Theor Appl Genet (2014) 127:2433–2451
DOI 10.1007/s00122-014-2388-6

ORIGINAL PAPER

Multiple loss-of-function *5-O-glucosyltransferase* alleles revealed in *Vitis vinifera*, but not in other *Vitis* species

Yingzhen Yang · Joanne A. Labate · Zhenchang Liang · Peter Cousins · Bernard Prins · John E. Preece · Mallikarjuna Aradhya · Gan-Yuan Zhong

Received: 27 March 2014 / Accepted: 23 August 2014 / Published online: 11 September 2014

Abstract

Key message **Wild and loss-of-function alleles of the *5-O-glucosyltransferase* gene responsible for synthesis of diglucoside anthocyanins in *Vitis* were characterized. The information aids marker development for tracking this gene in grape breeding.**

Abstract Anthocyanins in red grapes are present in two glycosylation states: monoglucoside (3-*O*-glucoside) and diglucoside (3, 5-di-*O*-glucoside). While monoglucoside anthocyanins are present in all pigmented grapes, diglucoside anthocyanins are rarely found in the cultivated grape species *Vitis vinifera*. Biochemically 3-*O*-glucoside anthocyanins can be converted into 3,5-di-*O*-glucoside anthocyanins by a 5-*O*-glucosyltransferase. In this study, we surveyed allelic variation of the *5-O-glucosyltransferase* gene (*5GT*) in 70 *V. vinifera* ssp. *vinifera* cultivars, 52 *V. vinifera* ssp. *sylvestris* accessions, 23 *Vitis* hybrid grapes, and 22 accessions of seven other *Vitis* species. Eighteen *5GT* alleles with apparent loss-of-function mutations, including seven premature stop codon mutations and six frameshift indel mutations, were discovered in *V. vinifera*, but not in the other *Vitis* species. A total of 36 *5GT* alleles without apparent loss-of-function mutations (W-type) were identified. These W-type alleles were predominantly present in wild *Vitis* species, although a few of them were also found in some *V. vinifera* accessions. We further evaluated some of these *5GT* alleles in producing diglucoside anthocyanins by analyzing the content of diglucoside anthocyanins in a set of representative *V. vinifera* cultivars. Through haplotype network analysis we revealed that *V. vinifera* ssp. *vinifera* and its wild progenitor *V. vinifera* ssp. *sylvestris* shared many loss-of-function *5GT* alleles and extensive divergence of the *5GT* alleles was evident within *V. vinifera*. This work advances our understanding of the genetic diversity of *5GT* and provides a molecular basis for future marker-assisted selection for improving this important wine quality trait.

Communicated by Reinhard Toepfer.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-014-2388-6) contains supplementary material, which is available to authorized users.

Y. Yang · P. Cousins · G.-Y. Zhong (✉)
USDA-ARS Grape Genetics Research Unit, Geneva, NY 14456, USA
e-mail: ganyuan.zhong@ars.usda.gov

J. A. Labate · G.-Y. Zhong
USDA-ARS Plant Genetic Resources Unit, Geneva, NY 14456, USA

Z. Liang
Institute of Botany, The Chinese Academy of Sciences, Beijing 100093, People's Republic of China

Present Address:
P. Cousins
E. & J. Gallo Winery, Modesto, CA 95353, USA

B. Prins · J. E. Preece · M. Aradhya
USDA National Clonal Germplasm Repository, UCD, One Shields Avenue, Davis, CA 95616-8683, USA

Introduction

Grapes are widely cultivated for fresh fruits as well as for processed products such as wine, juice, raisins, and jam. Grape berries are rich in polyphenolic compounds providing many health and nutritional benefits to consumers. About 60–87 % of the total polyphenolics in colored grape berries are anthocyanins (Liang et al. 2011, 2012),

whose backbone is anthocyanidin. In grape, anthocyanidins can be stabilized by glycosylation at the 3 or 3 and 5 positions of the C ring to form 3-*O*-glucoside anthocyanin or 3,5-*O*-diglucoside anthocyanin, respectively. While glucose at the 3 position is added by UDP-glucose: flavonoid 3-*O*-glucosyltransferase (Ford et al. 1998), 5-*O*-glucosyltransferase is responsible for adding a glucose at the 5 position (Janvary et al. 2009).

Anthocyanins are the main contributors to wine color and they exist in both monomeric and polymeric forms by forming adducts with different chemicals in wine. Monoglucoside and diglucoside anthocyanins behave differently in wine chemistry due to their structural differences. Wines with monoglucoside anthocyanins were less color stable to heat and light than wines with diglucoside pigments (Kim et al. 2010; Robinson et al. 1966; Vanburen et al. 1968). While 5-*O*-glucoside in the diglucoside anthocyanin has dual effects on the equilibrium between color and colorless forms and stability against oxidation/heat/light (Bishop and Nagel 1984; Garcia-Viguera and Bridle 1999; Mazza and Brouillard 1987; Sims and Morris 1985), diglucoside anthocyanins are highly undesirable in wine maturation and aging chemistry. Studies on muscadine grape (*V. rotundifolia*) wines and European red wines (mostly made from *V. vinifera*) provided support for the dual roles of the 5-*O*-glucoside in wine chemistry. Compared to wines made from *V. vinifera* var. Cabernet Sauvignon, wines made from muscadine grapes started with a better color, but they were very susceptible to browning and overall loss of color quality during processing and storage (Sims and Morris 1985). The fact that the phenolic profile of most muscadine wines was less complex compared to *vinifera* wines further suggested that the anthocyanins in muscadine wines were unable to form many complex chemicals such as those found in *vinifera* wines (Sims and Morris 1985, 1986). The different behaviors of monoglucoside and diglucoside anthocyanins in wine chemistry might contribute to the observation that wines made from hybrid grapes are often considered inferior to those made from *V. vinifera* of which most do not produce any diglucoside anthocyanins (Liang et al. 2011).

The contents of diglucoside anthocyanins in grape berries vary tremendously among and within different *Vitis* species. While all anthocyanins in the muscadine grapes are diglucosylated (Ballinger et al. 1973; Huang et al. 2009; Sandhu and Gu 2010; You et al. 2012), most cultivated *V. vinifera* and its wild progenitor, *V. vinifera* ssp. *sylvestris*, produce berries with no detectable diglucoside anthocyanins (Liang et al. 2011; Revilla et al. 2010). On the other hand, North American wild *Vitis* species such as *V. labrusca*, *V. rupestris* and *V. riparia* produce both monoglucoside and diglucoside anthocyanins, with the latter comprising 30–60 % of total anthocyanins (Liang et al. 2011; Mazza 1995). Some wild Asian *Vitis* species such as *V. amurensis* and *V. coignetiae* contain much higher amounts of diglucoside anthocyanins (50–90 % of total anthocyanins) (De la Cruz et al. 2012; Liang et al. 2011; Mazza 1995; Zhao et al. 2010; Zhu et al. 2012). It is interesting to note that not all wild *Vitis* species contain high amounts of diglucoside anthocyanins. For example, diglucoside anthocyanins were not detectable in certain *V. cinerea* accessions (Anderson et al. 1970; De la Cruz et al. 2012). Many hybrid grapes between *V. vinifera* and North American wild *Vitis* species (e.g. *V. labrusca* and *V. riparia*) have been developed to improve *V. vinifera* cultivars for resistance to phylloxera and powdery mildew diseases. Depending on pedigrees and breeding histories, diglucoside anthocyanin contents in these hybrid grapes varied from 0 to 60 % of the total anthocyanins (Balík et al. 2013; De Rosso et al. 2012; Flamini and Tomasi 2000; Li et al. 2013; Robinson et al. 1966). For example, 'Chelois' (Seibel 10878) did not contain any detectable diglucoside anthocyanins while more than half of the anthocyanin contents in the hybrid 'Rosette' (Seibel 1000) were diglucosylated (Robinson et al. 1966). A recent study of ten hybrid grapes suggested that some of them did not contain detectable diglucoside anthocyanins (Balík et al. 2013). Because most *V. vinifera* do not produce diglucoside anthocyanins, absence of diglucoside anthocyanins could be used as a criterion for identifying purebred *V. vinifera* grapes (Picariello et al. 2012, 2014).

A causal correlation between the *5-O-glucosyltransferase* (*5GT*) gene and diglucoside anthocyanin production was established in a mapping population between 'Regent' (producing diglucoside anthocyanin) and 'Lemberger' (not producing diglucoside anthocyanin). One 'Regent' *5GT* allele encoded a truncated protein due to a premature stop codon and no diglucoside anthocyanins were produced in F1 plants carrying this *5GT* allele. F1 plants with the other 'Regent' *5GT* allele, which encoded a full-length protein, produced diglucoside anthocyanins (Hausmann et al. 2009, 2010). The functional *5GT* allele was named as *5GT-Cha* and the *5GT* allele with the premature stop codon was named as *5GT-Dia* after the parents of 'Regent', a cross between the interspecific hybrid 'Chambourcin' and *V. vinifera* cv. 'Diana' (Janvary et al. 2009). Two reverse mutations were required to restore *5GT-Dia* allele's 5-*O*-glucosyltransferase function, including elimination of the premature stop codon mutation and a Valine to Leucine transition at the amino acid sequence position 121(amino acid 110 in Fig. 1) located in the vicinity of a catalytic site (Janvary et al. 2009). The authors argued that the two mutations in *5GT-Dia* offered an explanation as to why revertants for this allele have never been reported in breeding programs.

To determine whether other forms of *5GT* alleles are present in *V. vinifera* and other *Vitis* species and to examine their relationships, we surveyed the coding sequence

