



## Neutral invertases in grapevine and comparative analysis with Arabidopsis, poplar and rice

Alberto Nonis, Benedetto Ruperti, Alessandro Pierasco, Aurelie A. Canaguier, Anne-Francoise A.-F. Adam-Blondon, Gabriele Di Gaspero, Giannina Vizzotto

### ► To cite this version:

Alberto Nonis, Benedetto Ruperti, Alessandro Pierasco, Aurelie A. Canaguier, Anne-Francoise A.-F. Adam-Blondon, et al.. Neutral invertases in grapevine and comparative analysis with Arabidopsis, poplar and rice. *Planta*, 2008, 229 (1), pp.129-142. 10.1007/s00425-008-0815-0 . hal-02668994

**HAL Id: hal-02668994**

**<https://hal.inrae.fr/hal-02668994>**

Submitted on 31 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Neutral invertases in grapevine and comparative analysis with *Arabidopsis*, poplar and rice

Alberto Nonis · Benedetto Ruperti · Alessandro Pierasco ·  
Aurelie Canaguier · Anne-Françoise Adam-Blondon ·  
Gabriele Di Gaspero · Giannina Vizzotto

Received: 10 June 2008 / Accepted: 2 September 2008 / Published online: 18 September 2008  
© Springer-Verlag 2008

**Abstract** Neutral invertases (NIs, EC 3.2.1.26) cleave sucrose to glucose and fructose. They are encoded by a small gene family of 9 members in the *Arabidopsis* genome, 8 in rice, 16 in poplar and 9 in *Vitis vinifera* (L.). The grapevine NIs were identified in the 8.4X genome assembly of the quasi-homozygous line PN40024. In addition, alleles of three NIs were sequenced in the heterozygous cultivar ‘Cabernet Sauvignon’. Analyses of sequence variation between alleles, homoeologous and paralogous copies in grapevine and their orthologues in *Arabidopsis*, poplar and rice are provided. In grapevine, NIs were classified into four  $\alpha$  NIs and five  $\beta$  NIs and subsequently grouped into hierarchical clades using a combination of

evidence including amino acid identity, exon/intron structure, rate of synonymous substitutions ( $K_s$ ) and chromosomal distribution. Estimation of  $K_s$  proved the ancient origin of all NIs and the lack of expansion by gene duplication past the event of polyploidisation. We then focused on transcription analysis of five NIs for which evidence of expression was available from expressed sequence tag databases. Among these, four NIs consisted of pairs of homoeologous copies, each pair lying on a pair of chromosomes duplicated by polyploidy. Unequal expression of homoeologous genes was observed by quantitative RT-PCR in leaf, flower, seed and root tissues. Since NIs might play significant roles in fruit and wine quality, NIs expression was monitored in flesh and skin of ‘Merlot’ berries and shown in parallel with the suite of changes that accompany fruit ripening, including glucose and fructose accumulation.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00425-008-0815-0) contains supplementary material, which is available to authorized users.

A. Nonis · B. Ruperti · A. Pierasco · G. Di Gaspero (✉) ·  
G. Vizzotto  
Dipartimento di Scienze Agrarie e Ambientali,  
University of Udine, via delle Scienze 208, 33100, Udine, Italy  
e-mail: gabriele.digaspero@uniud.it

A. Canaguier · A.-F. Adam-Blondon  
UMR de Génétique Végétale, INRA-CNRS-UEVE,  
2, Rue Gaston Crémieux, CP5708, 91057 Evry Cedex, France

G. Di Gaspero  
Istituto di Genomica Applicata,  
Parco Scientifico e Tecnologico Luigi Danieli,  
via Jacopo Linussio 51, Udine, Italy

*Present Address:*

B. Ruperti  
Dipartimento di Agronomia Ambientale e Produzioni Vegetali,  
University of Padua, viale dell’Università 16, Legnaro,  
Padua, Italy

**Keywords** Expression divergence · Fruit development · Homoeologous genes · Ripening · *Vitis vinifera*

## Abbreviations

BAC	Bacterial artificial chromosome
Chr	Chromosome
DAFB	Days after full bloom
EST	Expressed sequence tag
FPC	Fingerprint contig
Indel	Insertion/deletion
NJ	Neighbor-joining
NI	Neutral invertase
SNP	Single nucleotide polymorphism
SSCP	Single strand conformational polymorphism
$K_s$	Substitutions per synonymous site
$K_a$	Substitutions per non-synonymous site
UTR	Untranslated region
WGD	Whole genome duplication

## Introduction

Sugars are important components of fruit quality. Hexoses sweeten the pulp of grapes sold for fresh consumption and are the substrates for wine yeast fermentation. Ethanol is indispensable for preservation of wine and contributes to the complexity of sensory perception. Sugar accumulation in the fruit is a limiting factor in cool climates for achieving an adequate ethanol concentration in wines. The wine industry copes with this limitation by enriching must with sucrose from other plants or by removing water by means of reverse osmosis. Horticultural practises for improving sugar content include reduction of crop load by cluster thinning and phloem girdling of shoots proximal to the clusters. Both techniques are time-consuming and conveniently apply only to the production of upmarket wines.

Beside their economic value, sugars and the pathways of their biosynthesis, storage and signalling have stirred much interest because they are at the core of fruit metabolism (Sheen et al. 1999; Smeekens 2000; Koch 2004; Gibson 2005). The grape berry has active photosynthetic skin and its flesh burns all of the imported sugars for sustaining the initial phase of cell division and growth. Soon after maturation has been triggered, the epicarp specialises in the synthesis of secondary metabolites and the mesocarp sinks sucrose. Drop of cell turgor and softening are the earliest tangible reporters of the onset of maturation in the grape berry (Terrier et al. 2005; Thomas et al. 2006). This change is accompanied by an increase in apoplastic solutes in the mesocarp and a shift of sugar trafficking from symplastic to apoplastic phloem unloading (Zhang et al. 2006; Wada et al. 2008). Then, hexose accumulation and, in red grapes, onset of anthocyanin biosynthesis become detectable within a few days.

Sucrose is cleaved to UDP-glucose and fructose by sucrose synthases and to glucose and fructose by invertases (Winter and Huber 2000; Koch 2004). Invertases are classified into three sub-families based on their biochemical properties and their presumed sub-cellular localisation (Tymowska-Lalanne and Kreis 1998). Vacuolar and cell wall invertases have an acidic pH-optimum and are associated with osmoregulation, responses to wounding or pathogen infection and cold-induced sweetening across different plant taxa (reviewed in Sturm and Tang 1999; Roitsch and Gonzalez 2004). The activity of cell wall invertases increases in the grape berry at the inception of fruit maturation (Zhang et al. 2006), while expression of vacuolar invertases is down-regulated at the same time (Davies and Robinson 1996). Neutral invertases (NI) have an alkaline/neutral pH-optimum and are expressed in sugar accumulating sinks (Gallagher and Pollock 1998; Sturm et al. 1999; Bosch et al. 2004), including soft fruit (Nonis et al. 2007).

NIs are encoded by small gene families in the completely sequenced plant genomes, with 9 members in *Arabidopsis thaliana*, 8 in rice and 16 in poplar (The Arabidopsis Genome Initiative 2000; Ji et al. 2005; Bococek et al. 2008). Exon–intron structure and amino acid identity led Ji et al. (2005) to distinguish  $\alpha$  and  $\beta$  NIs in rice. The four rice  $\alpha$  NIs have a six-exon structure, while the four  $\beta$  NIs are encoded by four exons each. All NIs are very similar at the C-terminus region, which differentiates them from cell wall and vacuolar invertases (Roitsch and Gonzalez 2004). NIs were presumed to localise in the cytoplasm owing to the pH-optimum, however, the presence of peptide signals that target the mature proteins in mitochondria and chloroplasts has recently called this into question (Ji et al. 2005). Sub-cellular localisation in both organelles was demonstrated in rice, spinach and *Arabidopsis* (Murayama and Handa 2007; Vargas et al. 2008).

Low transcript levels and faint enzyme activity were measured across organs and species, leading to the hypothesis that NIs are expressed at a basal level (Gallagher and Pollock 1998; Sturm et al. 1999; Bosch et al. 2004; Roitsch and Gonzalez 2004). However, expression of a peach NI was modulated during fruit development and in response to exogenous sucrose and growth regulators (Nonis et al. 2007). In cell cultures of sugarcane, NI activity peaked during the exponential phase of growth and antisense NI lines grew poorly, suggesting that NI contributes to sucrose mobilisation and to monosaccharide supply for respiration and cell growth (Rossouw et al. 2007). Mutants defective for a NI showed an inhibition of root elongation in *Arabidopsis*, a phenotype which resulted from altered hexose metabolism (Qi et al. 2007; Lou et al. 2007).

