Table 3(A)].

Our results also indicate that the observed phenotypic correlations between the three monoterpenes are, in large part, of genetic origin. In this study we detected both 'common' QTLs, namely loci regulating the accumulation of distinct molecules, and linalool-specific QTLs. These findings are in agreement with the significant correlation existing between the analyzed compounds (stronger between nerol and geraniol than between linalool and nerol or geraniol), which was also reported by Versini et al. (1993) and Doligez et al. (2006). From a chemical point of view, nerol and geraniol share a common precursor, geranyl-pyrophosphate (GPP), which produces also (*S*)-linalool in a one-step reaction catalyzed by (*S*)-linalool synthase (Ebang-Oke et al. 2003). However, additional distinct mechanisms for linalool synthesis could exist and nerol could be derived from the transformation of geraniol (Guardiola et al. 1996; Luan et al. 2005). The major QTL on LG 5 exhibited smaller effects for linalool than for nerol and geraniol in both populations analyzed in this study. However, the lower variance explained for the linalool QTL detected on this chromosome does not seem to be linked with the geraniol/linalool ratio in the aromatic parent, which proved to be high in Moscato Bianco (9.51), as well as in Muscat of Hamburg (Doligez et al. 2006), but low in Italia (0.13).

CG markers co-localizing with QTLs

DXS acts upstream in monoterpene biosynthesis: it encodes DXS, which is the first enzyme implicated in the non-mevalonate pathway of IPP biosynthesis. The correlated high concentration of linalool, nerol, geraniol (Tables 2S, 3S and 4S) is probably the consequence of an excess of a common precursor compound, i.e. IPP, which, in turn, is expected to derive from a favorable allele at locus *DXS*. Many investigations support a regulatory role of *DXS* in terpene biosynthesis at the transcriptional level. Estévez et al. (2001) observed a positive correlation between the accumulation of several plastidic isoprenoids and the content of DXS in transgenic *Arabidopsis* plants. They concluded that this enzyme catalyzes a limiting step in the DOXP/MEP pathway in plants. A positive correlation between the level of *DXS* transcript and the production of specific isoprenoids has been observed in several other plant systems (Bouvier et al. 1998; Chahed et al. 2000; Gong et al. 2006; Han et al. 2003; Khemvong and Suvachittanont 2005; Lange et al. 1998; Lois et al. 2000; Veau et al. 2000; Walter et al. 2000; Wungsintaweekul et al. 2008). Luan and Wüst (2002) analyzed the incorporation of labeled 1-deoxy-D-xylulose into linalool and geraniol in grape berries and suggested that

DXS is a limiting enzyme for plastidic isoprenoid biosynthesis also in grapevine. Furthermore, it is interesting to note that the gene *DXS1* was found to be differentially expressed during berry ripening (Pilati et al. 2007), in distinct grape berry tissues and in water-deficient compared with well-watered conditions (Grimplet et al. 2007).

The locus *CND41* in the nuclear genome encodes a protein, which non-specifically binds to chloroplast DNA. Characterization of antisense transgenic tobacco cells indicated that decreases in the CND41 protein increased the chloroplastic gene transcripts (Nakano et al. 1997). This observation suggests that CND41 is involved in the negative regulation of gene expression in chloroplasts, although the actual mechanism is not clear. The CND41 also has a protease activity, by which it may degrade transcriptional factors or RNA polymerase and decrease the level of chloroplast transcripts. CND41 might be implicated not only in the regulation of gene expression, but also in the biogenesis of the functional apparatus of chloroplasts and in the degradation of denatured proteins. This protein may function under stress conditions, which lower the cytosolic pH to its optimal value (Murakami et al. 2000). Nakano et al. (2003) reported a role for CND41 in the control of chloroplast development and gibberellic acid (GA) biosynthesis in an antisense tobacco transformant. A direct regulation effect of CND41 on the expression of the genes involved in the DOXP/MEP pathway should be excluded since they are encoded by the nuclear genome and then translocated to the plastids. Also this gene (TC57402) was found to be differentially expressed during berry ripening (Pilati et al. 2007).