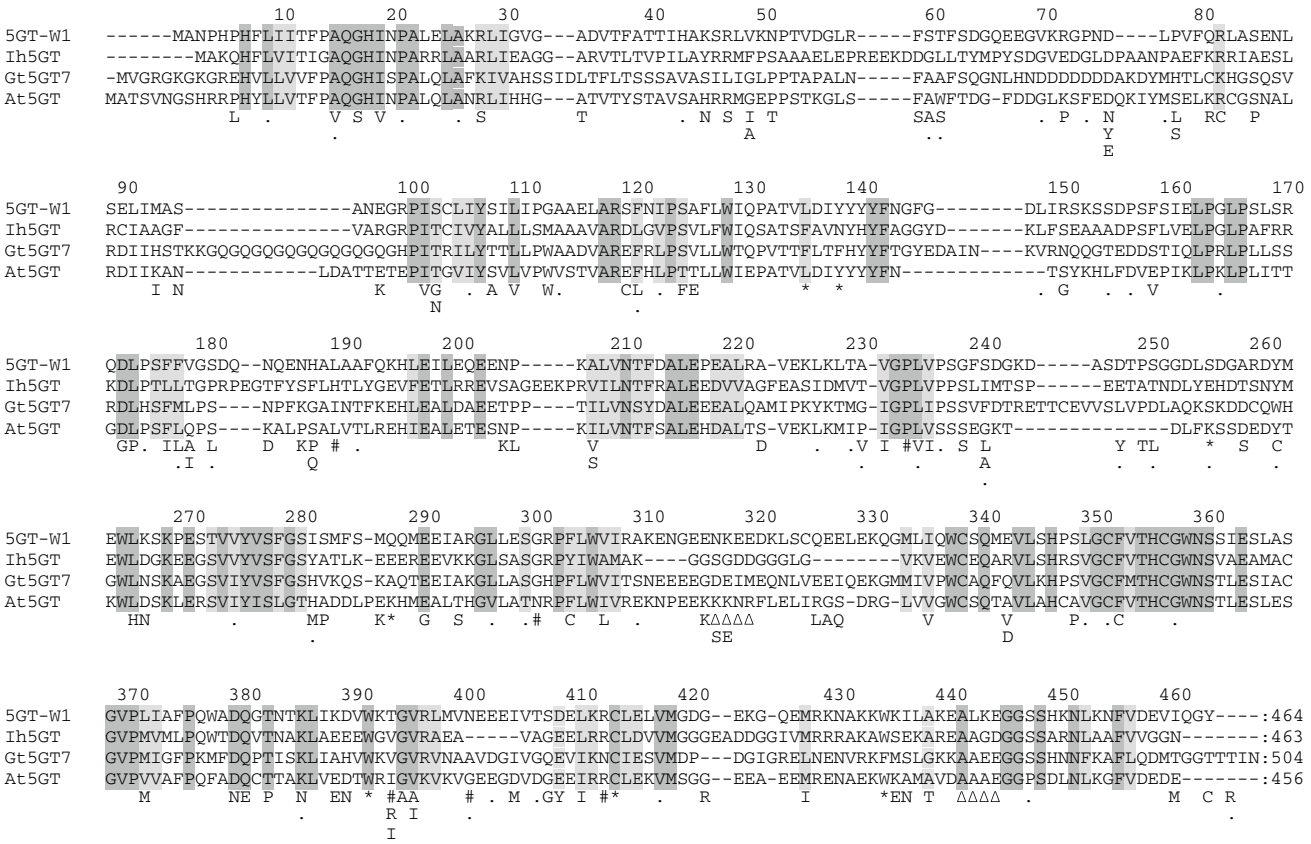

Fig. 1 Distribution and alignment of observed amino acid changes in *Vitis* 5GT with deduced amino acid sequences of *5GT* genes from three other plant species. 5GT-W1: 5GT W1 allele observed in this study, which was identical to the 5GT-Cha allele reported by Janvary et al. (2009); Ih5GT: *Iris x hollandica* 5GT (BAD06874); Gt5GT7: *Gentiana triflora* 5GT (BAG32255); At5GT: *Arabidopsis thaliana* 5GT (NP_193146, AT4G14090). *Dark and light shading* indicate identical and similar amino acids, respectively. *Dashes* represent alignment gaps among sequences. The *numbers* represent amino acid positions in 5GT-W1. Below the alignment: *letters* indicate amino acid substitutions, "*Asterisk*" indicates a stop codon, "*Triangle*" indicates an amino acid deletion, "*Hash*" indicates a frameshift, and "*Dot*" indicates a silent mutation

variation of the *5GT* gene in 70 *V. vinifera* ssp. *vinifera* cultivars, 52 *V. vinifera* ssp. *sylvestris* accessions, 23 hybrid grapes involving various *Vitis* species, and two to five accessions each of *V. aestivalis*, *V. riparia*, *V. rupestris*, *V. labrusca*, *V. amurensis*, *V. cinerea* and *V. rotundifolia*. Surprisingly, many different *5GT* allelic forms of loss-of-function mutations were present in *V. vinifera* ssp. *vinifera* and ssp. *sylvestris*. These loss-of-function alleles, however, were not found in other *Vitis* species. We further analyzed the anthocyanin profiles of 27 *V. vinifera* cultivars representing various *5GT* genotypes. While most *V. vinifera* cultivars do not have functional *5GT* and produce no diglucoside anthocyanins, some exceptions exist. This work advances our understanding of the genetic and functional diversity of *5GT* in the *Vitis* species, *V. vinifera* in particular, and provides insights into how *5GT* has evolved within *V. vinifera*. It also provides the molecular basis for tracking various allelic forms of this critical wine quality gene in a grape breeding program for grape variety development.

Materials and methods

Plant materials

The experimental material was obtained from either USDA-Agricultural Research Service (USDA-ARS) *Vitis* clonal repository in Davis, California (identified as DVIT) or USDA-ARS *Vitis* clonal repository in Geneva, New York (identified as GVIT) (Table 1). USDA *Vitis* repository in Davis, California conserves a collection of more than 1,200 *V. vinifera* ssp. *vinifera* cultivars. Many of these *V. vinifera* cultivars are genetically related (Myles et al. 2011). We selected 67 cultivars of *V. vinifera* from the Davis repository for this study on the basis of the work of Myles et al. (2011). Among them, cultivars with many first-degree relationships, such as 'Aswad' and 'Muscat of Alexandria' were selected. Representative wine grapes such as 'Chardonnay' and 'Pinot Noir' and common table grapes such as 'Thompson Seedless' and 'Husseine' were also selected.

Table 1 Identities, sources, *5GT* genotypes, and 3,5-di-*O*-glucoside anthocyanin content of the grape accessions included in the present study

Accession ID	Species	Cultivar	Origin	*5GT* genotype		Content of diglucoside anthocyanins	
				5GT-allele1	*5GT-allele2*	Diglucoside anthocyanins/total anthocyanins	Source of data
DVIT 1395	*Aestivalis*	Lincecumii	California, USA	W20	W21	>30 %	Liang et al. (2012)
DVIT 1585	*Aestivalis*	Aestivalis	Illinois, USA	W19	W22		
GVIT 1285	*Amurensis*		Asia	W5	W4	>50–91 %	De La Cruz et al. (2012), Liang et al. (2012), Mazza (1995), Zhao et al. (2010), Zhu et al. (2012)
GVIT 709	*Amurensis*		Asia	W3	W3		
GVIT 171	*Cinerea*	B 17	Illinois, USA	W23	W24	<15 %	Anderson et al. (1970), De La Cruz et al. (2012), Liang et al. (2012)
GVIT 259	*Cinerea*	Ill 65	Illinois, USA	W23	W27		
GVIT 268	*Cinerea*	Resseguier 2	Texas, USA	W23	W25		
GVIT 269	*Cinerea*	B27	Illinois, USA	W23	W26		
DVIT 1263	Hybrid	Aestivalis Yeager	France	A1	W23		
DVIT 3106[a]	Hybrid	Regent	Germany	C1	W1	25.40 %	Balík et al. (2013)
GVIT 23	hybrid	Colobel (Seibel 8357)	France	A1	W36	9.00 %	De Rosso et al. (2012)
GVIT 35	Hybrid	Seyve-Villard 23.512	France	A1	W34	55.50 %	De Rosso et al. (2012)
GVIT 41	Hybrid	Baco Noir (Baco 1)	France	E	W13	36.5 %, 100 %	De Rosso et al. (2012), Robinson et al. (1966)
GVIT 42	Hybrid	Chelois (Seibel 10878)	France	A1	C2	0	De Rosso et al. (2012), Robinson et al. (1966)
GVIT 135	Hybrid	Vergennes	Vermont, USA	A1	W10		
GVIT 254	Hybrid	Isabella	USA	A1	W9	13 %	Flamini and Tomasi (2000)
GVIT 323	Hybrid	Seibel 8745	France	A1	W35	25 %	De Rosso et al. (2012)
GVIT 336	Hybrid	Seibel 5163	France	A1	W34	33.60 %	Liang et al. (2013)
GVIT 433	Hybrid	Wheeler	USA	W11	W19		
GVIT 437	Hybrid	Seibel 6339	France	A1	W13	59.90 %	Liang et al. (2013)
GVIT454	Hybrid	Seibel 7052	France	W34	W35	58.70 %	Liang et al. (2013)
GVIT 464	Hybrid	Seibel 5409	France	A1	A2	No anthocyanin[b]	Liang et al. (2013)
GVIT 468	Hybrid	Rosette (Seibel 1000)	France	F	W1	63.70 %	Robinson et al. (1966)
GVIT 480	Hybrid	Seibel 9280	France	A1	W15	0.67 %	Liang et al. (2013)
GVIT 496	Hybrid	Seibel 7162	France	A2	W34	46.70 %	Liang et al. (2013)
GVIT 514	Hybrid	Seibel A	France	W1	W34	55.10 %	Liang et al. (2013)
GVIT 522	Hybrid	Clinton	USA	W1	W9	25.58–36.7 %	De Rosso et al. (2012), Flamini and Tomasi (2000)
GVIT 1251	Hybrid	Kyoho[c]	Japan	A1/A1	W11/W19	57.50 %	Li et al. (2013)
GVIT 1334	Hybrid	Lutie	Tennessee, USA	W11	W19		
GVIT 1344	Hybrid	Galibert 238-35	France	A1	W19	3 %	De Rosso et al. (2012)
GVIT 1374	Hybrid	Shimek	British Columbia, Canada	A1	W10		
GVIT 1541	Hybrid	Alexander	New York, USA	A1	W9		

Table 1 continued

Accession ID	Species	Cultivar	Origin	*5GT* genotype		Content of diglucoside anthocyanins	
				5GT-allele1	*5GT-allele2*	Diglucoside anthocyanins/total anthocyanins	Source of data
GVIT 229	*Labrusca*	Dunkel 1	New York, USA	W10	W10	~40 %	Liang et al. (2012)
GVIT 878	*Labrusca*	Rem NE 11	New Hampshire, USA	W11	W19		
GVIT 994	*Labrusca*	Rem NE 4	Massachusetts, USA	W11	W19		
GVIT 1009	*Riparia*	Rem NE 22	Massachusetts, USA	W11	W11	50–64 %	Liang et al. (2012)
GVIT 1324	*Riparia*	Zumbrunnen	Wisconsin, USA	W12	W23		
GVIT 773	*Riparia*	1F	Colorado, USA	W12	W12		
GVIT 894	*Riparia*	Rem 85-76	New Hampshire, USA	W1	W9		
DVIT 1689	*Rotundifolia*	Olmo (U62-56)	USA	W29	W30	~100 %	Ballinger et al. (1973), Huang et al. (2009), Sandhu and Gu (2010), You et al. (2012)
DVIT 1690	*Rotundifolia*	Olmo (U62-61)	USA	W29	W32		
DVIT 1692	*Rotundifolia*	Olmo (U66-39)	Florida, USA	W29	W33		
DVIT 1695	*Rotundifolia*	Olmo (U66-55)	Florida, USA	W29	W31		
DVIT 2186	*Rotundifolia*	Noble	Georgia, USA	W28	W29		
GVIT 180	*Rupestris*	B 38	Texas, USA	W8	W13	50–83 %	Liang et al. (2012), Mazza (1995)
GVIT 596	*Rupestris*		USA	W7	W6		
DVIT 2009	*Sylvestris*		France	A1	F	~0 %	Revilla et al. (2010)
DVIT 2012	*Sylvestris*	L57	France	A1	W17		
DVIT 2014	*Sylvestris*	L19	France	A1	A2		
DVIT 2426.5	*Sylvestris*		Tunisia	A2	C2		
DVIT 2426.11	*Sylvestris*		Tunisia	A1	W16		
DVIT 2426.27	*Sylvestris*		Tunisia	B2	W16		
DVIT 2426.41	*Sylvestris*		Tunisia	A1	B3		
DVIT 2426.77	*Sylvestris*		Tunisia	A1	W15		
DVIT 2440.1	*Sylvestris*	Ayedere	Turkmenistan	G	W14		
DVIT 2444.1	*Sylvestris*		Turkmenistan	A1	A1		
DVIT 2446.1	*Sylvestris*		Turkmenistan	A1	C4		
DVIT 2854	*Sylvestris*	Olmo	NA[d]	W2	W2		
DVIT 3348.26	*Sylvestris*		Georgia	A1	B1		
DVIT 3348.27	*Sylvestris*		Georgia	A1	B1		
DVIT 3349.04	*Sylvestris*		Georgia	A1	G		
DVIT 3349.16[e]	*Sylvestris*		Georgia	A1	W1		
DVIT 3350.02[e]	*Sylvestris*		Georgia	B1	W19		
DVIT 3350.22[e]	*Sylvestris*		Georgia	B1	W1		
DVIT 3351.01	*Sylvestris*		Armenia	A1	W14		
DVIT 3351.28	*Sylvestris*		Armenia	A1	A1		
DVIT 3353.03	*Sylvestris*		Armenia	A1	B1		
DVIT 3353.35	*Sylvestris*		Armenia	A1	A1		
DVIT 3354.01	*Sylvestris*		Armenia	A1	D		
DVIT 3355.01	*Sylvestris*		Georgia	A1	A1		
DVIT 3355.06	*Sylvestris*		Georgia	A1	B1		
DVIT 3356.03	*Sylvestris*		Armenia	A1	A1		
DVIT 3356.08	*Sylvestris*		Armenia	A1	B1		
DVIT 3357.14	*Sylvestris*		Georgia	B1	B1		
DVIT 3357.16	*Sylvestris*		Georgia	A1	B1		