Our goals were to investigate the size and the genomic organisation of the NI gene family in grapevine with respect to other fully sequenced plant genomes and to compare NI expression patterns and sugar content in fleshy fruits that accumulate hexoses. For the first purpose, gene structures of the whole family were analysed in the quasi-homozygous line PN40024, used for genome sequencing, and allelic variation was studied in the heterozygous cultivar ‘Cabernet Sauvignon’. We showed that the nine NIs found in the grapevine genome have an ancient origin, which partially predates the polyploidisation events at the roots of dicot lineages (Jaillon et al. 2007; Ming et al. 2008). Homoeologous copies duplicated by polyploidy were largely retained in the present-day grapevines and have undergone expression divergence. For the second purpose, we showed that at least five NI genes were transcribed in two tissues of the berry, throughout ripening. Small changes in mRNA abundance of most NIs were temporally associated with hexose accumulation in berry flesh and skin.

## Materials and methods

### Sequence analysis

A tBLASTx search was performed on the 8.4X assembly of the grape genome with an *E*-value cutoff of  $e^{-04}$ , using NI proteins predicted in *Arabidopsis*, rice and poplar as query sequences (Vargas et al. 2003; Ji et al. 2005; Bockock et al. 2008). The gene models in PN40024 and their chromosomal positions were retrieved from the website <http://www.genoscope.cns.fr/Vitis>. Expressed sequence tags (ESTs) were identified by BLASTn of the grape EST database of the University of California, Davis, USA (<http://cgf.ucdavis.edu/>), using the PN40024 gene models as query sequences. Selective primers for five expressed NIs were designed using Primer3 (Supplementary Material S1). The exon–intron structure of the alleles from ‘Cabernet Sauvignon’ was predicted by comparison with the PN40024 gene models and with the ESTs.  $K_s$  and  $K_a/K_s$  ratio were calculated using DNAsp 4.0 (Rozas et al. 2003). Homoeologous regions containing NIs were determined by performing a BLASTp search of 20 translated genes surrounding a given NI: the corresponding blocks of DNA, assigned to a different chromosome with the same order of genes and with two or more surrounding genes showing the best reciprocal match with the original query, were then identified. The locations of these putative homoeologous regions were compared with the chromosomal arrangement of the three ancestral genomes as reported in Jaillon et al. (2007). Sequence alignments were performed with ClustalX. Neighbor-Joining (NJ) trees were constructed with MEGA (Kumar et al. 2004). Prediction of peptide targeting signals was performed using ChloroP, Predotar, TargetP and WoLF PSORT (Emanuelsson et al. 1999, 2000; Small et al. 2004; Horton et al. 2006). Regulatory *cis*-elements were predicted using PlantCARE (Lescot et al. 2002).

### Sequencing alleles

A 6X bacterial artificial chromosome (BAC) library of ‘Cabernet Sauvignon’ was screened by PCR using specific primers for NIs (Adam-Blondon et al. 2005). A three-dimensional pooling scheme was used to identify the position of the positive BACs in the library. Positive BACs were confirmed by PCR using a single BAC as a template, and identified in the corresponding contig of the physical map (Moroldo et al. 2008). All BACs within a positive contig were tested with the specific primers of the corresponding gene member. The positive BAC clones were then used for validating the selectivity of each paralogue-specific primer pair by amplifying all BAC clones harbouring NIs and ruling out cross-amplification from BACs harbouring paralogous NIs. Amplicons of positive BACs were run in a

single strand conformational polymorphism (SSCP) gel, as described by Di Gaspero et al. (2007), in parallel with the amplicons generated from ‘Cabernet Sauvignon’ and 46 offspring of ‘Cabernet Sauvignon’ in order to confirm the allelic status of NI variants. BAC DNA was extracted from one representative BAC per allele per gene, by means of alkaline lysis of bacterial cells followed by double acetate precipitation. NI alleles were sequenced by primer walking from 1.5 kb upstream of the start codon to the polyadenylation site, using an Applied Biosystems 3730xl sequencer and Sanger chemistry. Quality checks of the sequences were performed with Phred (Ewing et al. 1998) and sequences were assembled using Phrap (Ewing et al. 1998).

### Plant material, quantitation of sugars and gene expression

Vines of *Vitis vinifera* ‘Merlot’ grafted on rootstock SO4 were grown in North-Eastern Italy (46° 02' N; 13° 13' E; 88 masl) with a spur-pruned cordon training system, at a planting density of 4,000 vines/ha. At *véraison*, during the period when anthocyanin-pigmented and green berries coexist on the same cluster, samples of 40 green berries and 40 red berries were collected over the clusters of 12 plants and processed independently. Sampling was performed over a 73-day time span in the season of 2005, at eight time points from 45 to 118 days after full blooming (DAFB). Skins were peeled and seeds were removed from flesh for all subsequent analyses. Fully expanded leaves and adventitious rootlets were collected from canes grown in hydroponic culture. Flower buttons were collected 1 week before blooming and seeds were extracted from berries collected at 86 DAFB.

Sugars were extracted from 0.5 g of skin and flesh tissues in 2.5 ml of boiling 100% ethanol for 5 min, then twice in 60% ethanol, twice in 20% ethanol and twice in water for 10 min each step. The solutions recovered from every step of extraction were mixed. Glucose and fructose were measured using an enzymatic method reported by Vizzotto et al. (1996).

Total RNA was extracted using the protocol described by Iandolino et al. (2004), with slight modifications. Three grams of powdered tissue were added to 24 ml of 100 mM Tris–HCl, 25 mM EDTA, 2% CTAB, 2.5 M NaCl, 2% polyvinylpyrrolidone, 0.5 g/l spermidine, 2%  $\beta$ -mercaptoethanol buffer, at pH = 8.0 and kept at 65°C for 30 min. Debris was discarded after centrifugation at 6,000 g for 10 min. RNA was extracted twice by adding 25 ml of 24:1 chloroform:isoamyl alcohol, and centrifugation at 6,000 g for 10 min. RNA was precipitated overnight at 4°C by adding 0.2 volumes of 12 M LiCl to the aqueous phase. RNA pellet was recovered by a centrifugation at 16,500 g for 40 min, rinsed with 70% ethanol and re-suspended in RNase-free water. RNA integrity was checked by gel

electrophoresis. After DNase treatment, 1 µg of RNA was retro-transcribed and quantitative RT-PCR was performed as reported by Nonis et al. (2007). Transcripts were quantified upon normalisation with the gene coding for the ubiquitin conjugating enzyme (accession no. CF203457.1; sense primer, CTATATGCTCGCTGCTGACG; antisense primer AAGCCAGGCAGAGACAACCTC), based on three technical replicates. Expression levels were reported as arbitrary units by comparing the threshold cycle of each gene with that of the housekeeping gene. Melting curves were generated for confirming the presence of a unique amplicon.

## Results

### Number and gene structure of NIs in the grapevine genome

Neutral invertases were identified in the 8.4X genome assembly of PN40024. A tBLASTx search was performed using all NIs known in *Arabidopsis*, poplar and rice with an *E*-value cutoff of  $e^{-04}$ . Nine genes showing significant identity with plant NIs were predicted in the grapevine genome based on this search (Table 1, and additional information in Supplementary Material S1). Eight of these genes are located across chromosomes 3, 5, 6, 13, 14 and 18. Chromosomes 6 and 18 harbour two copies each, but each copy is located a few millions base pairs from the other one on the same chromosome. No evidence for local gene duplication or intra-chromosomal clustering was found. The ninth gene member is located on scaffold\_203, which has not yet been assigned to a chromosome. The relationship of orthology with *Arabidopsis* and poplar was assessed by the best reciprocal match at the protein level and is given in Supplementary Material S1. ESTs for five NIs out of the nine grapevine members were identified by the best reciprocal match in the UC Davis CGF public database (<http://cgf.ucdavis.edu/home/>). The corresponding ESTs were originally sequenced from cDNA of different organs in the cultivars ‘Chardonnay’, ‘Cabernet Sauvignon’ and ‘Pinot noir’ (Table 1), which justified the occurrence of some varietal single nucleotide polymorphisms (SNPs) in comparison with the BLAST query. We proposed a nomenclature for the grapevine NIs from *VvNII* to -9, assigning priority to members with evidence of expression (*VvNII* to -5). Three other NIs (*VvNI6*, -7, -9) matched the same EST attributed to *VvNII*, but the level of significance of these matches was far lower compared with *VvNII* (Table 1, Supplementary Material S1). *VvNI8* did not match any EST at an *E*-value  $< e^{-04}$ .