FAH1 corresponds to the gene encoding fumarylacetoacetase, an enzyme of the hydrolase class that catalyzes the cleavage of fumarylacetoacetate to form acetoacetate and fumarate. This reaction is a step in the tyrosine catabolic pathway.

No genes underlying the QTL for linalool content on LG 2 of Moscato Bianco map could be identified since in Pop2 we mapped only DXS and the four CG markers, which turned out to be monomorphic in Pop1. Good candidate loci for this QTL could be genes acting downstream in the biosynthetic pathway of linalool, such as the (*S*)-linalool synthase (Ebang-Oke et al. 2003) or loci regulating it.

In silico analysis

Based on the in silico analysis of Pinot Noir genomic sequence (Velasco et al. 2007) a physical length of 267 kb was attributed to the mean 1.3 cM-long 1-LOD interval of the QTL on Italia LG 5 for the content of linalool, nerol and geraniol (Table 4). This distance is significantly shorter than the 8.3 cM-long VrZAG79-VVC6 interval reported by Doligez et al. (2006), which is estimated to correspond to more than 1.5 Mb. As a consequence, a relatively small (27) number of genes was predicted herein, thus giving strength to the hypothesis of an involvement of *DXS* in the regulation of Muscat aroma. On the contrary, the 1-LOD interval of the QTL on LG 10 for the content of linalool is still too broad and the number of potential genes (ca 200) is still too high to encourage focusing on one of them. Moreover, just outside the VVIH01-VrZAG67 interval two additional CGs for the control of linalool level were predicted to exist, which encode linalool synthase. Based on these findings and on literature reporting a regulatory role for *DXS* in terpene biosynthesis, we decided to further investigate gene copy number and organization in grapevine only for *DXS*.

Multi-copy gene families encode *DXS* and a number of genes involved in isoprenoid biosynthesis in plants, which may facilitate the tightly regulated expression of isoenzymes with roles specific to certain tissues, developmental stages and/or environmental challenges (Lange and Ghassemian 2003). The existence of a small *DXS* gene family has been suggested for *Arabidopsis* (Estévez et al. 2000; Rodríguez-Concepción and Boronat 2002), *Ginkgo biloba* (Kim et al. 2006), *Medicago truncatula* (Walter et al. 2002), *Morinda citrifolia* (Han et al. 2003), Norway spruce (Phillips et al. 2007), oil palm (Khemvong and Suvachittannont 2005), rice (Kim et al. 2005) and even for the purple non-sulfur bacterium *Rhodobacter capsulatus* (Hahn et al. 2001). By using TC56417 as query to mine the grapevine whole genome sequence, we identified six contigs encoding ortholog *DXS* sequences. To our knowledge this is the first evidence of the existence of *DXS* genes class 1, 2 and 3 in grape. *DXS2 B* and *D*, which were both located on LG 15, could represent allelic variants of the same gene (Table 5). A high level of similarity at the nucleotide level was observed among the coding regions of all the identified *DXS* sequences, as revealed by the presence of seven exons with the same length (Fig. 3b), whereas the level of similarity among intron regions was very low even within the same gene class (Fig. 3a). These data encourage verifying the existence of different spatial and temporal expression patterns of *DXS* isoforms and investigating their biological role in grapevine.

Conclusions

In this work we combined metabolite profiling with QTL analysis and CG approach to identify key genetic factors underlying Muscat aroma in grapevine. Our results indicate that a large part of the genetic variability for this trait might be due to pleiotropic or closely linked genes, but independent genes controlling specific branch points in terpenoid biosynthesis are probably also implicated. Interesting associations between QTLs and CGs were identified as well, which support their involvement in the phenotypic varia-

tion under study (Price 2006). In particular, the co-localization of the CG marker DXS and the QTL on LG 5 for the content of linalool, nerol and geraniol suggests a role of DXS in regulating metabolic flux through the DOXP/MEP pathway. To our knowledge, these findings represent the first indication of a putative key gene in the genetic determination of Muscat flavor in grapevine and are in complete agreement with observations made in other plant systems. However, further experiments are in progress to confirm *DXS* as candidate for QTLs, like fine mapping, association genetics, transcriptomic profiling and functional testing. Given success, this CG might be used as an efficient molecular marker to help in selecting desirable alleles, with a potential application in crop improvement.