Table 1 continued

Accession ID	Species	Cultivar	Origin	*5GT* genotype		Content of diglucoside anthocyanins	
				5GT-allele1	*5GT-allele2*	Diglucoside anthocyanins/total anthocyanins	Source of data
DVIT 3555.12	*Sylvestris*		Azerbaijan	A1	B1		
DVIT 3556.19	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3556.23	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3557.32	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3557.33	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3603.19	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3603.25	*Sylvestris*		Azerbaijan	A1	W14		
DVIT 3604.05	*Sylvestris*		Azerbaijan	D	W14		
DVIT 3604.08	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3605.14	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3605.19	*Sylvestris*		Azerbaijan	A1	G		
DVIT 3607.02	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3607.42	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3608.12	*Sylvestris*		Azerbaijan	A1	B1		
DVIT 3608.17	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3609.08	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3609.13	*Sylvestris*		Azerbaijan	A1	B1		
DVIT 3612.23	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3612.28	*Sylvestris*		Azerbaijan	A1	B1		
DVIT 3614.19	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3614.24	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 3615.28	*Sylvestris*		Azerbaijan	A1	D		
DVIT 3615.33	*Sylvestris*		Azerbaijan	A1	A1		
DVIT 353	*Vinifera*	Bhokari	India	B1	C5	No anthocyanin	
DVIT 371	*Vinifera*	Chaouch	Turkey	A1	W15	No anthocyanin	
DVIT 400	*Vinifera*	Eununi	Israel	A2	W16	Detectable[f]	This study
DVIT 402	*Vinifera*	Fahri	Afghanistan	C3	E	No anthocyanin	
DVIT 421	*Vinifera*	Husseine	Afghanistan	A1	W16	No anthocyanin	
DVIT 427	*Vinifera*	Kali Sahebi	India	B3	C3	Not detectable	This study
DVIT 439	*Vinifera*	Kisumashi	Afghanistan	B1	C3	No anthocyanin	
DVIT 465	*Vinifera*	Muscat of Alexandria	Egypt	A1	A1	No anthocyanin	
DVIT 480	*Vinifera*	Olivette Blanche	France/Tunisia	A1	B3	No anthocyanin	
DVIT 496	*Vinifera*	Queen of the Vineyard	Hungary	C3	C6	No anthocyanin	
DVIT 501	*Vinifera*	Rhazaki De Crete	Greece	A1	C3	No anthocyanin	
DVIT 535	*Vinifera*	Thompson Seedless	Turkey	B1	C3	No anthocyanin	
DVIT 541	*Vinifera*	Turkish Grape	Turkey	C1	D	No anthocyanin	
DVIT 548	*Vinifera*	Vorosmarthy	Hungary	A1	A2	No anthocyanin	
DVIT 558	*Vinifera*	Sidezites Proimo	Greece	B3	W15	Detectable	Liang et al. (2011); This study
DVIT 562	*Vinifera*	Itonychi Mavro	Greece	A1	B3	Not detectable	This study
DVIT 563	*Vinifera*	Ajimi	Iraq	A1	W15	Detectable	This study
DVIT 565	*Vinifera*	Askary	Iran	C3	C3	No anthocyanin	
DVIT 569	*Vinifera*	Dais-el-anz	Iraq	W16	W16	No anthocyanin	
DVIT 580	*Vinifera*	Keshmesh	Iran	B1	C3	No anthocyanin	

Table 1 continued

Accession ID	Species	Cultivar	Origin	*5GT* genotype		Content of diglucoside anthocyanins	
				5GT-allele1	*5GT-allele2*	Diglucoside anthocyanins/total anthocyanins	Source of data
DVIT 582	*Vinifera*	Kreatza	Croatia	A1	B1	No anthocyanin	
DVIT 597	*Vinifera*	Salomani	Iraq	B1	C8	Not detectable	This study
DVIT 633	*Vinifera*	Aramon	France	A1	E	Not detectable	This study
DVIT 671	*Vinifera*	Bonarda	Italy	A1	A2	Not detectable	This study
DVIT 675	*Vinifera*	Cabernet Franc	France	A1	A2	Not detectable	This study
DVIT 681	*Vinifera*	Carignane	Spain	A2	B1	Not detectable	This study
DVIT 706	*Vinifera*	Divromo	Greece	C9	W15	No anthocyanin	
DVIT 738	*Vinifera*	Gewürztraminer	Italy	A2	B1	Not detectable (light color)	
DVIT 774	*Vinifera*	Katta Kurgan	Russia Federation	A1	C3	No anthocyanin	
DVIT 826	*Vinifera*	Merlot	France	A1	A1	Not detectable	This study
DVIT 829	*Vinifera*	Mezes	Austria	A1	W15	No anthocyanin	
DVIT 841	*Vinifera*	Olmo 39098	Portugal	A1	A1	No anthocyanin	
DVIT 883	*Vinifera*	Pardala	Greece	A1	C9	Not detectable	This study
DVIT 928	*Vinifera*	Red Veltliner	Germany	A1	A1	Not detectable	This study
DVIT 954	*Vinifera*	Sémillon	France	A2	C2	No anthocyanin	
DVIT 981	*Vinifera*	Thiakon	Greece	E	E	No anthocyanin	
DVIT 990	*Vinifera*	Touriga	Portugal	A1	A1	Not detectable	This study
DVIT 992	*Vinifera*	Trincadeiro	Portugal	A1	B1	No anthocyanin	
DVIT 1021	*Vinifera*	Zalovitico	Greece	E	W18	Detectable	This study
DVIT 1036	*Vinifera*	Kontocladi	Greece	C6	W15	No anthocyanin	
DVIT 1066	*Vinifera*	Pinot Noir	France	A2	B1	Not detectable	This study
DVIT 1071	*Vinifera*	Prokupac	Serbia	C6	W18	Detectable	This study
DVIT 1119	*Vinifera*	Plavac Mali	Serbia/Montenegro	A1	A1	Not detectable	This study
DVIT 1123	*Vinifera*	Sauvignon Gris	France	A2	C2	No anthocyanin	
DVIT 1326	*Vinifera*	Black Kishmish	Russia Federation	C3	W16	Detectable	Liang et al. (2011); This study
DVIT 1332	*Vinifera*	Chardonnay	France	A1	A2	No anthocyanin	
DVIT 1342	*Vinifera*	Zinfandel	Croatia	A1	A2	Not detectable	This study
DVIT 1350	*Vinifera*	Cabernet Sauvignon	France	A1	A2	Not detectable	Li et al. (2013); This study
DVIT 1982	*Vinifera*	Dabouki	Israel	A1	A1	No anthocyanin	
DVIT 2054	*Vinifera*	Asma	Russia Federation	A1	C7	Not detectable	This study
DVIT 2057	*Vinifera*	Bastardo	Portugal	B1	C2	Not detectable	This study
DVIT 2059	*Vinifera*	Ezerjo	Hungary	E	W15	No anthocyanin	
DVIT 2080	*Vinifera*	Grassa De Cotnari	Romania	A1	E	No anthocyanin	
DVIT 2083	*Vinifera*	Red Roumi	Egypt	A1	B3	Not detectable	This study
DVIT 2084	*Vinifera*	Khalili	Afghanistan	C3	C3	No anthocyanin	
DVIT 2085	*Vinifera*	Aswad	Yemen	A1	E	NA	
DVIT 2086	*Vinifera*	Vranac	Serbia/Montenegro	A1	C6	Not detectable	This study
DVIT 2173	*Vinifera*	Oubeidy	Lebanon	A1	C9	No anthocyanin	
DVIT 2341	*Vinifera*	Trollinger	Germany	B1	C1	Not detectable	This study
DVIT 2635	*Vinifera*	Doradillo	Spain	A1	A1	No anthocyanin	
DVIT 2636	*Vinifera*	Fayoumi	Egypt	A1	E	No anthocyanin	
DVIT 2648	*Vinifera*	Perle de Csaba	Hungary	C6	W14	No anthocyanin	
DVIT 2654	*Vinifera*	Sultana Crimson	Turkey	B1	C6	Not detectable	Liang et al. (2011)
DVIT 2659	*Vinifera*	Tavriz	Serbia	A1	E	No anthocyanin	

Table 1 continued

Accession ID	Species	Cultivar	Origin	*5GT* genotype		Content of diglucoside anthocyanins	
				5GT-allele1	*5GT-allele2*	Diglucoside anthocyanins/total anthocyanins	Source of data
DVIT 2913	*Vinifera*	Flame Tokay	Algeria	C3	D	Not detectable	This study
DVIT 3002	*Vinifera*	Chasselas Rouge	France	A2	A2	Not detectable	This study
DVIT 3077	*Vinifera*	Sabal Kanskoi	Russia	C3	D	NA	
GVIT 1022	*Vinifera*	Madeleine Sylvaner	Germany/France?	A2	C1	No anthocyanin	
GVIT 1371	*Vinifera*	White Riesling	Germany	A2	E	No anthocyanin	
GVIT 1434	*Vinifera*	Ehrenfelser	Germany	A1	A2	No anthocyanin	

[a] 'Regent' was not included in this study but listed as a reference for its genotype and diglucoside anthocyanin content from literatures (Janvary et al. 2009; Balík et al. 2013)

[b] The grape berries are white/green or non-colored

[c] 'Kyoho' is a tetraploid grape containing two copies each of the alleles W and A1

[d] Information is not available

[e] These three *sylvestris* accessions were collected along a major road and a major river near a small city in Georgia, raising the possibility that they are interspecific hybrids; and their taxonomic identities are yet to be confirmed

[f] The diglucoside anthocyanin peaks were detected, but their relative amount were not quantified

Furthermore, cultivars without first-degree relatives such as 'Zalovitico' and 'Itonychi Mavro' were added to increase the chance of finding novel *5GT* alleles. In addition to *V. vinifera*, 52 accessions of *V. vinifera* ssp. *sylvestris*, the presumed wild progenitor of cultivated *vinifera* grapevines, and five accessions of *V. rotundifolia* and two accessions of *V. aestivalis* were sampled from the Davis repository. Two to four accessions of *V. riparia*, *V. rupestris*, *V. labrusca*, *V. cinerea* and *V. amurensis*, 23 accessions of various *Vitis* hybrids, and three accessions of *V. vinifera* cultivars were from the USDA-ARS *Vitis* clonal repository in Geneva, New York.