The predicted exon–intron structure of *VvNII* to -9 is shown in Fig. 1. *VvNII*, -6, -7 and -8 have a four-exon

structure, peculiar to the  $\beta$  group of NIs. *VvNI2*, -3, -4 and -5 have six-exon gene models, which is a structural hallmark of  $\alpha$  NIs. For *VvNI9*, the automatic annotator GAZE predicted six exons (Jaillon et al. 2007). However, the length of single exons and the total length of the virtual cDNA recall a substantial similarity with the  $\beta$  group of NIs, as if the second and the third introns were erroneously predicted or had originated within the second exon of a former four-exon gene. *VvNI4*, -5 and -7 contained large introns. The first intron of *VvNI4* spanned 3,637 bp and contained a LINE transposable element. The third intron of *VvNI5* (~9.7 kb) and the second intron of *VvNI7* (~8.2 kb) contained repetitive elements of unknown classes. The 12 conserved domains, characteristic of the rice NIs described by Ji et al. (2005), were identified in the amino acid sequences of all nine *VvNIs*.

### Allelic variation of NI genes in a heterozygous genotype

Primer pairs were designed for the five NIs (*VvNII* to -5) for which evidence of transcription was indicated by the presence of the corresponding EST in public databases (Supplementary Material S1). These primers allowed for the PCR identification and sequencing of allelic sequences of *VvNII*, -2 and -3 (AM930845-50) in ‘Cabernet Sauvignon’ genomic DNA cloned into a BAC vector. BACs containing *VvNI4* and -5 could not be found in the same library. Owing to the heterozygous nature of ‘Cabernet Sauvignon’ and its effects on the physical map (Moroldo et al. 2008), each BAC contig contained two sets of NI clones, which aligned to two different regions (Supplementary Material S2). In order to demonstrate that the pairs of sequenced NIs from each contig represented alleles of the same gene and to disregard the possibility that paralogous copies of NIs were duplicated over a short physical distance, the expected segregation in the gametes produced by ‘Cabernet Sauvignon’ was confirmed using 46 offspring. The extent of allelic variation in *VvNII*, -2 and -3 is dependent on the region analysed. Indels were present in the 5'-UTR and in the 3'-UTR of *VvNI3*, and in the introns of *VvNII* and -2. Neither indels nor premature stop codons were found in the exons, which indicate that none of the alleles have structural features that might prevent functionality. SNPs were found at a frequency of 1 SNP every 61, 127 and 75 bp in non-coding regions, and at a frequency of 1 SNP every 287, 185 and 184 bp in the exons of *VvNII*, -2 and -3, respectively. The ratio of non-synonymous substitution rate vs. synonymous rate ( $K_a/K_s$ ) was 0.06 for *VvNI*, 0.24 for *VvNI2* and 0.17 for *VvNI3*, indicating purifying selection for the alleles of ‘Cabernet Sauvignon’. Indels longer than 10 bp were found in the alleles of *VvNII* and -2 in non-coding regions. For instance, two adjacent indels of 36 and 66 bp were found between the *VvNI2* alleles of

**Table 1** Neutral invertase genes identified in grapevine

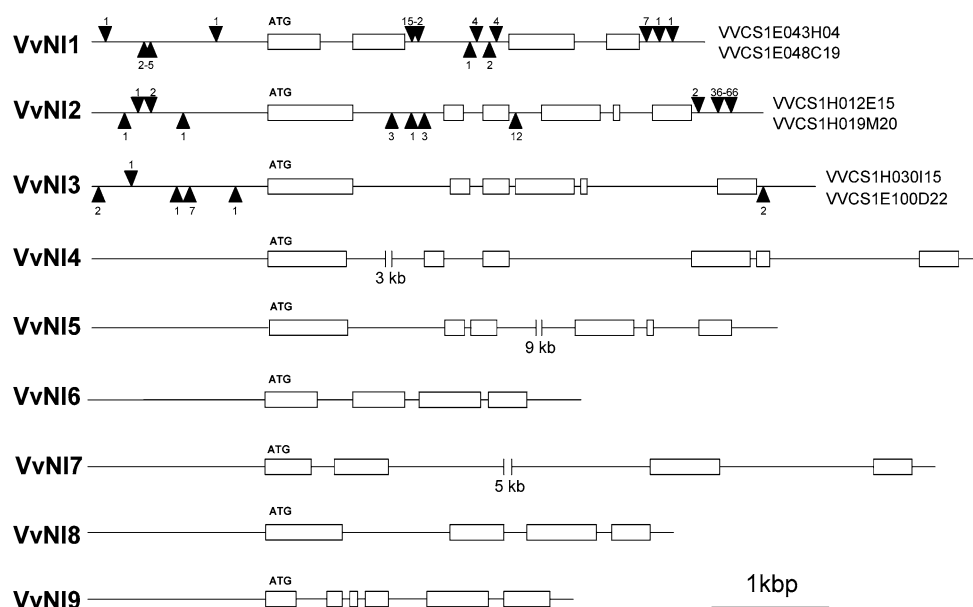
Nickname	Gene model in PN40024 <sup>a</sup>	Chromosomal position <sup>b</sup>	Alleles in 'Cabernet Sauvignon' <sup>c</sup>	The highest identity with ESTs <sup>d</sup>	Score (bits); E-value <sup>e</sup>	At orthologue <sup>f</sup>	Pt orthologue <sup>f</sup>
VVNI1	GSVIVT00036029001	Chr 3: 1,010,600..1,014,009	AM930845-VVCS1E043H04; AM930846-VVCS1E048C19	CTG1033913 (CB969644, CB969682, CB968665, CB968626)	2016; 0.0	At-A/N-InvK	PtNIN9
VVNI2	GSVIVT00033188001	Chr 5: 1,558,376..1,562,437	AM930847-VVCS1H012E15; AM930848-VVCS1H019M20	CTG1034488 (CF403671, CD004479, CD003978, CB036042)	1931; 0.0	At-A/N-InvC	–
VVNI3	GSVIVT00022575001	Chr 14: 668,231..672,442	AM930849-VVCS1H030I15; AM930850-VVCS1E100D22	CTG1035435 (CD006537, CB342019, CB342090)	2260; 0.0	–	PtNIN5
VVNI4	GSVIVT00022057001	Chr 6: 23,629,966..23,638,970	–	CF607582	1011; 0.0	–	–
VVNI5	GSVIVT00034936001	Chr 13: 9,340,522..9,354,127	–	CF602913	490; e-137	At-A/N-InvE	PtNIN3
VVNI6	GSVIVT00014421001	Chr 18: 12,149,471..12,151,638	–	CTG1033913 (CB969644, CB969682, CB968665, CB968626)	496 e-139	–	PtNIN10
VVNI7	GSVIVT00031814001	Chr 18, random: 2,662,729..2,673,261	–	CTG1033913 (CB969644, CB969682, CB968665, CB968626)	252; 4e-66	At-A/N-InvI	PtNIN12
VVNI8	GSVIVT00008639001	Chr Un: 51,283,491..51,286,704	–	–	–	–	PtNIN10
VVNI9	GSVIVT00017350001	Chr 6: 13,175,320..13,177,856	–	CTG1033913 (CB969644, CB969682, CB968665, CB968626)	52; 7e-06	At-A/N-InvI	PtNIN12

<sup>a</sup> <http://www.genoscope.cns.fr/Vitis>
<sup>b</sup> 8.4X assembly at <http://www.genoscope.cns.fr/Vitis>
<sup>c</sup> NCBI accession no.—BAC clone ID

<sup>d</sup> Contig sequences (CTG) are deposited at the UC Davis CGF public database (<http://cgf.ucdavis.edu/home/>), NBI accession numbers are given for individual ESTs (in parenthesis) and for singletons

<sup>e</sup> BLASTn search; threshold  $E < e^{-04}$ 
<sup>f</sup> Orthologues in *Arabidopsis thaliana* (At) and *Populus trichocarpa* (Pt), based on <http://www.genoscope.cns.fr/Vitis> prediction, gene nicknames are reported as defined by Vargas et al. (2003) and Bocoek et al. (2008)





**Fig. 1** Gene structure of nine neutral invertases found in the grapevine genome. NIs were sequenced either in the heterozygous cultivar ‘Cabernet Sauvignon’ (*VvNI1*, -2 and -3) from allelic BAC clones (VVCS1E043H04 and VVCS1E048C19; VVCS1H012E15 and VVCS1H019M20; VVCS1H030I15 and VVCS1E100D22, respectively), or retrieved from the genome assembly of the nearly homozy-

gous line PN40024 (*VvNI4* through -9). Introns are shown as *black lines*, exons are shown as *white boxes*. Indels found in non-coding regions of either allelic sequence are shown as *black triangles* above and underneath the consensus line. The size of each indel is reported as base pairs (bp). SNPs are not indicated in this graph and are discussed in the text. The *scale bar* represents 1 kb

‘Cabernet Sauvignon’ (AM930847-48). The homozygous line PN40024 carried the allele with the deletions (gene model no. GSVIVT00033188001). The heterozygous ‘Pinot noir’ is likely to carry two alleles differing by the indels, like ‘Cabernet Sauvignon’, which caused a chimerical assembly in the corresponding contig of the ‘Pinot noir’ genome sequence (accession no. AM447505.1 nucleotides 3180-9066 reverse complement; Velasco et al. 2007), as assessable by pairwise vs. multiple ClustalX alignments of AM930847, AM930848, GSVIVT00033188001 and AM447505.1.