Acknowledgments This research was jointly funded by the EU grant (INCO project MASTER, contract ICA4-2001-10065), the autonomous province of Trento, Italy (QTL-A project) and the Fondazione CARITRO Trento, Italy (Advanced Biology Program). We would like to thank also Girolamo Fanizza and Marco Stefanini for the management, respectively of Pop1 and Pop2, Alessandro Cestaro for his support in bioinformatic analysis, and Silvia Carlin for her contribution in metabolite profiling.

References

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783–795

Birney E, Clamp M, Durbin R (2004) GeneWise and genomewise. Genome Res 14:988–995

Bouvier F, d'Harlingue A, Suire C, Backhaus RA, Camara B (1998) Dedicated roles of plastid transketolases during the early onset of isoprenoid biogenesis in pepper fruits. Plant Physiol 117:1423–1431

Chahed K, Oudin A, Guivarc'h N, Hamdi S, Chénieux JC, Rideau M, Clastre M (2000) 1-deoxy-D-xylulose 5-phosphate synthase from periwinkle: cDNA identification and induced gene expression in terpenoid indole alkaloid-producing cells. Plant Physiol Biochem 38:559–566

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Clastre M, Bantignies B, Feron G, Soler E, Ambid C (1993) Purification and characterization of geranyl diphosphate synthase from *Vitis vinifera* L. cv Muscat de Frontignan cell cultures. Plant Physiol 102:205–211

Costantini L, Battilana J, Lamaj F, Fanizza G, Grando MS (2008) Berry and phenology-related traits in grapevine (*Vitis vinifera* L.): from Quantitative Trait Loci to underlying genes. BMC Plant Biol 8:38

Crespan M, Milani N (2001) The Muscats: a molecular analysis of synonyms, homonyms and genetic relationships within a large family of grapevine cultivars. Vitis 40:23–30

da Silva FG, Iandolino A, Al-Kayal F, Bohlmann MC, Cushman MA, Lim H, Ergul A, Figueroa R, Kabuloglu EK, Osborne C, Rowe J, Tattersall E, Leslie A, Xu J, Baek J, Cramer GR, Cushman JC, Cook DR (2005) Characterizing the grape transcriptome. Analysis of expressed sequence tags from multiple *Vitis* species and development of a compendium of gene expression during berry development. Plant Physiol 139:574–597

Doligez A, Audiot E, Baumes R, This P (2006) QTLs for muscat flavour and monoterpenic odorant content in grapevine (*Vitis vinifera* L.). Mol Breeding 18:109–125

Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res 19:4008

Ebang-Oke JP, de Billerbeck GM, Ambid C (2003) Temporal expression of the *Lis* gene from *Vitis vinifera* L., cv. Muscat de Frontignan. In: Le Quéré JL, Etiévant PX (eds) Proceedings of the 10th Weurman flavour research symposium "Flavour research at the dawn of the twenty-first century", Beaune (France), 25–28 June 2002. Lavoisier, Paris, France, pp 321–325

Estévez JM, Cantero A, Romero C, Kawaide H, Jiménez LF, Kuzuyama T, Seto H, Kamiya Y, León P (2000) Analysis of the expression of *CLA1*, a gene that encodes the 1-deoxyxylulose 5-phosphate synthase of the 2-C-methyl-D-erythritol-4-phosphate pathway in *Arabidopsis*. Plant Physiol 124:95–103

Estévez JM, Cantero A, Reindl A, Reichler S, Leo P (2001) 1-deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. J Biol Chem 276:22901–22909