The *sylvestris* genomic DNA samples were extracted using a CTAB-based method (Doyle and Doyle 1987). All other genomic DNA samples were extracted using Qiagen DNeasy 96 Plant Kit (Qiagen, Valencia, CA, USA).

Genomic PCR, PCR cloning and sequencing of PCR products and plasmids

The grapevine *5GT* gene has no introns. A pair of primers P1 and P2 (P1, 5′ctgctaca**ATG**GCGAATCCT CAC3′, located in the junction of 5′ UTR and the coding sequence; P2, 5′gcaaaccgtataccgctaatgattcagtac3′, located in 3′ UTR) were used to amplify the entire 1,398-bp *5GT* coding sequence by genomic polymerase chain reaction (PCR) (see Fig. 1S for primer locations on the *5GT* coding sequence). The PCR reaction consisted of 0.2 µl genomic DNA (~10 ng), 10 µl 5xPhusion HF buffer, 1 µl dNTP (10 mM each), 2 µl P1 (10 µM), 2 µl P2 (10 µM) and 0.5 µl Phusion high-fidelity polymerase (New England Biolabs, Ipswich, MA, USA) in 50 µl reaction. The PCR conditions included 60 s at 98 °C for initial denaturation, followed by 35–40 cycles of 98 °C for 10 s, 60 °C for 30 s and 72 °C for 45 s, and 5 min at 72 °C for final extension. The PCR products were cleaned using Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics, Danvers, MA, USA). The purified PCR products were sequenced with Primer 3 (P3, 5′CCGAAAGCACTGGTAAACAC CTTTG3′, forward primer approximately 620 bps downstream from the ATG codon) and Primer 4 (P4, 5′CTTG GACTTGAGCCATTCCATGTAGTC3′, reverse primer approximately 800 bps from the ATG codon). When double peaks were observed over a long stretch of sequence in the chromatograms, the alleles in the samples were interpreted as being heterozygous for insertion/deletion (indel) polymorphisms. Additional sequencing was performed using P1, P2, P5 (5′TCTGATGGAAAAGATGCGTC3′, forward primer approximately 720 bps downstream from the ATG codon) or P6 (5′GGCCAAAATTTTCCATTTCTTGGC3′, reverse primer approximately 1,300 bps downstream from the ATG codon), depending on the locations of the encountered indels. PCR products were also cloned into a *pCR8/GW/TOPO* vector (Invitrogen, Carlsbad, CA) and the *5GT* inserts were sequenced using the primers located in the *pCR8/GW/TOPO* vector.

5GT sequence analysis

The *5GT* chromatographic sequence data were analyzed using Sequencher software (Gene Codes Corporation, Ann Arbor, MI, USA). The *5GT-Cha* was used as the reference

sequence (Janvary et al. 2009). *5GT* sequences acquired through cloning and plasmid sequencing were compared with the results of direct PCR sequencing to verify that two alleles from an individual sample contained all of the polymorphisms revealed by PCR sequencing. Most *5GT* alleles were recovered by PCR cloning and plasmid sequencing. Some were determined by direct PCR sequencing, when their PCR sequencing profiles showed no allelic variation.

Analysis of diglucoside anthocyanin content

Contents of diglucoside anthocyanins were previously assayed and reported for many *Vitis* species and hybrids (Balík et al. 2013; Ballinger et al. 1973; De la Cruz et al. 2012; De Rosso et al. 2012; Flamini and Tomasi 2000; Huang et al. 2009; Li et al. 2013; Liang et al. 2012, 2013; Mazza 1995; Robinson et al. 1966; Sandhu and Gu 2010; You et al. 2012; Zhao et al. 2010). The relevant results were summarized and included in Table 1. Although the reported results from various studies were not directly comparable due to the differences in detection methods, physiological conditions of the harvested grape berries, and the growing environments, they should provide reasonable references for the relative amount of diglucoside anthocyanins produced and, particularly, for qualitative assessment of presence or absence of diglucoside anthocyanins in the *Vitis* cultivars and species reported.

In contrast, very limited analysis was carried out for *V. vinifera* due to the widely accepted assumption that cultivated *V. vinifera* do not produce diglucoside anthocyanins. In this study, we determined the HPLC profiles of diglucoside anthocyanins in the berries of 27 *V. vinifera* cultivars covering various combinations of the *5GT* alleles (Table 1). Briefly, about 20 mature berries from each cultivar were sampled from vineyard and kept on wet ice before further processed. The berries were frozen and berry skins were peeled on dry ice, and then ground into powder in liquid nitrogen. Sample extraction and subsequent HPLC analysis followed the protocols as described by Liang et al. (2011). Because we were mainly interested in a qualitative assessment of presence or absence of diglucoside anthocyanins, only one bulk sample, with no biological duplicate, was analyzed for each cultivar. The identities of various diglucoside anthocyanin peaks in the *V. vinifera* samples were determined by using the HPLC profiles of berries from *V. amurensis* and 'Concord' grapes as references. Both *V. amurensis* and 'Concord' grapes contain many different diglucoside anthocyanins (Wu and Prior 2005; Zhao et al. 2010).

Haplotype networks

Haplotype networks were generated from aligned nucleotide sequences using TCS software (Templeton et al. 1992). These were used to illustrate mutational steps between alleles including multifurcations and/or reticulations. The software uses parsimony to minimize the number of convergent or parallel mutations while allowing for recombination. Two analyses were performed: (1) all sampled *V. vinifera* ssp. *vinifera* and *V. vinifera* ssp. *sylvestris* (26 alleles representing 244 sampled genes) and (2) all non-*Vinifera* wild species (27 alleles representing 44 sampled genes). To examine relationships among alleles predicted to be functional versus non-functional, including non-conservative amino acid changes, stop codons and frameshift mutations, were drawn on the haplotype networks based on amino acid translation of alignments.

Results

The grape *5GT* gene encodes a peptide of 464 amino acid residues with tandem double stop codons. Its coding region has 1,398 bps with no introns. This was consistent with the observation in a previous study in which two grape *5GT* alleles were identified (Janvary et al. 2009). Similarly, no intron was found in the *Arabidopsis 5GT* gene (At4g14090) (http://www.ncbi.nlm.nih.gov/gene?cmd=retrieveandlist_uids=827046). *5GT* genes were also identified from several other plant species including *Perilla frutescens* and *Petunia hybrida*. However, the intron status of the *5GT* genes from these species was unknown because they were isolated as cDNA clones (Yamazaki et al. 1999, 2002).

A total of 54 grape *5GT* alleles were discovered in this study (Tables 1, 2, Table 1S). To determine the relative functional importance and distribution patterns of allelic variations (mutations), we aligned the deduced amino acid sequences of *5GT* genes from *Arabidopsis thaliana* (*At5GT*), *Gentiana triflora* (*Gt5GT7*), *Iris x hollandica* (*Ih5GT*), and *Vitis riparia* (*5GT-W1* of this study) (Fig. 1). The grape *5GT-W1* of this study was identical to the *5GT-Cha* allele previously reported (Janvary et al. 2009). Among 172 mutations found in this study, 47 were silent (Fig. 1; Table 1S). Although the mutations were scattered across the entire *5GT* coding sequence (Fig. 1), there were apparently several mutation hotspots. For example, multiple mutations were detected in the 15-bp region from nucleotide positions 946 to 960 (amino acid positions 316–320) (Fig. 1; Table 1S). Two types of in-frame deletions were found in this region, including a 12-bp deletion specific to *V. cinerea* (corresponding amino acid changes noted as NKEE317-320Δ: deletion (Δ) of 4 amino acids NKEE at amino acid positions 317–310) and a 3-bp deletion (E320Δ) specific to *V. rotundifolia* and one *V. vinifera* allele (Table 2). Furthermore, three missense mutations were found in this region, including mutations at nucleotide sequence positions 946 specific to *V. rotundifolia* (E316K:

amino acid change from E to K at the amino acid position 316), 950 specific to a *V. vinifera* allele (N317S) and 952 specific to *V. aestivalis* (K318E). Apparently, this region is variable for both amino acid composition and number of amino acid residues (Fig. 1). It is also interesting to note that at certain amino acid positions as many as three forms of mutational changes were found. For example, D76E, D76 N, and D76Y were respectively found in *V. amurensis*, *V. vinifera* and *V. rotundifolia* (Fig. 1; Table 1S).

5GT alleles in *V. vinifera* ssp. *vinifera* and *V. vinifera* ssp. *sylvestris*

We identified 26 *5GT* alleles from 70 *vinifera* cultivars and 52 *sylvestris* accessions (Tables 1, 2). Eighteen of them contained premature stop codons and/or frameshift mutations (Table 2, Table 1S) and they were likely nonfunctional. Indeed, 3,5-*O*-diglucoside anthocyanins were not detected in the cultivars containing homozygous or combinations of these non-functional alleles (Table 1). These loss-of-function *5GT* alleles were classified into seven types (A, B, C, D, E, F and G) on the basis of their key mutational features (i.e., premature stop codons and/or frameshift mutations) (Table 2). A1, the most common *5GT* allele (~44 % or 107 out of 244 alleles) in both *vinifera* and *sylvestris*, contained one single base pair deletion at nucleotide position 901 resulting in a frameshift (G301#: a frameshift (#) replacing the amino acid G at the amino acid position 301) (Table 2, Table 1S). Allele A2 contained one additional 2-bp deletion at nucleotide positions 1,182–1,183 (T394#) in addition to the 1-bp deletion at position 901. However, A2 was much less frequent than A1 (about 8 %) (Tables 1, 2). The key feature for the B type of mutation was the deletion of one base pair at nucleotide position 700 (P234#). There were three alleles classified as B type (B1 to B3). One of them, B3, also contained a premature stop codon at position 407 (L136*: a stop codon (*) replacing L at the amino acid position 136). Nine alleles were assigned to C type (C1 to C9). Their key feature was the presence of a premature stop codon at nucleotide position 1,242 (C414*). The previously discovered nonfunctional *5GT-Dia* allele (Hausmann et al. 2009, 2010; Janvary et al. 2009) belongs to this type and was referred as C1 in this study (Tables 1, 2). C1 was not a common *5GT* allele and observed only in three *vinifera* samples in this study (~1 %) (Table 1). C1 was not found in *sylvestris* accessions surveyed in this study. While the C-type *5GT* alleles were present in both *vinifera* and *sylvestris*, it seems that additional mutations have accumulated in *vinifera* cultivars. Among the nine C-type alleles, eight were found in *vinifera*, but only two in *sylvestris* (Table 2). For the D, E, F or G type, only one *5GT* allele was identified in each. The D type contained a premature stop codon at nucleotide position 417 (Y139*) and a 7-bp deletion at nucleotide positions 1,203–1,209 (N401#). The E type contained a 1-bp deletion at nucleotide position 568 (L190#) and a premature stop codon at nucleotide position 1,175 (W392*). The F type contained a premature stop codon at nucleotide position 1,305 (W435*) and the G type contained a 68-bp insertion (sequence duplication) at nucleotide position 1,237 (R413#). The E type was identified only in *vinifera* while F and G types were recovered only in *sylvestris* (Tables 1, 2). Some of these *5GT* alleles had additional mutations in conserved amino acid residues (Table 2, Table 1S).