#### Evolution of the NI gene family across taxa of fully sequenced plants

Three analyses were conducted to determine the relationships among orthologous NIs in fully sequenced species and among paralogous NIs within a species: assessment of the level of amino acid identity, the relative age based on the number of synonymous substitutions per synonymous site ( $K_s$ ), and the colinearity of the gene environment in DNA blocks containing NI copies.

In the first approach, grape NIs were aligned with 9 NIs from *Arabidopsis*, 8 from rice and 16 from poplar, and a NJ tree was constructed (Fig. 2). This dendrogram also served for orthology prediction. Two main groups were identified among grapevine NIs, corresponding to  $\alpha$  and  $\beta$  NIs, which largely agreed with the classification based on exon–intron

structure (Fig. 1). Four NIs *VvNI2*, -3, -4 and -5 were included in the  $\alpha$  clade. Within this group, two branches were identified. The first contained *VvNI2* and -3, which have 75% amino acid identity. *VvNI2* is located on Chr 5 and *VvNI3* is located on Chr 14. The chromosomes in which these two gene copies lie are homoeologous, derived from an ancient event of hybridisation or whole genome duplication (WGD) shared by the major dicot phyla (Jaillon et al. 2007). The pairs of poplar genes *PtNIN1*, -16, and *PtNIN2*, -5, that are orthologous to *VvNI2* and -3 respectively, grouped together in a clade termed  $\alpha 1$ , which was statistically supported by bootstrap analysis. The second clade of the  $\alpha$  group ( $\alpha 2$ ) contained *VvNI4* and -5, which share 75% amino acid identity. *VvNI4* is located on Chr 6 and *VvNI5* is located on Chr 13, which are homoeologous chromosomes. Members of the  $\alpha 1$  clade shared 64–72% identity with the  $\alpha 2$  NIs. The  $\beta$  group included *VvNI1*, -6, -7, -8 and -9. *VvNI1* is located on Chr 3 and is orthologous to the poplar genes *PtNIN9*, -11. *VvNI1* shares 80–82% identity with *VvNI6* and -7. *VvNI6* and -7 are both located on Chr 18, which is homoeologous to Chr 3. Even though placed on the same chromosome, *VvNI6* and -7 did not originate from a recent intra-chromosomal duplication, since their origin is as ancient as that of the  $\alpha$  NI homoeologous copies (see below). The most likely explanation is that a chromosomal rearrangement brought these genes together on the same chromosome. *VvNI9* which apparently showed an  $\alpha$  NI six-exon structure, grouped together with  $\beta$  NIs and



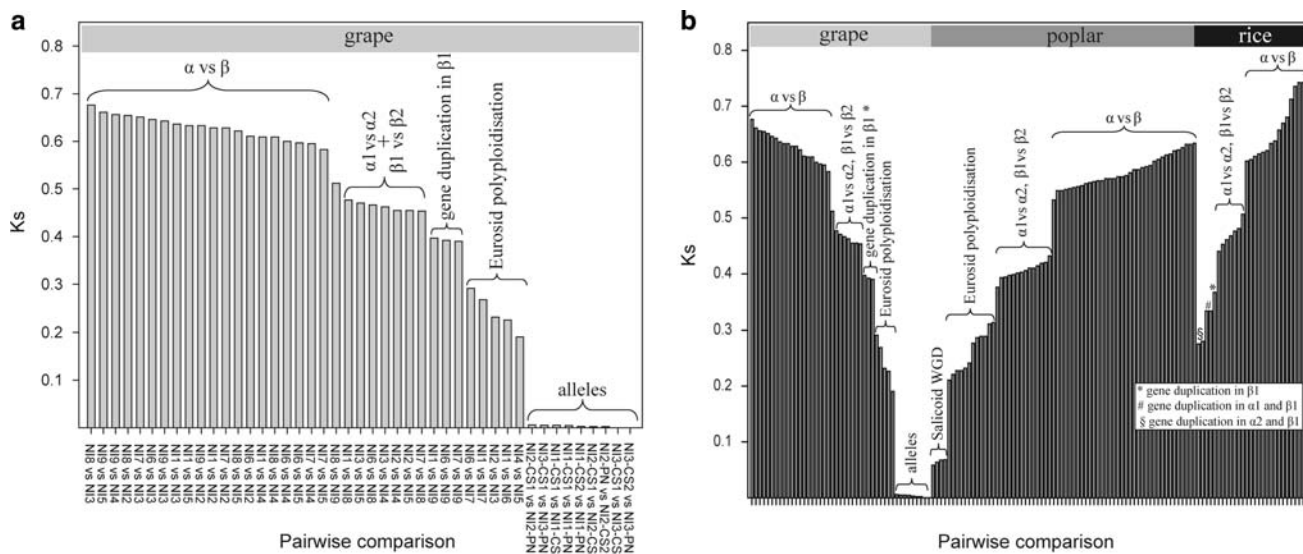


was located on Chr 6, ~10 Mb away from the location of the  $\alpha$  *VvNI4* on the same chromosome. The fifth  $\beta$  NI, *VvNI8*, orthologous to *PtNIN10*, -13, could not be assigned to any chromosome, based on the information of the 8.4X assembly.

The relative age of NI copies was estimated using the number of synonymous substitutions per synonymous site ( $K_s$ ). Pairwise comparison between  $\alpha$  and  $\beta$  NIs gave  $K_s$ -values between 0.58 and 0.67 (Fig. 3a). Similar values were obtained between members of these two clades in poplar and rice (Fig. 3b). Comparison between NIs from clades  $\alpha 1$  and  $\alpha 2$  gave  $K_s$ -values between 0.45 and 0.47. Duplications also occurred within the  $\beta$  group at the same evolutionary time, as witnessed by the  $K_s$ -values of 0.45–0.47 obtained from pairwise comparisons of *VvNI1*, -7 and -8. This led us to define the clades  $\beta 1$  and  $\beta 2$ , the first including *VvNI1*, -6 and -7, and the second including *VvNI8*, unless a weaker statistical support was obtained in the bootstrap analysis, compared with the corresponding  $\alpha 1$  and  $\alpha 2$  clades (Fig. 2). The above mentioned gene expansions within  $\alpha$  and  $\beta$  groups occurred in the common ancestor of flowering plants and were likely generated by large scale duplications. A gene duplication occurred later and specifically in the  $\beta 1$  clade, which generated gene copies with  $K_s$ -values of 0.39–0.40. Since these copies are distinguishable in rice as well ( $K_s = 0.37$  in *OsNIN6* vs. *OsNIN7*), this event most likely occurred before the separation of monocots and dicots, and accounts for the presence of the orthologues

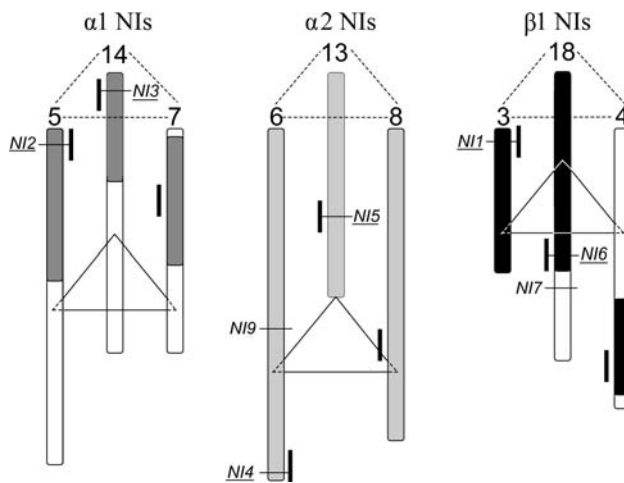
*OsNIN6* and *VvNI9* in the  $\beta 1$  clade of the NJ tree. The NI copies generated by the Eurosid polyploidisation had  $K_s$ -values of 0.19–0.29 and still lie on pairs of grapevine homoeologous chromosomes (Chr 5–14, 6–13, and 3–18, Jaillon et al. 2007). The orthologous genes in poplar (*PtNIN2* -5 vs. *PtNIN1*; *PtNIN3* -4 vs. *PtNIN6*; *PtNIN7* vs. *PtNIN8* -12 vs. *PtNIN9* -11) showed similar  $K_s$ -values. A further gene expansion occurred later in poplar due to the Salicoid WGD, with the new gene copies (*PtNIN2* vs. *PtNIN5*; *PtNIN3* vs. *PtNIN4*; *PtNIN8* vs. *PtNIN12*; *PtNIN9* vs. *PtNIN11*) showing  $K_s$ -values of 0.06–0.07. Neither large scale events nor tandem gene duplications involving NIs occurred after the Eurosid polyploidisation in the grapevine lineage. Two ancient rounds of NI gene duplication occurred in the ancestor of rice after the split of  $\alpha 1$  from  $\alpha 2$  and of  $\beta 1$  from  $\beta 2$ . The duplicated copies had  $K_s$ -values of 0.28 and 0.33, respectively. It is not clear whether they can be ascribed to WGDs followed by extensive gene loss, to a much greater extent than that observed in dicots, or they were caused by small scale duplications.