Gong YF, Liao ZH, Guo BH, Sun XF, Tang KX (2006) Molecular cloning and expression profile analysis of *Ginkgo biloba* DXS gene encoding 1-deoxy-D-xylulose 5-phosphate synthase, the first committed enzyme of the 2-C-methyl-D-erythritol 4-phosphate pathway. Planta Med 72:329–335

Grando MS, Bellin D, Edwards KJ, Pozzi C, Stefanini M, Velasco R (2003) Molecular linkage maps of *Vitis vinifera* L. and *V. riparia* Mchx. Theor Appl Genet 106:1213–1224

Grimplet J, Deluc LG, Tillett RL, Wheatley MD, Schlauch KA, Cramer GR, Cushman JC (2007) Tissue-specific mRNA expression profiling in grape berry tissues. BMC Genomics 8:187

Guardiola J, Iborra JL, Rodenas L, Canovas M (1996) Biotransformation from geraniol to nerol by immobilized cells (*V. vinifera*). Appl Biochem Biotechnol 56:169–180

Günata Z, Biron C, Sapis JC, Bayonove C (1989) Glycosidase activities in sound and rotten grapes in relation to hydrolysis of grape monoterpenyl glycosides. Vitis 28:191–197

Hahn FM, Eubanks LM, Testa CA, Blagg BSJ, Baker JA, Poulter CD (2001) 1-Deoxy-D-xylulose 5-phosphate synthase, the gene product of open reading frame (ORF) 2816 and ORF 2895 in *Rhodobacter capsulatus*. J Bacteriol 183:1–11

Han YS, Roytrakul S, Verberne MC, van der Heijden R, Linthorst HJM, Verpoorte R (2003) Cloning of a cDNA encoding 1-deoxy-D-xylulose 5-phosphate synthase from *Morinda citrifolia* and analysis of its expression in relation to anthraquinone accumulation. Plant Sci 164:911–917

Huang X, Madan A (1999) CAP3: a DNA sequence assembling program. Genome Res 9:868–877

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pè ME, Valle G, Morgante M, Caboche M, Adam-Blondon AF, Weissenbach J, Quétier F, Wincker P (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–467

Khemvong S, Suvachittanont W (2005) Molecular cloning and expression of a cDNA encoding 1-deoxy-D-xylulose-5-phosphate synthase from oil palm *Elaeis guineensis* Jacq. Plant Sci 169:571–578

Kim BR, Kim SU, Chang YJ (2005) Differential expression of three 1-deoxy-D-xylulose-5-phosphate synthase genes in rice. Biotechnol Lett 27:997–1001

Kim SM, Kuzuyama T, Chang YJ, Song KS, Kim SU (2006) Identification of class 2 1-deoxy-D-xylulose 5-phosphate synthase and

1-deoxy-D-xylulose 5-phosphate reductoisomerase genes from *Ginkgo biloba* and their transcription in embryo culture with respect to ginkgolide biosynthesis. Planta Med 72:234–240

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Lange BM, Ghassemian M (2003) Genome organization in *Arabidopsis thaliana*: a survey for genes involved in isoprenoid and chlorophyll metabolism. Plant Mol Biol 51:925–948

Lange BM, Wildung MR, McCaskill D, Croteau R (1998) A family of transketolases that directs isoprenoid biosynthesis via a mevalonate-independent pathway. Proc Natl Acad Sci USA 95:2100–2104

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659

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

Lois LM, Rodríguez-Concepción M, Gallego F, Campos N, Boronat A (2000) Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. Plant J 22:503–513

Luan F, Wüst M (2002) Differential incorporation of 1-deoxy-D-xylulose into (3S)-linalool and geraniol in grape berry exocarp and mesocarp. Phytochemistry 60:451–459

Luan F, Mosandl A, Münch A, Wüst M (2005) Metabolism of geraniol in grape berry mesocarp of *Vitis vinifera* L. cv. Scheurebe: demonstration of stereoselective reduction, E/Z-isomerization, oxidation and glycosylation. Phytochemistry 66:295–303