Eight *5GT* alleles with no frameshift or premature stop mutations (referred as W-type) (W1, W2, W14, W15, W16, W17, W18 and W19) were identified in *vinifera* and *sylvestris* in this study (Tables 1, 2). These alleles, with the exceptions of W2 and W17 for which berry samples were not available for analysis, were likely functional since cultivars containing one of these alleles produced detectable 3,5-*O*-diglucoside anthocyanins (Table 1). Seven of the eight W-type *5GT* alleles, with the exception of W18, were found in *sylvestris*. In contrast, only four of them (W14, W15, W16 and W18) were found in *vinifera*. Among the 70 *vinifera* cultivars assayed, 14 of them contained at least one W-type *5GT* allele (Table 1). Both W14 and W15 contained a 12-bp in-frame deletion at the nucleotide positions 1,325-1,336 (ALKE442-445Δ) and a point mutation resulting in an amino acid substitution at the conserved amino acid residue 353 (F353C). Interestingly, W15 was confirmed to produce trace, but detectable amount of 3,5-*O*-diglucoside anthocyanins (e.g. DVIT 558 and DVIT 563, Table 1), suggesting that the 12-bp in-frame deletion did not knock out the function of the allele. W14 was found in the *vinifera* accession DVIT 2,648 (Table 1). This *vinifera* cultivar produces green/white berries, therefore it was not possible to determine whether or not W14 produces 3,5-*O*-diglucoside anthocyanins. Since W14 and W15 were very similar and W15 was functional, it is very likely that W14 was also functional. W16 and W17 contained mutations in three conserved amino acid residues, two of which were also present in the A, D and F types (Table 2). W16 was confirmed to possess some 5GT function (DVIT 1326, Table 1). W17 was found in a *sylvestris* accession (DVIT 2012) for which no berry samples were available for HPLC assay, therefore the functionality of W17 was not determined in this study. However, W17 was likely functional as it had similar mutations as W16. W18 was found only in *vinifera* (DVIT 1021 and DVIT 1071) and had missense mutations which were not located in highly conserved residues compared to W1 (Table 2, Table 1S). Some other point mutations in W18 were also present in the B, C and G types (Table 2). W18 was functional in producing detectable 3,5-*O*-diglucoside anthocyanin (DVIT 1021 and DVIT 1071, Table 1). Both W1 and W19 were found in *sylvestris* (DVIT 3350.22 and

Table 2 Key mutation sites and distribution of 5GT alleles in *Vitis* species

Allele[a]	Mutation feature and sites[b]				Distribution among species										
	Frame shift	Stop codon	In-frame deletion	Non-conservative amino acid changes[c]	*Vinifera*	*Sylvestris*	Hybrid	*Riparia*	*Labrusca*	*Aestivalis*	*Cinerea*	*Rotundifolia*	*Amurensis*	*Rupestris*	Total
A1	G301#			**K**386 N, **V**396A, P78L	43	64	16								123
A2	G301#,T394#			**K**386 N, **V**396A, P78L	17	2	2								21
B1	P234#			**E**291G, **V**396I	13	14									27
B2	P234#			**E**291G, **V**396I		1									1
B3	P234#	L136*		**E**291G, **V**396I	5	1									6
C1		C414*		**L**110 V, **V**307L	3										3
C2		C414*		**V**307L	3	1	1								5
C3		C414*		**L**266H, **V**307L	15										15
C4		C414*		**V**307L		1									1
C5		C414*		**V**307L, E343 V	1										1
C6		C414*		**V**307L, E343 V	6										6
C7		C414*,S256*,Q288*		E343 V	1										1
C8		C414*		**V**307L	1										1
C9		C414*		**V**307L	3										3
D	N401#	Y139*		**L**174P, **K**386 N, **V**396A	3	3									6
E	L190#	W392*	E320Δ		11		1								12
F		W435*		**K**386 N, **V**396A		1	1								2
G	R413#			**F**304C, S349P		3									3
W1						2	3	1							6
W2						2									2
W3													2		2
W4													1		1
W5													1		1
W6														1	1
W7														1	1
W8														1	1
W9							3	1							4
W10				**D**409Y			2		2						4
W11				**D**409Y			3	2	2						7
W12				**D**173G, A439T				3							3
W13				**D**173G, A439T			2							1	3
W14			ALKE442-445Δ	**F**353C	1	4									5
W15			ALKE442-445Δ	**F**353C, **Q**381E	7	1	1								9

Table 2 continued

Allele[a]	Mutation feature and sites[b]				Distribution among species										
	Frame shift	Stop codon	In-frame deletion	Non-conservative amino acid changes[c]	*Vinifera*	*Sylvestris*	Hybrid	*Riparia*	*Labrusca*	*Aestivalis*	*Cinerea*	*Rotundifolia*	*Amurensis*	*Rupestris*	Total
W16				**D**380 N, **K**386 N, **V**396A, **A**208S	5	2									7
W17				**D**380 N, **K**386 N, **V**396A, **A**208S, **S**103 N		1									1
W18				A126E, **V**307L	2										2
W19				**G**395A		1	4		2	1					8
W20				**G**395A						1					1
W21				**G**395A						1					1
W22				**G**395A						1					1
W23			NKEE317-320Δ	**S**103G, S283P, **T**383P			1	1			4				6
W24			NKEE317-320Δ	**S**103G, S283P, **T**383P							1				1
W25			NKEE317-320Δ	**S**103G, S283P, **T**383P, **I**19 V							1				1
W26			NKEE317-320Δ	**S**103G, S283P, **T**383P, **I**19 V							1				1
W27			NKEE317-320Δ	**S**103G, S283P, **T**383P, **I**19 V, **A**15 V							1				1
W28			E320Δ	P78S, **F**121L, K267 N, **I**102 V, **R**82C								1			1
W29			E320Δ	P78S, **F**121L, K267 N								5			5
W30			E320Δ	P78S, **F**121L, K267 N, **S**125F								1			1
W31			E320Δ	P78S, **F**121L, K267 N, **I**102 V								1			1
W32			E320Δ	P78S, **F**121L, K267 N, **S**125F								1			1
W33			E320Δ	P78S, **F**121L, K267 N, **I**102 V								1			1
W34							5								5
W35							2								2
W36				**G**395A			1								1
Total alleles					18	17	16	5	3	4	5	6	3	4	

[a] The underlined alleles were *vinifera*/*sylvestris*-specific and the W alleles which were found in *vinifera*/*sylvestris* were in bold

[b] W1 amino acid sequence was used as the reference. A change in the amino acid sequence is indicated by following the format K386 N in which amino acid K was replaced by N at the amino acid position 386. "#" stands for frameshift, "*" for stop codon, and "Δ" for in-frame deletion

[c] Bold and underlined bold letters stand for similar or identical amino acid residues, respectively, in the 5GT alignment in Fig. 1

DVIT 3350.02, respectively), but not in *vinifera*. W1 was also found in *V. riparia* (GVIT 894) and hybrid grapes involving *V. riparia* (e.g. GVIT 468). Similarly, W19 was present in *V. labrusca*, *V. aestivalis* and hybrid grape GVIT 1344 (Table 1). It is possible that these *sylvestris* accessions might be hybrids between *V. vinifera* and wild *Vitis* species.

5GT alleles in non-*V. vinifera Vitis* species

A total of 27 *5GT* alleles were identified from 22 accessions of *V. aestivalis*, *V. amurensis*, *V. cinerea*, *V. labrusca*, *V. riparia*, *V. rotundifolia* and *V. rupestris* (Tables 1, 2). None of the *5GT* alleles from these *Vitis* species contained premature stop codon or frameshift mutations (Table 2, Table 1S). These alleles are likely functional and many of them indeed produce abundant 3,5-di-*O*-glucoside anthocyanins (Table 1). W1, identical to the *5GT-Cha* allele previously reported (Janvary et al. 2009), was recovered from a *V. riparia* accession (Tables 1, 2). This is consistent with the fact that the *5GT-Cha* allele was cloned from the grape hybrid 'Regent' which contains *V. riparia* in its pedigree (Eibach and Töpher 2003).

As expected, *5GT* alleles from different *Vitis* species had their own unique features. All *5GT* alleles from *V. rotundifolia* (W28 to W33) contained an in-frame 3-bp deletion at nucleotide positions 958–960 (E320Δ). These alleles collectively contained 27 point mutations unique to *V. rotundifolia* (Table 1S). All the anthocyanins in *V. rotundifolia* berries were diglucosylated (Ballinger et al. 1973; Huang et al. 2009; Sandhu and Gu 2010; You et al. 2012). Apparently, the in-frame 3-bp deletion in the *V. rotundifolia 5GT* does not compromise the 5GT functionality in producing diglucoside anthocyanins. Three *5GT* alleles (W3, W4 and W5) were identified from two *V. amurensis* accessions containing no missense mutations in conserved residues (Table 2, Table 1S). It is no surprise that the relative amount of diglucoside anthocyanins (>50–91 % of total anthocyanins) was very high in *V. amurensis* (Table 1) (De la Cruz et al. 2012; Liang et al. 2012; Mazza 1995; Zhao et al. 2010; Zhu et al. 2012). Four *5GT* alleles, W6, W7, W8 and W13, were identified from two *V. rupestris* accessions. W6, W7, and W8 contained no missense mutations in conserved residues. W13 was very similar to W12, a common allele from *V. riparia*, and they contained mutations in conserved or semi-conserved residues (Table 2, Table 1S). In addition to W12, *V. riparia* had four other alleles including W1, W9, W11, and W23. W1 and W9 had no missense mutations in conserved residues, while the other two alleles did. The presence of some *5GT* alleles with missense mutations in conserved residues in *V. riparia* and *V. rupestris* may compromise the function of these alleles, which in turn explains why only about 50 % of the anthocyanins were diglucosylated in these species (Table 1) (Liang et al. 2012; Mazza 1995). Four *5GT* alleles (W19–W22) were found in *V. aestivalis* and they contained seven unique point mutations (Table 1S). Interestingly, W19 was also present in two of the three *V. labrusca* accessions (Tables 1, 2). In addition to W19, two other alleles, W10 and W11, were identified from *V. labrusca*. W10 and W11 shared a common missense mutation at nucleotide position 1,225 (D409Y) (Table 2, Table 1S). W11 was also found in one of the *V. riparia* accessions (GVIT 1009) (W11/W11) (Tables 1, 2). The diglucoside anthocyanin contents in *V. aestivalis* and *V. labrusca* were relatively low (30-40 % or more) compared with the other wild *Vitis* species (Liang et al. 2012). All six *5GT* alleles from these two species (W10, W11, W19 to W22) contained missense mutations on conserved (nucleotide position 1,184, G395A) or semi-conserved (nucleotide position 1,225, D409Y) residues (Table 2). *V. cinerea* is an interesting exception among the wild *Vitis* species regarding the production of diglucoside anthocyanins. It produces much less diglucoside anthocyanins (<15 %) compared with other wild *Vitis* species (Table 2) (Anderson et al. 1970; De la Cruz et al. 2012; Liang et al. 2012). In fact, diglucoside anthocyanins were not detectable at all in one *V. cinerea* (Englem.) Millardet accession (De la Cruz et al. 2012). All four *V. cinerea* accessions recovered in this study (W23–W27) carried nine unique polymorphisms (Table 1S). Besides the three missense mutations at conserved or semi-conserved residues (nucleotide position 307, S103G; nucleotide position 847, S283P; and nucleotide position 1,147, T383P), they all contained an in-frame 12-bp deletion at nucleotide positions 949–960 (NKEE317-320Δ) (Fig. 1). Apparently some of these mutations significantly compromised the production of diglucoside anthocyanins. It is interesting to note that W23 was also found in one of the *V. riparia* accessions (GVIT 1324), raising the question as to whether this *V. riparia* accession was actually a hybrid between *V. riparia* and *V. cinerea*.