Since the grape chromosomes group in triplet series, as homoeologous remnants from three ancestral genomes, a blast search for clusters of 20 genes around each NI was performed against the whole proteome, and the chromosomal position of the highest hit was retained (Fig. 4). The gene environment around *VvNI2* on Chr 5 and around *VvNI3* on Chr 14 was colinear within a DNA block of ~40 kb on Chr 7 between the positions 4,737,129 and



**Fig. 3** Relative age of neutral invertase genes and alleles in the grapevine genome (a) and comparison with poplar and rice genomes (b). Relative age was estimated using the number of synonymous substitutions per synonymous site ( $K_s$ ). In the x-axis of (a), genes *VvNI1* to -9 were abbreviated to NI1 to -9, alleles sequenced in the heterozygous ‘Cabernet Sauvignon’ (CS1 and CS2) and in the quasi-homozygous

line PN40024 (PN) are tagged with the corresponding abbreviations. Gene identifiers are not reported in the x-axis of (b) due to space constraints; genes within groups identified by pairwise comparison, as reported in this figure, are cross-linked to genes within groups identified by amino acid sequence identity and reported with the complete identifiers in Fig. 2



**Fig. 4** Genome location of  $\alpha 1$ ,  $\alpha 2$ ,  $\beta 1$  NIs (underlined) and other  $\beta$  NIs on three triplets of homoeologous chromosomes. Gene nomenclature *VvNI1* to -9 used in the text was abbreviated to *NI1* to -9 in this graph. The chromosome number is reported on top of the diagram of the chromosome; chromosomes are scaled and oriented as in Jaillon et al. (2007). The homoeologous portions within each chromosome triplet are drawn in light grey, dark grey and black, and are defined as reported in Jaillon et al. (2007). The position of each *NI* gene is indicated by a thin horizontal line, and colinearity of the gene environment around NIs and within each triplet is indicated by thick vertical lines

4,777,926. In a similar way, genes around *VvNI4* on Chr 6 and around *VvNI5* on Chr 13 scored the highest hits with genes confined within a 173-kb interval from position 16,511,172 to 16,684,133 on Chr 8. Yet, the genes in the surroundings of *VvNI1* on Chr 3 and of *VvNI6* on Chr 18 matched a block of genes with the same relative order on the 156-kb interval from position 14,713,273 to 14,869,211 on Chr 4. While this analysis confirmed the homoeologous nature of the  $\alpha 1$  pair *VvNI2*, -3, the  $\alpha 2$  pair *VvNI4*, -5, and the  $\beta 1$  pair *VvNI1*, -6, consistent with what was suggested by their  $K_s$ -values, it also identified the third homoeologous region where the third *NI* copy of each homoeologous series was expected to reside. Neither a complete open reading frame in the PN40024 gene prediction nor even a gene remnant showing a detectable similarity with NIs was found by BLAST in the homoeologous intervals on Chr 4, 7 and 8. In contrast, some of the surrounding genes assigned to gene categories different from NIs showed colinearity. Neither scaffold junctions nor major sequence gaps occurred in these intervals. These features might indicate that either the set of chromosomes including Chr 4, 7 and 8 were derived from a component genome with a lower level of gene conservation, or one copy for each of the  $\alpha 1$ ,  $\alpha 2$  and  $\beta 1$  NIs was randomly lost across the three constituent genomes. The genes and the regions homoeologous to *VvNI7*, -8 and -9 could not be found by this approach, as if those copies were lost and the regions in which they had resided underwent extensive rearrangements.

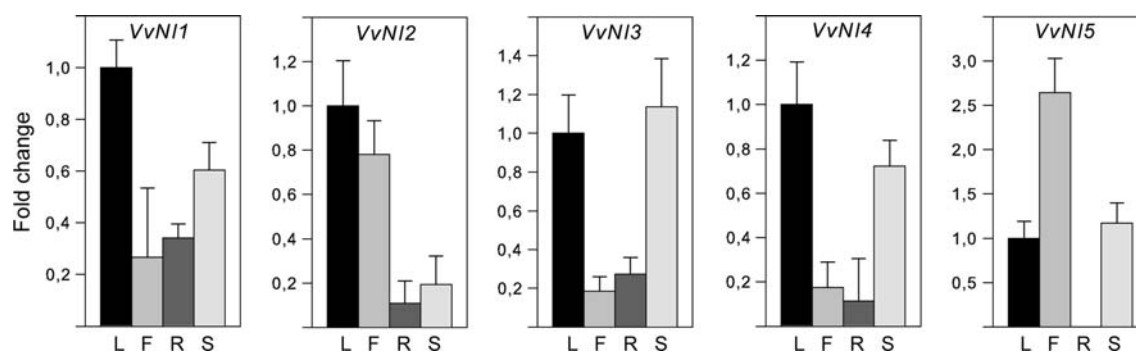
#### Predicted cellular targeting of NIs and sequence variation in the region upstream of the start codon

The grapevine  $\alpha 2$  *VvNI4* and -5 clustered with *OsNIN3* and *At-A/N-InvE* over the dendrogram of amino acid similarity (Fig. 2). The corresponding proteins in rice and *Arabidopsis* are localised in the chloroplast (Murayama and Handa 2007; Vargas et al. 2008). In silico prediction of sub-cellular targeting based on motifs in the N-terminal region also suggested a chloroplast localisation for *VvNI4*, while the predictions for *VvNI5* were inconsistent when different algorithms were used. The grapevine  $\alpha 1$  *VvNI2* and -3 clustered with *OsNIN1*, a rice gene coding for a mitochondrial *NI*, and with *At-A/N-InvA* and *At-A/N-InvC*, for which Murayama and Handa (2007) predicted high likelihood for mitochondrial localisation. However, none of the predictors used in this study supported the presumed mitochondrial localisation for the grapevine orthologues included in the  $\alpha 1$  clade. Finally, a cytosolic localisation of  $\beta$  *VvNI1* and -6 was likely, based on an in silico prediction.

Excess copies created by gene duplication retain functionality, provided that their coding sequences are under purifying selection, and may acquire specialisation by accumulating mutations in the promoter regions. Functional isogenes that underwent sub-functionalisation, i.e. divergent expression across organs and/or differential responses to stimuli, should thus harbour distinct regulatory elements. We searched for *cis*-element motifs in the 1.5 kb region upstream of the start codon using the PlantCARE database. The abundance of motifs associated with responses to endogenous stimuli (hormones and signalling specimens), external stresses and light is given in Supplementary Material S3. The number and the type of regulatory elements varied extensively across the nine *VvNIs*. The  $\alpha 1$  *VvNI2* and -3 had the highest number of known *cis*-elements. Among the  $\beta$  NIs, *VvNI7* had the lowest number of responsive motifs, while *VvNI8* was particularly rich in motifs sensitive to external stimuli. Motifs for tissue-specific expression were found, such as those conferring meristem (*VvNI2*, -3) and endosperm specificity (all but *VvNI7*). A transcriptional analysis of five out of the nine NIs, for which evidence of expression came from EST databases and which included pairs of homoeologous genes, was carried out in order to verify whether motif variation in their promoter regions resulted in differences in expression across organs or during berry ripening.

#### Expression patterns of homoeologous copies of NIs

Most *VvNIs* were expressed across all leaf, flower, root and seed tissues, but with an organ-specific regulation (Fig. 5). *VvNI1* was expressed with the highest relative abundance in leaf. *VvNI2* expression was much higher in leaf and flower



**Fig. 5** Expression of five neutral invertases in leaf (L), flower (F), root (R) and seed (S) tissues. Transcript levels are expressed as fold changes relative to a calibrator tissue, in this case the leaf. In each tissue, transcripts of each NI paralogue were amplified using specific primer

pairs (Supplementary Material S1) and normalised using the constitutive gene coding for the ubiquitin conjugating enzyme. Bars represent standard deviation of three technical replicates

than in root and seed. *VvNI3* and *VvNI4* transcripts were detected at higher levels in leaf and seed than in flower and root. *VvNI5* transcripts were more abundantly expressed in flower than in leaf and seed, and were not detected in root. In berry, gene expression of *VvNIs* was evaluated more in detail during the period of fruit development.