Lücker J, Bowen P, Bohlmann J (2004) *Vitis vinifera* terpenoid cyclases: functional identification of two sesquiterpene synthase cDNAs encoding (+)-valencene synthase and (−)-germacrene D synthase and expression of mono- and sesquiterpene synthases in grapevine flowers and berries. Phytochemistry 65:2649–2659

Mahmoud SS, Croteau RB (2002) Strategies for transgenic manipulation of monoterpene biosynthesis in plants. Trends Plant Sci 7:366–373

Martin DM, Bohlmann J (2004) Identification of *Vitis vinifera* (-)-α-terpineol synthase by in silico screening of full-length cDNA ESTs and functional characterization of recombinant terpene synthase. Phytochemistry 65:1223–1229

Mateo JJ, Jiménez M (2000) Monoterpene in grape juice and wines. J Chromatogr 881:557–567

Mathieu S, Terrier N, Procureur J, Bigey F, Günata Z (2005) A carotenoid cleavage dioxygenase from *Vitis vinifera* L.: functional characterization and expression during grape berry development in relation to C_{13}-norisoprenoid accumulation. J Exp Bot 56:2721–2731

Morgante M, Salamini F (2003) From plant genomics to breeding practice. Curr Opin Biotech 14:214–219

Morreel K, Goeminne G, Storme V, Sterck L, Ralph J, Coppieters W, Breyne P, Steenackers M, Georges M, Messens E, Boerjan W (2006) Genetical metabolomics of flavonoid biosynthesis in *Populus*: a case study. Plant J 47:224–237

Murakami S, Kondo Y, Nakano T, Sato F (2000) Protease activity of CND41, a chloroplast nucleoid DNA-binding protein, isolated from cultured tobacco cells. FEBS Lett 468:15–18

Nakano T, Murakami S, Shoji T, Yoshida S, Yamada Y, Sato F (1997) A novel protein with DNA binding activity from tobacco chloroplast nucleoids. Plant Cell 9:1673–1682

Nakano T, Nagata N, Kimura T, Sekimoto M, Kawaide H, Murakami S, Kaneko Y, Matsushima H, Kamiya Y, Sato F, Yoshida S (2003) CND41, a chloroplast nucleoid protein that regulates plastid development, causes reduced gibberellin content and dwarfism in tobacco. Physiol Plant 117:130–136

Paran I, Zamir D (2003) Quantitative traits in plants: beyond the QTL. Trends Genet 19:303–306

Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. Mol Breed 7:275–291

Phillips MA, Walter MH, Ralph SG, Dabrowska P, Luck K, Urós EM, Boland W, Strack D, Rodríguez-Concepción M, Bohlmann J, Gershenzon J (2007) Functional identification and differential expression of 1-deoxy-D-xylulose 5-phosphate synthase in induced terpenoid resin formation of Norway spruce (*Picea abies*). Plant Mol Biol 65:243–257

Pilati S, Perazzolli M, Malossini A, Cestaro A, Demattè L, Fontana P, Dal Ri A, Viola R, Velasco R, Moser C (2007) Genome-wide transcriptional analysis of grapevine berry ripening reveals a set of genes similarly modulated during three seasons and the occurrence of an oxidative burst at vèraison. BMC Genomics 8:428

Price AH (2006) Believe it or not, QTLs are accurate! Trends Plant Sci 11:213–216

Remington DL, Ungerer MC, Purugganan MD (2001) Map-based cloning of quantitative trait loci: progress and prospects. Genet Res Camb 78:213–218

Ribérau-Gayon P, Boidron JN, Terrier A (1975) Aroma of muscat grape varieties. J Agric Food Chem 23:1042–1047

Ribéreau-Gayon P, Glories Y, Maujean A, Dubourdieu D (2000) The chemistry of wine: stabilization and treatments. In: Wiley J and Sons Ltd (eds) Handbook of enology, vol 2, pp 187–206

Rodríguez-Concepción M, Boronat A (2002) Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiol 130:1079–1089

Salamov A, Solovyev V (2000) Ab initio gene finding in *Drosophila* genomic DNA. Genome Res 10:516–522