5GT alleles in hybrid grapes

Twenty-three hybrid grapes were surveyed for their *5GT* genotypes, including eleven Seibel hybrid grapes and some other widely cultivated cultivars such as 'Isabella', 'Clinton', and 'Kyoho' (Table 1). Some of these hybrid grapes have very complex genetic backgrounds. For example, many Seibel hybrid grapes have *V. aestivalis*, *V. cinerea*, *V. labrusca*, *V. riparia*, *V. rupestris* and *V. vinifera* in their pedigrees (http://www.ars-grin.gov/cgi-bin/npgs/acc/display.pl?1001422).

Fifteen out of the 23 hybrid grapes each contained one *vinifera*-specific *5GT* allele (A1, A2, C2, E, F, or W15) and one non-*vinifera 5GT* allele (Table 1). Most of these hybrids were reported to produce diglucoside anthocyanins (Table 1). Eleven different W-type *5GT* alleles were

recovered, including W1, W9 to W11, W13, W15, W19, W23, and W34–W36. W34, W35 and W36 were unique to the Seibel hybrid grapes, but their parental sources could not be identified from the *Vitis* species investigated in this study. Interestingly, three of the hybrid grapes contained *vinifera*-specific *5GT* alleles only. Specifically, 'Seibel 10878' contained A1 and C2, 'Seibel 5409' (with green/white berries) contained A1 and A2, and 'Seibel 9280' contained A1 and W15. 'Seibel 10878' did not produce diglucoside anthocyanins while 'Seibel 9280' did, although in very small quantity (0.67 % of total anthocyanins), further confirming that the *vinifera*-specific W15 weakly conferred *5GT* function (Table 1) (De Rosso et al. 2012; Liang et al. 2013; Robinson et al. 1966). Five other hybrid grapes, 'Seibel 7052' (W34/W35), 'Seibel A' (W1/W34), 'Clinton' (W1/W9), 'Lutie' (W11/W19) and 'Wheeler' (W11/W19) contained only non-*vinifera 5GT* alleles. It was not unexpected that 'Clinton' contained two W-type *5GT*s since it likely contains *V. riparia* and *V. labrusca* in its pedigree (http://www.ars-grin.gov/cgi-bin/npgs/acc/display.pl?1006882). Because Seibel grapes were often created from multiple crosses between various hybrid grapes, some progenies derived from such crosses could contain only *vinifera* or non-*vinifera 5GT* alleles. As expected, those hybrids carrying one or more W-type alleles produced diglucoside anthocyanins (Table 1).

It is worth noting that the tetraploid hybrid grape 'Kyoho' contained two W-type *5GT* alleles (W11 and W19) and two copies of A1 alleles based on the sequencing chromatograms of its genomic PCR products (data not shown) (Table 1). Both W11 and W19 are common alleles of *V. labrusca* (Table 1). 'Kyoho' is a tetraploid hybrid from the cross between cv. 'Ishihara-wase' and cv. 'Centennial'. These two parents were likely tetraploid mutants from *V. labruscana* cv. 'Campbell Early' and *V. vinifera* cv. 'Rozaki', respectively (Okamoto 2007). The *5GT* genotype of 'Kyoho' is consistent with the expectation that 'Kyoho' inherited both *5GT* alleles from each parent (Table 1).

5GT haplotype network

To further elucidate the possible evolutionary relationships among various *5GT* alleles, two haplotype networks were constructed, one for the *V. vinifera* complex involving both *sylvestris* and *vinifera* (Fig. 2), and the other for non-*vinifera Vitis* species (Fig. 3). The *5GT* haplotype network of *sylvestris* and *vinifera* showed that the most frequent allele A1 was shared by both subspecies (Fig. 2). There were many branches leading to less frequent alleles that were many mutational steps removed from A1. A striking pattern was that the majority of branches across the entire network involved stop codons and/or frameshift mutations, which likely resulted in generation of non-functional alleles. Only the four branches that led to alleles W14 through W19 did not involve stop codon or frameshift mutations, but each of these four branches included amino acid substitutions in conserved residues. No recombinational events were evident.

The haplotype network of wild *Vitis* species alleles showed that non-conservative amino acid changes and in-frame deletions were distributed sparsely throughout most of the network (Fig. 3). *V. rotundifolia* alleles were distinct from other *Vitis* species alleles. Several wild species alleles were shared (W1, W11, W19, W23) or closely related (W12 and W13) between species. There was a single reticulation that involved alleles W8 in *V. rupestris* and W9 in *V. riparia*.

Discussion

V. vinifera is the most widely cultivated *Vitis* species and produces grapes with superior quality of fruits from which most high quality wines are made. Because *V. vinifera* is poorly adapted or tolerant to pests, temperature extremes, drought, and other forms of biotic and abiotic stresses, grape breeders address these issues by introgressing adaptive traits from wild *Vitis* species. However, these wild adaptive traits are often accompanied by poor fruit quality and low yield. For example, most of the wines produced from hybrid grapes, such as those involving the North American species *V. labrusca*, have a strong "foxy" flavor, which is viewed as inferior quality by many consumers. Undesirable biochemical properties such as "foxiness" in fruits and wines can be evaluated through taste, but more reliably they can be determined through biochemical analyses. However, molecular markers are favored these days as a method of choice for tracking the presence or absence of these compounds in a breeding line when the genes or genomic regions responsible for the synthesis of these compounds are known. The association of methyl anthranilate with "foxy" flavor in grape fruits and wines and the discovery of the anthraniloyl-CoA:methanol anthraniloyl transferase gene responsible for the synthesis of methyl anthranilate (Wang and De Luca 2005) have made it possible to develop and use molecular markers for tracking the "foxy" gene in a grape breeding program. Such a marker-assisted selection tool should also be feasible for tagging different *5GT* alleles reported in this study, since the biochemical process of the *5GT* gene regulating the synthesis of diglucoside anthocyanins has been recently elucidated (Janvary et al. 2009).

One critical requirement for developing effective molecular markers to track a gene of interest in a breeding program is the knowledge of the allelic forms of the gene and their associated functionalities or phenotypes. Two *5GT* alleles, one functional (*5GT-Cha*) and the other

nonfunctional (*5GT-Dia*), were previously identified (Hausmann et al. 2010, 2009; Janvary et al. 2009). However, how representative these two *5GT* alleles are in the *Vitis* species, *V. vinifera* in particular, is unknown. The present study identified 18 alleles with apparent loss-of-function mutations (premature stop codons and frameshift indels) in *V. vinifera*, including the previously reported nonfunctional *5GT-Dia* allele (Janvary et al. 2009). In addition, 36 W-type *Vitis 5GT* alleles, including the previously discovered *5GT-Cha* allele, were identified in this study. While additional functional and non-functional *5GT* alleles could be identified from other *Vitis* materials which were not included in this study, we believe that most *Vitis 5GT* alleles, particularly those in *V. vinifera*, have likely been captured in this study. Discovery of these allelic forms and their associated sequence information in this study will provide the foundation for designing effective molecular markers to track these alleles in grape breeding programs.

A very high level of polymorphisms was observed in *Vitis 5GT*. Fifty-four *5GT* alleles collectively carrying 172 mutational variations were found in this study, on average with one mutation every 8 bps in the *5GT* coding region. At some amino acid sites more than one mutant form was present. Within *V. vinifera*, a total of 82 mutations were found with an average of one mutation every 16 bps. Similar levels of polymorphisms (one mutation every 21 bps) were found within each of the two *V. vinifera* subspecies: *vinifera* (68 mutated sites) and s*ylvestris* (67 mutated sites). These levels of polymorphism were much higher than the average nucleotide polymorphism reported in grape genome sequences (one SNP per 60 to 100bps) (Lijavetzky et al. 2007; Riahi et al. 2013; Velasco et al. 2007). The fact that the *5GT* gene has no introns, but yet maintains such a high level of polymorphisms is very interesting. What is particularly intriguing was that many forms of mutations, including frameshift indels, premature stop codons, and point mutations, were involved in generating the diverse *5GT* alleles. This highly polymorphic locus offers an excellent diagnostic tool for tracking the origins of *5GT* alleles in a grape breeding program or in wine-making quality control processes. Indeed, in most cases, the *5GT* allelic profiles of hybrid grapes clearly identified them as hybrid progeny between *V. vinifera* and a non-*vinifera Vitis* species, with one non-functional *5GT* allele from *V. vinifera* and one functional W-type *5GT* allele from a non-*vinifera* parent (Table 1). The utility of *5GT* alleles as diagnostic markers was also clearly demonstrated for interpreting the genetic origin of the tetraploid grape hybrid 'Kyoho' (see "Results").

The wide presence of *5GT* alleles with apparent loss-of-function mutations in *V. vinifera* ssp. *vinifera* explains why most *vinifera* cultivars do not produce diglucoside anthocyanins. Among the 70 *vinifera* cultivars investigated, 56 carried only those apparent loss-of-function alleles and, as expected, they had no detectable diglucoside anthocyanins. It was important to note that 14 of the 70 *vinifera* cultivars investigated in this study each carried one W-type *5GT* allele. However, eight of these 14 cultivars produced green/white berries and therefore their ability to produce diglucoside anthocyanins could not be determined. The remaining six *vinifera* cultivars with red-colored berries contained W15, W16 or W18 alleles and produced some diglucoside anthocyanins. Knowing which *vinifera* cultivars carry W-type *5GT* alleles in the *vinifera* germplasm pool is important, particularly when these accessions are involved in breeding. What is intriguing was the observation that some white grapes carry W-type *5GT* alleles, although no diglucoside anthocyanins were produced due to lack of anthocyanins. It is expected that when such white grape germplasm are crossed with those red grapes which do not have functional *5GT* alleles, progeny from the crosses could potentially produce unwanted diglucoside anthocyanins.

Our haplotype network analyses provided an opportunity to understand the evolutionary relationships among various *5GT* alleles. It should be noted that the haplotype networks constructed in this study were used as phenetic representations, i.e., as illustrations of relatedness as estimated by mutations, rather than as phylogenetic reconstructions. The mode and tempo of evolution of this gene in *Vitis* is complicated by reticulation and convergent evolution. Genus-wide gene flow also circulates these mutations among different, at least, sympatric species adding to extensive homoplasy within and among lineages. For example, although allele G in *sylvestris* branched out from allele C9 in *vinifera* (Fig. 2), this did not imply that G directly evolved from C9. In fact, *V. vinifera* ssp. *sylvestris* was the progenitor of cultivated grape *V. vinifera* ssp. *vinifera* and it was not likely that allele G was derived from ssp. *vinifera*. At the same time, a common ancestor to the two subspecies is not known in the Caucasus where ssp. *vinifera* was historically domesticated. This pattern could represent gene flow between the two subspecies, incomplete lineage sorting or a sampling artifact. Similarly, because allele W12 was one mutational step from allele W13, this did not imply that the *V. riparia* allele evolved directly from the *V. rupestris* allele (Fig. 3). There is a possibility of reticulate evolution within the Series *Ripariae* to which these two species belong.