Fruit growth was monitored from 45 DAFB until harvest (118 DAFB) in the season of 2005. The growth curve was consistent with a double-sigmoid pattern (Fig. 6a). Berry growth entered the lag phase at 59 DAFB. Increase of berry weight resumed at 72 DAFB, and then grew progressively with ripening. Glucose and fructose content were measured separately in flesh and in berry skin. Sugars were barely detectable in flesh of green berries at 45 DAFB (Fig. 6b). A modest but detectable increase in sugar content occurred in green berries during the lag phase from 59 to 72 DAFB, before anthocyanin accumulation had begun. This increase was measured in both flesh and skin. Sugar concentration was substantially higher in flesh of red berries compared with green berries sampled on the same day over the same clusters (Fig. 6b). This difference was less evident in skin (Fig. 7a).

In the fruit flesh after the onset of maturation, the level of expression of all *NI* members was generally low (Fig. 6c). *VvNI2* and -3 were expressed slightly more than the others in terms of relative level during ripening. As for the trends, *VvNI3*, -4 and -5 were weakly expressed in green berries until 59 DAFB. Then, *VvNI3*, -4 and -5 were up-regulated at 66 DAFB, in concomitance with the end of the lag phase of growth and with the rapid increase of sugar concentration, which manifested itself in a discontinuous step in the curve of sugar accumulation. *VvNI3* and -4 also showed a continuous trend of increased expression in the later stage of ripening. Transcripts of *VvNI1* and -2 were more abundant in the flesh of pre-*véraison* berries and then levelled off.

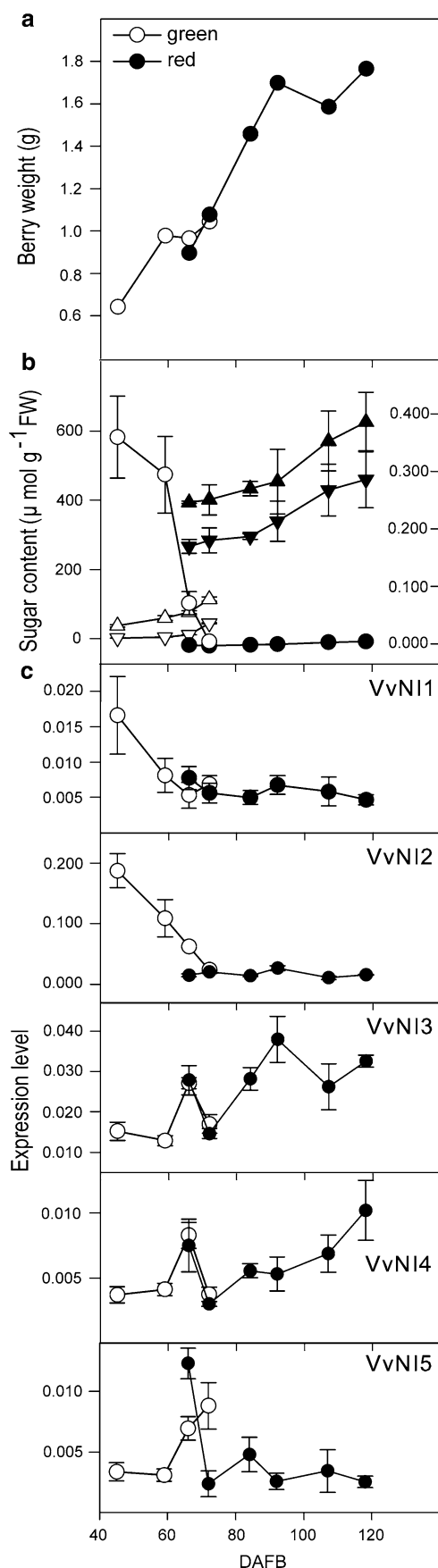
In order to compare transcriptional levels of *NIs*, sugar accumulation and anthocyanin biosynthesis, gene expres-

sion was also assayed in skin (Fig. 7). In terms of trends, expression mirrored the patterns observed in the mesocarp, though the transcript levels in skin were higher in absolute terms. At *véraison*, *VvNI2* expression decreased concomitantly with the onset of maturation while *VvNI3*, -4 and -5 transcripts increased at the same time. Expression of *VvNI3* and -4 remained activated during ripening, then declined as fruit approached full maturity.

As for the homoeologous copies, *VvNI2* and -3 showed a divergent temporal regulation in the same tissue during fruit development, with *VvNI2* more transcribed at the pre-*véraison* stage than during ripening in both skin and pulp (and *VvNI3* vice versa). The other homoeologous pair, *VvNI4* and -5 showed more similar patterns in fruit. Divergent expression between both pairs of homoeologous copies was also evident by comparing different organs (Fig. 5).

## Discussion

Nine *NIs* are present in the grapevine genome. *NIs* are located on at least six different chromosomes. One gene copy of  $\alpha 1$  *NIs*, one gene copy of  $\alpha 2$  *NIs* and more than one copy of  $\beta$  *NIs* of the present-day grapevines already existed in the progenitors before the Eurosid polyploidisation. One additional copy for each of the ancestral gene members was retained on the homoeologous chromosomes generated by polyploidy. As three component genomes contributed to the grapevine haploid content (Olmo 1976; Jaillon et al. 2007), the third *NI* copy of the homoeologous series that include either *VvNI1* -6, *VvNI2* -3 or *VvNI4* -5 was searched for, but not found. We were able to identify the third region homoeologous to each of the DNA segments containing the gene pairs *VvNI1* -6; *VvNI2* -3 or *VvNI4* -5, by looking at the conservation and colinearity of gene environment around each *NI*. Although this analysis confirmed the homoeologous nature of the *NI* copies within the  $\alpha 1$ ,  $\alpha 2$  and  $\beta 1$

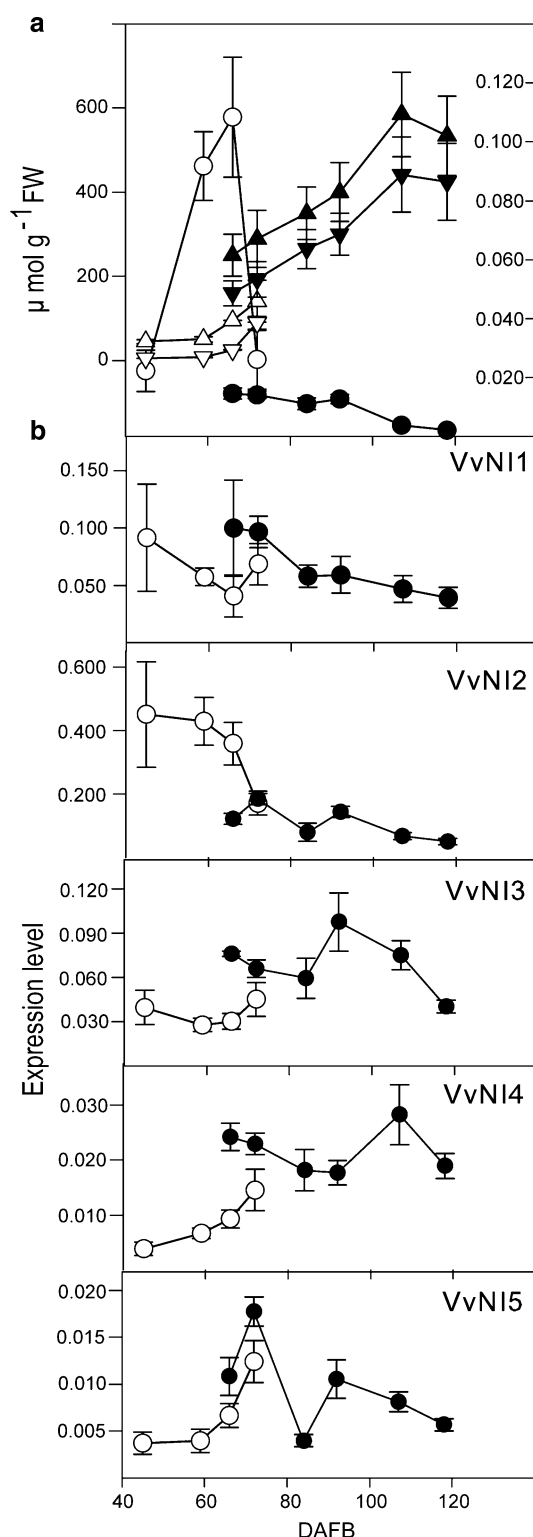


**Fig. 6** **a** Pattern of fruit growth, expressed as berry weight, from 45 to 118 days after full blooming (DAFB). **b** Fructose (*inverted triangle*) and glucose (*triangle*) concentration in the mesocarp of green (*open symbols*) and red berries (*solid symbols*) are scaled to the y-axis on the left. The expression of the brassinosteroid 6-oxidase gene *VvBR6OX1* (Symons et al. 2006) was reported as a molecular marker of the onset of maturation (*open and solid circle*), and it is referenced to the y-axis on the right hand side of the graph as arbitrary units. **c** Expression of five neutral invertases in the mesocarp. Transcript levels are expressed as arbitrary units, normalised using the constitutive gene coding for the ubiquitin conjugating enzyme. Bars represent standard deviation of three technical replicates