Salvi S, Tuberosa R (2005) To clone or not to clone plant QTLs: present and future challenges. Trends Plant Sci 10:297–304

Sanguinetti CJ, Dias Neto E, Simpson AJ (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. Biotechniques 17:914–921

Sevini F, Marino R, Grando MS, Moser S, Versini G (2004) Mapping candidate genes and QTLs for aroma content in grape. Acta Hortic 652:439–446

Small I, Peeters N, Legeai F, Lurin C (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics 4:1581–1590

Strauss CR, Wilson B, Gooley PR, Williams PJ (1986) Role of monoterpenes in grape and wine flavour. In: Parliament T, Croteau R (eds) Biogeneration of aromas. American Chemical Society, Washington, pp 222–242

Van Ooijen JW, Boer MP, Jansen RC, Maliepaard C (2002) MapQTL® 4.0 Software for the calculation of QTL positions on genetic maps. Plant Research International, Wageningen

Veau B, Courtois M, Oudin A, Chénieux JC, Rideau M, Clastre M (2000) Cloning and expression of cDNAs encoding two enzymes of the MEP pathway in *Catharanthus roseus*. BBA Gene Struct Exp 1517:159–163

Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, FitzGerald LM, Vezzulli S, Reid J, Malacarne G, Iliev D, Coppola G, Wardell B, Micheletti D, Macalma T, Facci M, Mitchell JT, Perazzolli M, Eldredge G, Gatto P, Oyzerski R, Moretto M, Gutin N, Stefanini M, Chen Y, Segala C, Davenport C, Demattè L, Mraz A, Battilana J, Stormo K, Costa F, Tao Q, Si-Ammour A, Harkins T, Lackey A, Perbost C, Taillon B, Stella A, Solovyev V, Fawcett J, Sterck L, Vandepoele K, Grando MS, Toppo S, Moser C, Lanchbury J, Bogden R, Skolnick M, Sgaramella V, Bhatnagar SK, Fontana P, Gutin A, Van de Peer Y, Salamini F, Viola R (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoSONE 2(12):e1326

Versini G, Dalla Serra A, Dell'Eva M, Scienza A, Rapp A (1988) Evidence of some glycosidically bound new monoterpenes and norisoprenoids in grapes. In: Schreier P (ed) Bioflavour '87. W de Gruyter, Berlin, Germany, pp 161–170

Versini G, Dalla Serra A, Monetti A, De Micheli L, Mattivi F (1993) Free and bound grape aroma profiles variability within the family of muscat-called varieties. In: Bayonove C, Crouzet J, Flanzy C, Martin JC, Sapis JC (eds) Proceedings of the International Symposium "Connaissance aromatique des cépages et qualité des vins", Montpellier (France), 9–10 February 1993. Revue Française d'Oenologie, Lattes, France, pp 12–21

Wagner R (1967) Etude de quelques disjonctions dans des descendances de Chasselas, Muscat Ottonel et Muscat à petits grains. Vitis 6:353–363

Walter MH, Fester T, Strack D (2000) Arbuscolar mycorrhizal fungi induce the non-mevalonate methylerythritol phosphate pathway of isoprenoid biosynthesis correlated with accumulation of the 'yellow pigment' and other apocarotenoids. Plant J 21:571–578

Walter MH, Hans J, Strack D (2002) Two distantly related genes encoding 1-deoxy-D-xylulose 5-phosphate synthases: differential regulation in shoots and apocarotenoid-accumulating mycorrhizal roots. Plant J 31:243–254

Wungsintaweekul J, Sirisuntipong T, Kongduang D, Losuphanporn T, Ounaroon A, Tansakul P, De-Eknamkul (2008) Transcription profiles analysis of genes encoding 1-deoxy-D-xylulose 5-phosphate synthase and 2C-methyl-D-erythritol 4-phosphate synthase in plaunotol biosynthesis from *Croton stellatopilosus*. Biol Pharm Bull 31:842–856