Homoplasious mutations were evident in both networks (not shown), an indication that these networks are single depictions of many different possible representations of distances among alleles. In *V. vinifera* the predominant pattern showed extensive loss-of-function mutations that were accumulated throughout the majority of branches. This pattern strongly contrasted with the non-*Vinifera* species network, in which none of the branches contained loss-of-function mutations and in which non-conservative amino

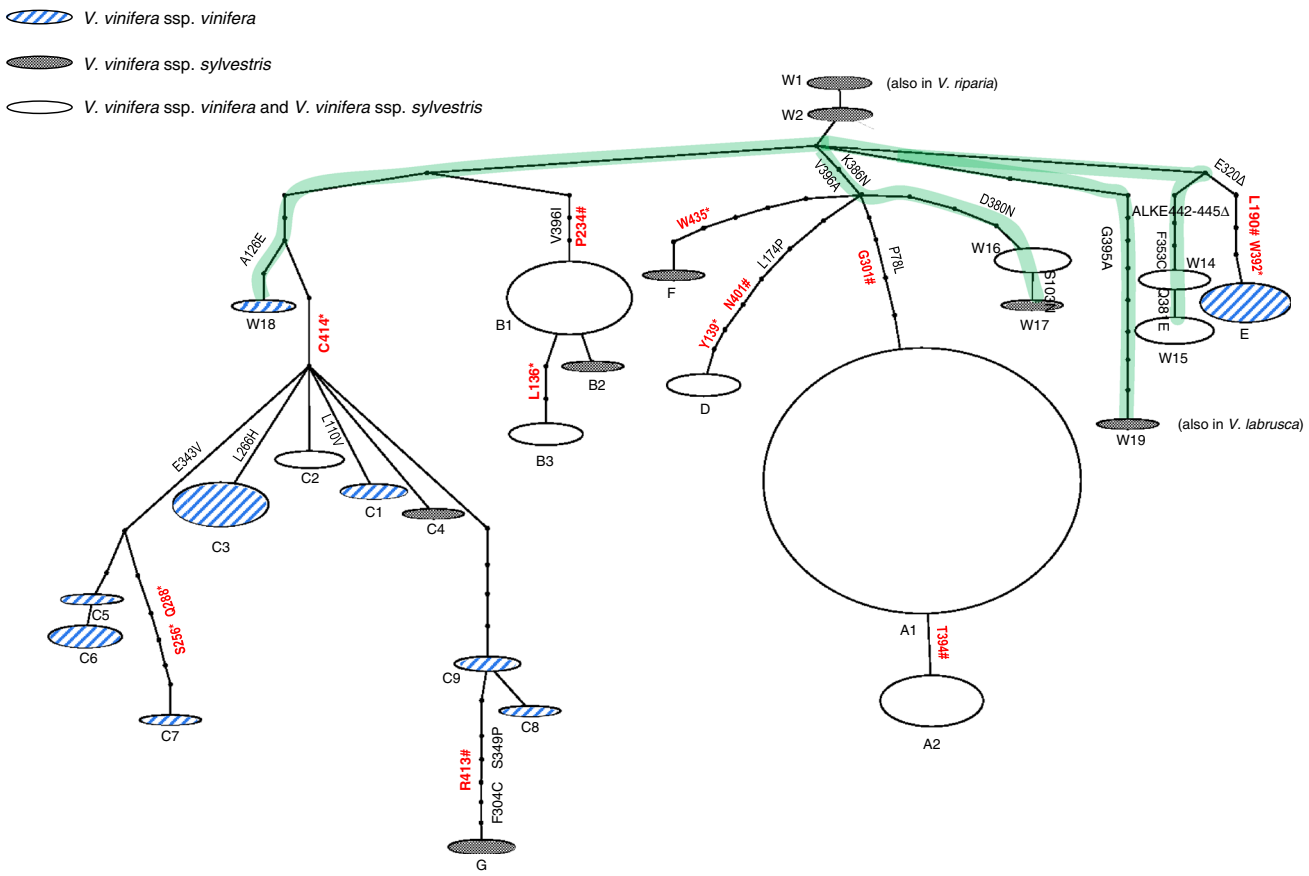

Fig. 2 Haplotype network depicting 26 *5GT* alleles found in *V. vinifera* ssp. *vinifera* and *V. vinifera* ssp. *sylvestris*, rooted by allele W1. Area of ellipse is proportional to sample size of an allele (total $n = 244$). Each branch between small nodes represents one mutational step (SNP or indel). All non-highlighted branches lead to frameshift or STOP codon mutations (*red colored letters*). Non-conservative amino acid changes and large indels are shown. Description of the mutational changes follows the same format as in Table 2 (color figure online)

acid changes were relatively rare. *V. vinifera* ssp. *vinifera* and *V. vinifera* ssp. *sylvestris* shared many same or similar *5GT* alleles, supporting the conclusions drawn from many previous studies that these two subspecies share close evolutionary relationships; a common assumption is that *sylvestris* represents the progenitor of *vinifera* (McGovern 2003; This et al. 2006). S*ylvestris* possessed two unique loss-of-function alleles (alleles F and G, Table 2) which were not detected in *vinifera*. This could be simply explained by the fact that *vinifera* captured only a portion of the s*ylvestris 5GT* alleles during domestication from s*ylvestris*, or that our survey of *vinifera* was not exhaustive. The fact that allele F was found in the hybrid grape 'Seibel 1000' which has *vinifera* in its pedigree, supported the latter explanation. Another possibility was that some of the unique alleles in s*ylvestris* were newly generated after *vinifera*'s domestication. Similarly, *vinifera* had a unique loss-of function allele (allele E) which was not found in *sylvestris*. While finite sample size could explain the failure to capture the missing *vinifera* allele in *sylvestris*, it is also possible that allele E was newly evolved in *vinifera*. The fact that *vinifera* has accumulated many new additional mutations superimposed on the loss-of-function allele resulting from a stop codon at nucleotide position 1242 (C414*) (alleles C1, C3, C5 to C9, Table 2) may provide further indication that the *5GT* gene has rapidly evolved independently in *vinifera* after its domestication.

By associating the allelic forms of *5GT* with the presence or absence of diglucoside anthocyanins, we were able to offer some explanations for why diglucoside anthocyanins were produced in some *V. vinifera* cultivars, but not in others. Similarly, we revealed that, while all the wild *Vitis* accessions produce diglucoside anthocyanins, the relative amount of diglucosylated anthocyanins varied among species, suggesting that some W-type *5GT* alleles exerted stronger effects than others and/or the genetic background played a role in producing diglucoside anthocyanins. Indeed, we observed that some *5GT* alleles had much higher levels of gene expression than others (unpublished data), presumably due to mutations in the promoter regions or other relevant regulatory elements. This information will enable breeders to decide which W-type *5GT* alleles should

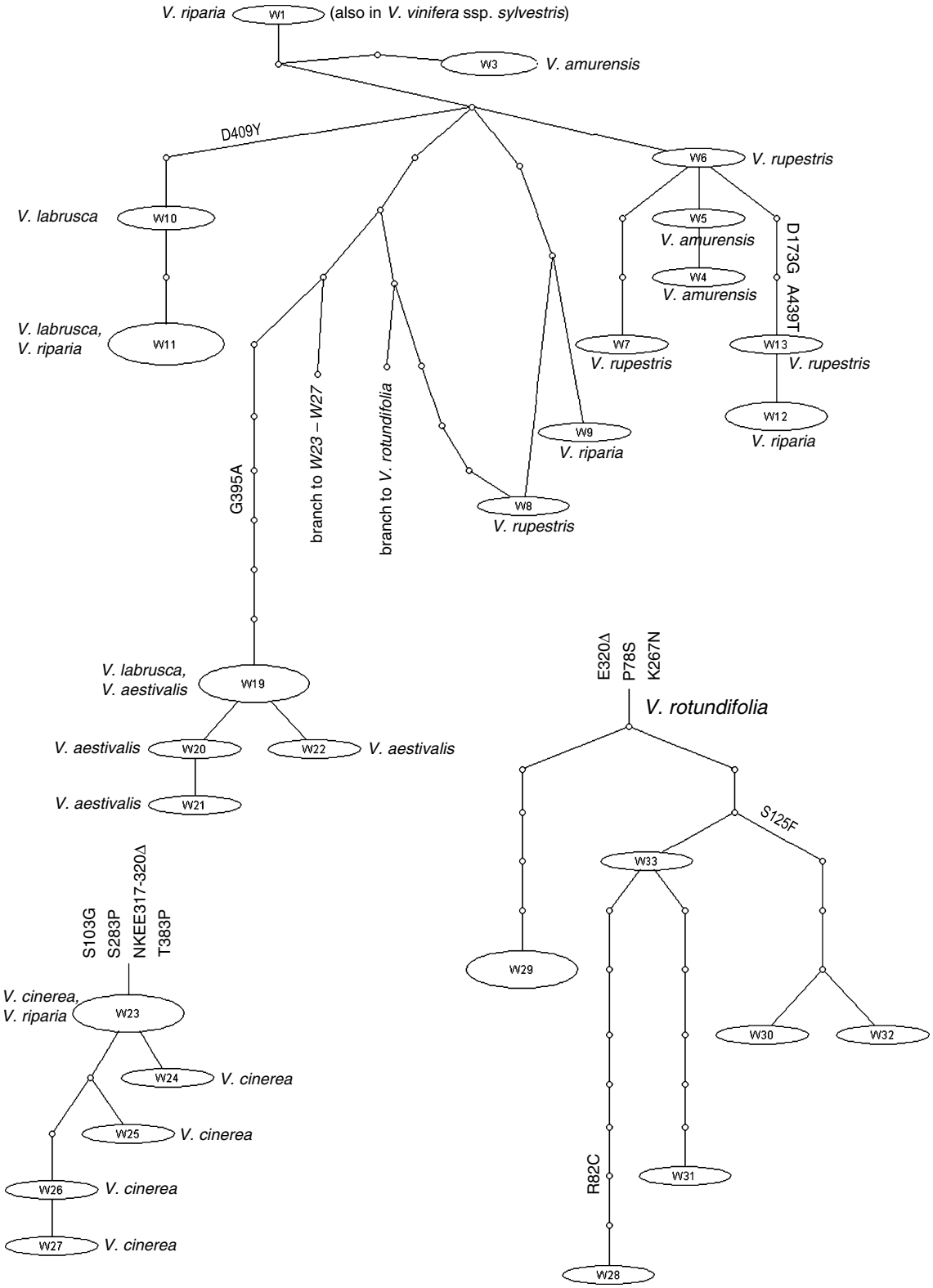

Fig. 3 Haplotype network depicting 27 *5GT* alleles found in wild grape species, rooted by allele W1. Area of ellipse is proportional to sample size of an allele (total $n = 44$). Each branch between small nodes represents one mutational step (SNP or indel). Non-conservative amino acid changes and large indels are shown. Description of the mutational changes follows the same format as in Table 2

be introduced and predict how much diglucoside anthocyanins might be produced when traits of interest from these wild *Vitis* species are incorporated into *V. vinifera*. The possibility of making such predictions was well demonstrated in the hybrid grapes investigated in this study. For example, 'Seibel 10878' and 'Seibel 1000' are two hybrid grapes well-known for their difference in producing diglucoside anthocyanins. 'Seibel 10878' mainly produced monoglucoside anthocyanins while 'Seibel 1000' produced abundant diglucoside anthocyanins (Robinson et al. 1966). Our study showed that 'Seibel 10878' contained two loss-of-function *5GT* alleles (A1 and C2) while 'Seibel 1000' contained a W-type *5GT* allele (W1) and an F allele (Tables 1, 2). The relative amount of diglucoside anthocyanins appeared to correlate well with the copy number of W-type *5GT* alleles present in hybrid grapes. For example, the hybrid 'Regent', which has a hybrid background of *V. vinifera*, *V. riparia* and some other *Vitis* species, contained about half the amount of diglucoside anthocyanins relative to *V. riparia* accessions (25.4 % versus 50–64 %) (Balík et al. 2013). Similarly, the content of diglucoside anthocyanins in 'Isabella', which carried one copy of W-type *5GT* allele (A1/W9), was about 13 % while it was 25–37 % in 'Clinton' which had two W-type *5GT* alleles (W1/W9) (De Rosso et al. 2012). Another example was that the hybrid table grape 'Kyoho' produced similar amount of diglucoside anthocyanins as *V. labrusca* (Table 1) (Li et al. 2013; Liang et al. 2012). This might be explained by the fact that 'Kyoho' contained two copies of W-type *5GT* alleles (W11 and W19), which enabled the hybrid to produce a similar amount of diglucoside anthocyanins as that in its wild parental species *V. labrusca*.