clades, in all three cases the expected third copy of each *NI* triplet was not present within the homoeologous block of DNA. The possible explanations are that either a copy was randomly deleted, or it early mutated into a pseudogene then becoming no longer distinguishable from junk DNA, or the third component genome was much more diverse than the others. After the Eurosid polyploidisation, the whole gene family of  $\alpha$  and  $\beta$  *NIs* did not undergo any expansion through tandem duplication, neither in grape nor in poplar. The increased number of *NIs* in poplar is solely explained by the recent WGD which occurred in the Salicaceae lineage (Tuskan et al. 2006). The *NI* story in the monocotyledonous rice is probably more intricate. The chromosomal dispersion of *NIs* (Franki et al. 2006) and the relative age of the gene members indicate an evolutionary scenario similar to that of the dicots analysed in this paper, with a few ancient copies and absence of recent proliferation by tandem duplication, but with more extensive rearrangement and gene loss.

Despite the lack of functional studies, the pairs of homoeologous  $\alpha$  *NIs* have maintained their potential functionality in the grape plant, based on the intact exon–intron structure, the retention of the amino acid domains conserved across plant *NIs*, and the evidence of transcription. In order to assess whether homoeologous genes underwent expression divergence, we analysed the transcriptional levels of homoeologous copies of  $\alpha 1$  and  $\alpha 2$  *NIs* in different organs of the grape plant at a single development stage and in different tissues of the grape berry during the course of ripening. The expression pattern of every gene member was spatially or temporally governed in a manner that suggests some member-specificity. Casneuf et al. (2006) showed that gene fate after duplication depends upon the functional category of the duplicated genes and the type of duplication. Genes involved in signal transduction, carbohydrate metabolism and hydrolase activity were preferentially retained after WGD events, and quickly diverged in their expression pattern in *Arabidopsis*. The patterns observed in grapevine *NIs* strengthens these general observations made in other plants for genes duplicated by polyploidy (described and reviewed in Adams et al. 2003; Blanc and





**Fig. 7** **a** Fructose (inverted triangle) and glucose (triangle) concentration in the skin of green (white symbols) and red berries (black symbols). Concentrations are referred to the y-axis on the left. The expression of the brassinosteroid 6-oxidase gene *VvBR6OX1* (Symons et al. 2006) was reported as a molecular marker of the onset of maturation (open and solid circle), and it is referenced to the y-axis on the right hand side of the graph as arbitrary units. **b** Expression of five neutral invertases in berry skin. Transcript levels are expressed as arbitrary units, normalised using the constitutive gene coding for the ubiquitin conjugating enzyme. Bars represent standard deviation of three technical replicates

ogous NIs within a clade. For instance,  $\alpha 1$  NIs differed from  $\alpha 2$  NIs for the presence of several abscisic acid responsive elements and the absence of ethylene-responsive motifs. Abscisic acid and ethylene are at the core of hormonal control of ripening in grape berry, and may establish links between sugar and hormone signalling (Leon and Sheen 2003). Within a clade, *VvNI2* had motifs implicated in response to gibberellin and wounding, which were absent in the homoeologous *VvNI3*. Vice versa, *VvNI3* had motifs responsive to auxin, heat stress and low temperatures. The abundance of *cis*-elements that are unique to each gene member might provide molecular explanations for the observed expression divergence.

Our work contributes to a wealth of information that claims a role for NIs in basal metabolism. Several clues suggest that some, if not all, NI members might be involved in day-to-day processes of cell maintenance. Most of the gene copies are transcribed at low levels in most organs and developmental stages in grape. Orthologues of NIs between monocots and dicots are still easily recognisable when amino acid sequences from grapevine, poplar, *Arabidopsis* and rice are compared. Additional gene members were largely retained after the Eurosid polyploidisation, most of the additional members generated by the Salicoid WGD did not lose functionality in poplar (Bocock et al. 2008), and the alleles of *VvNI1*, -2 and -3 sequenced in grapevine showed  $K_a/K_s$  ratios predictive of purifying selection. All of this suggests that the ancient copies present in the progenitor of flowering plants were not dispensable, and the new copies generated by WGDs provided some evolutionary advantage, which led to their retention. Although it seems that more is better, NIs have not expanded by tandem duplication either in grapevine or in the other sequenced plants. This is in contrast with the fate of genes involved in secondary metabolism, such as flavonoid, terpene and stilbene biosynthetic pathways, which have extensively expanded by local duplications in grapevine (Castellarin et al. 2006; Jaillon et al. 2007).

Genes involved in sugar metabolism, including the genes encoding enzymes of sucrose hydrolysis, have a dual role in energy balance and in regulation of fruit growth (Koch 2004). They ultimately determine yield and

Wolfe 2004; Haberer et al. 2004; Adams and Wendel 2005). Indeed, mutations that accumulate in *cis*-elements upstream of the coding region are the molecular basis for sub-functionalisation of additional gene copies. Responsive motifs that were predicted in silico by PlantCARE varied between NIs of different clades and also between homoeol-



quality of the crop in fruit trees. In the grape berry, sucrose cleavage is mainly attributed to the expression and the activity of cell wall invertases and sucrose synthases (Zhang et al. 2006; Deluc et al. 2007) or vacuolar invertases (Davies and Robinson 1996). In contrast, the implication of NIs in fruit development has remained in the background. NIs have long been thought to regulate sucrose metabolism in mature internodes of sugar cane, in sugar beet and carrot roots (reviewed in Ranwala et al. 1991) and in peach fruit (Nonis et al. 2007), which all store sucrose. The grapevine NIs, for which expression profiles were studied in this paper, are somewhat regulated at transcriptional level in the mesocarp and exocarp of ripening berries that accumulate glucose and fructose. Changes in transcript levels were coincident with the beginning of hexose accumulation. These data provide further hints that some NIs might fit into the network of sugar accumulation, sensing and signalling, and not solely act as housekeeping genes. While it was expected that growing tissues and storage organs require sucrose mobilisation, which might explain the contribution of the activity of some NIs in hexose release, the significance of the simultaneous expression of multiple gene copies in a single tissue remains obscure. If one accepts Murayama and Handa's (2007) assumption that  $\alpha 1$  NIs are localised in the mitochondria,  $\alpha 2$  NIs in the chloroplast and  $\beta$  NIs in the cytosol, it is tempting to hypothesise that transcripts of at least one NI gene copy with predicted organelle targeting increased in the grape berry at *véraison*. In the view of the recent findings that assign a non-sucrolytic role to sucrose synthases located in the mitochondria (Subbaiah et al. 2006), some NIs might function similarly in sensing and regulating sugar transport in these organelles.

**Acknowledgements** We thank S.D. Castellarin (University of Udine) for sampling 'Merlot' berries, J.-M. Aury (Genoscope, Paris) for help in the identification of homoeologous regions, R. Marconi (University of Udine) and I. LeClainche (INRA) for help in pooling and screening of BAC clones, A. Fiori (University of Udine) for SSCP analysis, F. Cattonaro and I. Jurman (Istituto di Genomica Applicata, Udine) for sequencing NI alleles in 'Cabernet Sauvignon', E. De Paoli (University of Udine) for advice in the use of Phred and Phrap software and C. Coleman for proofreading the manuscript.

## References

- Adam-Blondon AF, Bernole A, Faes G, Lamoureux D, Pateyron S, Grando MS, Caboche M, Velasco R, Chalhou B (2005) Construction and characterization of BAC libraries from major grapevine cultivars. *Theor Appl Genet* 110:1363–1371
- Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. *Curr Opin Plant Biol* 8:135–141
- Adams KL, Cronn R, Percifield R, Wendel JF (2003) Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc Natl Acad Sci USA* 100:4649–4654

- Blanc G, Wolfe KH (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* 16:1679–1691
- Bocock PN, Morse AM, Dervinis C, Davis JM (2008) Evolution and diversity of invertase genes in *Populus trichocarpa*. *Planta* 227:565–576
- Bosch S, Grof CPL, Botha FC (2004) Expression of neutral invertase in sugarcane. *Plant Sci* 166:1125–1133
- Casneuf T, De Bodt S, Raes J, Maere S, Van de Peer Y (2006) Non-random divergence of gene expression following gene and genome duplications in the flowering plant *Arabidopsis thaliana*. *Genome Biol* 7:13
- Castellarin SD, Di Gaspero G, Marconi R, Nonis A, Peterlunger E, Paillard S, Adam-Blondon A-F, Testolin R (2006) Colour variation in red grapevines (*Vitis vinifera* L.): genomic organisation, expression of flavonoid 3'-hydroxylase, flavonoid 3', 5'-hydroxylase genes and related metabolite profiling of red cyanidin/blue delphinidin-based anthocyanins in berry skin. *BMC Genomics* 7:12
- Davies C, Robinson SP (1996) Sugar accumulation in grape berries. Cloning of two putative vacuolar invertase cDNAs and their expression in grapevine tissues. *Plant Physiol* 111:275–283
- Deluc LG, Grimplet J, Wheatley MD, Tillett RL, Quilici DR, Osborne C, Schooley DA, Schlauch KA, Cushman JC, Cramer GR (2007) Transcriptomic and metabolite analyses of Cabernet Sauvignon grape berry development. *BMC Genomics* 8:429
- Di Gaspero G, Cipriani G, Adam-Blondon AF, Testolin R (2007) Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for R-gene candidates. *Theor Appl Genet* 114:1249–1263
- Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* 8:978–984
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300:1005–1016
- Ewing B, Hillier L, Wendl M, Green P (1998) Basecalling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8:175–185
- Francki MG, Walker E, Forster JW, Spangenberg G, Appels R (2006) Fructosyltransferase and invertase genes evolved by gene duplication and rearrangements: rice, perennial ryegrass, and wheat gene families. *Genome* 49:1081–1091
- Gallagher JA, Pollock CJ (1998) Isolation and characterization of a cDNA clone from *Lolium temulentum* L. encoding for a sucrose hydrolytic enzyme which shows alkaline/neutral invertase activity. *J Exp Bot* 49:789–795
- Gibson S (2005) Control of plant development and gene expression by sugar signalling. *Curr Opin Plant Biol* 8:93–102
- Haberer G, Hindemitt T, Meyers BC, Mayer KF (2004) Transcriptional similarities, dissimilarities, and conservation of *cis*-elements in duplicated genes of *Arabidopsis*. *Plant Physiol* 136:3009–3022
- Horton P, Park K-J, Obayashi T, Nakai K (2006) Protein subcellular localization prediction with WoLF PSORT. In: Proceedings of the 4th annual Asia Pacific bioinformatics conference APBC06, Taipei, Taiwan, pp 39–48
- Iandolino AB, Goes da Silva F, Lim H, Choi H, Williams LE, Cook DR (2004) High-quality RNA, cDNA, and derived EST libraries from grapevine (*Vitis vinifera* L.). *Plant Mol Biol Rep* 22:269–278
- Jaillon O, Aury JM, Noel B, Policriti A, Clepet C et al (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449:463–467
- Ji XM, Van den Ende W, Van Laere A, Cheng SH, Bennett J (2005) Structure, evolution, and expression of the two invertase gene families of rice. *J Mol Evol* 60:615–634

- Koch K (2004) Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Curr Opin Plant Biol* 7:235–246
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5:150–163
- Leon P, Sheen J (2003) Sugar and hormone connections. *Trends Plant Sci* 8:110–116
- Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res* 30:325–327
- Lou Y, Gou JY, Xue HW (2007) PIP5K9, an *Arabidopsis* phosphatidylinositol monophosphate kinase, interacts with a cytosolic invertase to negatively regulate sugar-mediated root growth. *Plant Cell* 19:163–181
- Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* 452:991–996
- Moroldo M, Paillard S, Marconi R, Legeai F, Canaguier A, Cruaud C, De Berardinis V, Guichard C, Brunaud V, Le Clainche I, Scalabrín S, Testolin R, Di Gaspero G, Morgante M, Adam-Blondon AF (2008) A physical map of the heterozygous grapevine ‘Cabernet Sauvignon’ allows mapping candidate genes for disease resistance. *BMC Plant Biol* 8:66
- Murayama S, Handa H (2007) Genes for alkaline/neutral invertase in rice: alkaline/neutral invertases are located in plant mitochondria and also in plastids. *Planta* 225:1193–1203
- Nonis A, Ruperti B, Falchi R, Casatta E, Thamashebi SE, Vizzotto G (2007) Differential expression and regulation of a neutral invertase encoding gene from peach (*Prunus persica*): evidence for a role in fruit development. *Physiol Plant* 129:436–446
- Olmo HP (1976) Grapes. In: Simmonds NW (ed) *Evolution of crop plants*. Longman, London, pp 249–298
- Qi X, Wu Z, Li J, Mo X, Wu S, Chu J, Wu P (2007) AtCYT-INV1, a neutral invertase, is involved in osmotic stress-induced inhibition on lateral root growth in *Arabidopsis*. *Plant Mol Biol* 64:575–587
- Ranwala AP, Iwanami S-S, Masuda H (1991) Acid and neutral invertases in the mesocarp of developing muskmelon (*Cucumis melo* L. cv Prince) fruit. *Plant Physiol* 96:881–886
- Roitsch T, Gonzalez M-C (2004) Function and regulation of plant invertases: sweet sensations. *Trends Plant Sci* 9:606–613
- Rossouw D, Bosch S, Kossmann J, Botha FC, Groenewald J-H (2007) Downregulation of neutral invertase activity in sugarcane cell suspension cultures leads to a reduction in respiration and growth and an increase in sucrose accumulation. *Funct Plant Biol* 34:490–498
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497
- Sheen J, Zhou L, Jang J-C (1999) Sugars as signalling molecules. *Curr Opin Plant Biol* 2:410–418
- 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
- Smeekens S (2000) Sugar-induced signal transduction in plants. *Annu Rev Plant Physiol Plant Mol Biol* 51:49–81
- Sturm A, Tang G-Q (1999) The sucrose-cleaving enzymes of plants are crucial for development, growth and carbon partitioning. *Trends Plant Sci* 4:401–407
- Sturm A, Hess D, Lee HS, Lienhard S (1999) Neutral invertase is a novel type of sucrose-cleaving enzyme. *Physiol Plant* 107:159–165
- Subbiah CC, Palaniappan A, Duncan K, Rhoads DM, Huber SC, Sachs MM (2006) Mitochondrial localization and putative signaling function of sucrose synthase in maize. *J Biol Chem* 281:15625–15635
- Symons GM, Davies C, Shavrukov Y, Dry IB, Reid JB, Thomas MR (2006) Grapes on steroids. Brassinosteroids are involved in grape berry ripening. *Plant Physiol* 140:150–158
- Terrier N, Glissant D, Grimplet J, Barrieu F, Abbal P, Couture C, Ageorges A, Atanassova R, Léon C, Renaudin JP, Dédaldéchamp F, Romieu C, Delrot S, Hamdi S (2005) Isogene specific oligo arrays reveal multifaceted changes in gene expression during grape berry (*Vitis vinifera* L.) development. *Planta* 222:832–847
- Initiative TheArabidopsisGenome (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815
- Thomas TR, Matthews MA, Shackel KA (2006) Direct in situ measurement of cell turgor in grape (*Vitis vinifera* L.) berries during development and in response to plant water deficits. *Plant Cell Environ* 29:993–1001
- Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I et al (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313:1596–1604
- Tymowska-Lalanne Z, Kreis M (1998) The plant invertases: physiology, biochemistry and molecular biology. *Adv Bot Res* 28:72–117
- Vargas W, Cumino A, Salerno GL (2003) Cyanobacterial alkaline/neutral invertases. Origin of sucrose hydrolysis in the plant cytosol? *Planta* 216:951–960
- Vargas WA, Pontis HG, Salerno GL (2008) New insights on sucrose metabolism: evidence for an active A/N-Inv in chloroplasts uncovers a novel component of the intracellular carbon trafficking. *Planta* 227:795–807
- Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A et al (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS ONE* 2(12):e1326
- Vizzotto G, Pinton R, Varanini Z, Costa G (1996) Sucrose accumulation in developing peach fruit. *Physiol Plant* 96:225–230
- Wada H, Shackel KA, Matthews MA (2008) Fruit ripening in *Vitis vinifera*: apoplastic solute accumulation accounts for pre-*véraison* turgor loss in berries. *Planta* 227:1351–1361
- Winter H, Huber SC (2000) Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. *Crit Rev Plant Sci* 19:31–67
- Zhang XY, Wang XL, Wang XF, Xia GH, Pan QH, Fan RC, Wu FQ, Yu XC, Zhang DP (2006) A shift of phloem unloading from symplasmic to apoplastic pathway is involved in developmental onset of ripening in grape berry. *Plant Physiol* 142:220–232