In conclusion, most *V. vinifera* cultivars carried only non-functional *5GT* alleles and therefore did not produce diglucoside anthocyanins. Many of these non-functional alleles resulted from frameshift indels and premature stop codons. Several red-colored *V. vinifera* cultivars, including 'Sidezites Proimo' and 'Zalovitico', carried one W-type *5GT* allele and produced detectable diglucoside anthocyanins. However, these genotypes were not widely represented in *V. vinifera* germplasm (Myles et al. 2011). Interestingly, several *V. vinifera* grapes with green/white berries also carried W-type *5GT* alleles, but they did not produce diglucoside anthocyanins due to the lack of anthocyanins. The functional *5GT* alleles from these cultivars can be effectively tracked by using allele-specific markers, should these cultivars be used in crosses in a grape breeding program. Similarly, we can also follow W-type *5GT* alleles in a grape breeding program when wild *Vitis* germplasm is introgressed into *V. vinifera*. Furthermore, we can use the *5GT* allele information of this study to develop a PCR-based diagnostic tool for predicting not only the presence of diglucoside anthocyanins, but also the origins of the contributing parent(s) in grape cultivars and their derived wines. Such a PCR-based tool will be much more sensitive and efficient than the current biochemical approaches, such as use of HPLC, for detecting the presence of diglucoside anthocyanins for wine quality control. In addition to the establishment of functional roles of various forms of *5GT* alleles in producing diglucoside anthocyanins and molecular basis for development of molecular markers for tracking these alleles in a grape breeding program, this study also revealed the networks of how these alleles are related to each other in the *Vitis* genus. The observations of many forms of mutations and a high level of polymorphisms involved in the *5GT* gene suggested that some *Vitis* genes can evolve rather rapidly.

Author contributions GYZ and YY conceived and designed the project; PC, BP, JP, MA and YY contributed materials; YY collected the sequencing data and, jointly with ZL, the HPLC data; YY, JAL and GYZ analyzed the data and drafted the paper; All authors read and approved the final submission.

Acknowledgments We wish to thank Andy Humiston of Department of Horticulture, Cornell University for providing his assistance in HPLC analysis of anthocyanins and Thomas Chao, Dawn Dellefave and Bill Srmack of USDA-ARS, Plant Genetic Resources Unit for their assistance in collecting tissue samples from the USDA-ARS Geneva *Vitis* Clonal Repository. USDA-ARS is an equal opportunity provider and employer. The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable. This project was funded by the United States Department of Agriculture, Agricultural Research Service, CRIS Project Number 1910-21000-020-00D.

Conflict of interest The authors declare that they have no conflict of interest.

References

Anderson DW, Julian EA, Kepner RE, Webb AD (1970) Chromatographic investigation of anthocyanin pigments in *Vitis cinerea*. Phytochemistry 9:1569–1578

Balík J, Kumšta M, Rop O (2013) Comparison of anthocyanins present in grapes of *Vitis vinifera* L. varieties and interspecific hybrids grown in the Czech Republic. Chem Pap 67:1285–1292

Ballinger WE, Maness EP, Nesbitt WB, Carroll DE (1973) Anthocyanins of black grapes of 10 clones of *Vitis rotundifolia*, Michx. J Food Sci 38:909–910

Bishop PD, Nagel CW (1984) Characterization of the condensation product of malvidin 3,5-diglucoside and catechin. J Agr Food Chem 32:1022–1026

De la Cruz AA, Hilbert G, Riviere C, Mengin V, Ollat N, Bordenave L, Decroocq S, Delaunay JC, Delrot S, Merillon JM, Monti JP, Gomes E, Richard T (2012) Anthocyanin identification and composition of wild *Vitis* spp. accessions by using LC-MS and LC-NMR. Anal Chim Acta 732:145–152

De Rosso M, Tonidandel L, Larcher R, Nicolini G, Ruggeri V, Dalla Vedova A, De Marchi F, Gardiman M, Flamini R (2012) Study of anthocyanic profiles of twenty-one hybrid grape varieties by liquid chromatography and precursor-ion mass spectrometry. Anal Chim Acta 732:120–129

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Eibach R, Töpher R (2003) Success in resistance breeding: "REGENT" and its steps into the market. Acta Hort (ISHS) 603:687–691

Flamini R, Tomasi D (2000) The anthocyanin content in berries of the hybrid grape cultivars Clinton and Isabella. Vitis 39:79–81

Ford CM, Boss PK, Hoj PB (1998) Cloning and characterization of *Vitis vinifera* UDP-glucose : flavonoid 3-*O*-glucosyltransferase, a homologue of the enzyme encoded by the maize Bronze-1 locus that may primarily serve to glucosylate anthocyanidins in vivo. J Biol Chem 273:9224–9233

Garcia-Viguera C, Bridle P (1999) Influence of structure on colour stability of anthocyanins and flavylium salts with ascorbic acid. Food Chem 64:21–26

Hausmann L, Neumann K, Eibach R, Zyprian E, Töpfer R (2009) Development of a molecular marker for an anthocyanin *5-O-glucosyl transferase* homologous gene of *Vitis* ssp. correlated with anthocyanin 3,5-diglucoside formation in berry skin. Acta Hort (ISHS) 827:457–460

Hausmann L, Eibach R, Toepfer R (2010) Alleles of anthocyanin 5-glucosyltransferase of *Vitis vinifera* are non-functional. Austr J Grape Wine Res 16(Suppl 1):A15

Huang ZL, Wang BW, Williams P, Pace RD (2009) Identification of anthocyanins in muscadine grapes with HPLC-ESI-MS. Lwt-Food Sci Technol 42:819–824

Janvary L, Hoffmann T, Pfeiffer J, Hausmann L, Topfer R, Fischer TC, Schwab W (2009) A double mutation in the anthocyanin 5-*O*-glucosyltransferase gene disrupts enzymatic activity in *Vitis vinifera* L. J Agr Food Chem 57:3512–3518

Kim M, Yoon SH, Jung M, Choe E (2010) Stability of meoru (*Vitis coignetiea*) anthocyanins under photochemically produced singlet oxygen by riboflavin. New Biotechnol 27:435–439

Li Y, Ma R, Xu Z, Wang J, Chen T, Chen F, Wang Z (2013) Identification and quantification of anthocyanins in Kyoho grape juice-making pomace, Cabernet Sauvignon grape winemaking pomace and their fresh skin. J Sci Food Agric 93:1404–1411

Liang ZC, Owens CL, Zhong GY, Cheng LL (2011) Polyphenolic profiles detected in the ripe berries of *Vitis vinifera* germplasm. Food Chem 129:940–950

Liang ZC, Yang YZ, Cheng LL, Zhong GY (2012) Polyphenolic composition and content in the ripe berries of wild *Vitis* species. Food Chem 132:730–738

Liang ZC, Yang YZ, Cheng LL, Zhong GY (2013) Characterization of polyphenolic metabolites in grape hybrids. Vitis 52:51–59

Lijavetzky D, Cabezas JA, Ibanez A, Rodriguez V, Martinez-Zapater JM (2007) High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera L.*) by combining a re-sequencing approach and SNPlex technology. BMC Genom 8:424

Mazza G (1995) Anthocyanins in grapes and grape products. Crit Rev Food Sci 35:341–371

Mazza G, Brouillard R (1987) Color stability and structural transformations of cyanidin 3,5-diglucoside and four 3-deoxyanthocyanins in aqueous solutions. J Agr Food Chem 35:422–426

McGovern PE (2003) Ancient wine: the search for the origins of viniculture. Princeton University Press, Princeton

Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK, Prins B, Reynolds A, Chia JM, Ware D, Bustamante CD, Buckler ES (2011) Genetic structure and domestication history of the grape. Proc Natl Acad Sci USA 108:3530–3535

Okamoto G (2007) Poor berry set in tetraploid grapes-causes and improvement of vineyard practices. J ASEV Jpn 18:94–106

Picariello G, Ferranti P, Chianese L, Addeo F (2012) Differentiation of *Vitis vinifera* L. and hybrid red grapes by matrix-assisted laser desorption/ionization mass spectrometry analysis of berry skin anthocyanins. J Agr Food Chem 60:4559–4566

Picariello G, Ferranti P, Garro G, Manganiello G, Chianese L, Coppola R, Addeo F (2014) Profiling of anthocyanins for the taxonomic assessment of ancient purebred *V. vinifera* red grape varieties. Food Chem 146:15–22

Revilla E, Carrasco D, Benito A, Arroyo-Garcia R (2010) Anthocyanin composition of several wild grape accessions. Am J Enol Viticult 61:536–543

Riahi L, Zoghlami N, Dereeper A, Laucou V, Mliki A, This P (2013) Single nucleotide polymorphism and haplotype diversity of the gene *NAC4* in grapevine. Ind Crop Prod 43:718–724

Robinson WB, Weirs LD, Bertino JJ, Mattick LR (1966) The relation of anthocyanin composition to color stability of New York state wines. Am J Enol Vitic 17:178–184

Sandhu AK, Gu LW (2010) Antioxidant capacity, phenolic content, and profiling of phenolic compounds in the seeds, skin, and pulp of *Vitis rotundifolia* (Muscadine grapes) as determined by HPLC-DAD-ESI-MS^n. J Agr Food Chem 58:4681–4692

Sims CA, Morris JR (1985) A comparison of the color components and color stability of red wine from Noble and Cabernet Sauvignon at various pH levels. Am J Enol Viticult 36:181–184

Sims CA, Morris JR (1986) Effects of acetaldehyde and tannins on the color and chemical age of red muscadine (*Vitis rotundifolia*) wine. Am J Enol Viticult 37:163–165

Templeton AR, Crandall KA, Sing CF (1992) A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data.III. cladogram estimation. Genetics 132:619–633

This P, Lacombe T, Thomas MR (2006) Historical origins and genetic diversity of wine grapes. Trends Genet 22:511–519

Vanburen JP, Bertino JJ, Robinson WB (1968) Stability of wine anthocyanins on exposure to heat and light. Am J Enol Viticult 19:147–154

Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, FitzGerald LM et al (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS ONE 2:e1326

Wang J, De Luca V (2005) The biosynthesis and regulation of biosynthesis of Concord grape fruit esters, including 'foxy' methylanthranilate. Plant J 44:606–619

Wu X, Prior RL (2005) Systematic identification and characterization of anthocyanins by HPLC-ESI-MS/MS in common foods in the United States: fruits and berries. J Agr Food Chem 53:2589–2599

Yamazaki M, Gong Z, Fukuchi-Mizutani M, Fukui Y, Tanaka Y, Kusumi T, Saito K (1999) Molecular cloning and biochemical characterization of a novel anthocyanin 5-*O*-glucosyltransferase by mRNA differential display for plant forms regarding anthocyanin. J Biol Chem 274:7405–7411

Yamazaki M, Yamagishi E, Gong ZZ, Fukuchi-Mizutani M, Fukui Y, Tanaka Y, Kusumi T, Yamaguchi M, Saito K (2002) Two flavonoid glucosyltransferases from *Petunia hybrida*: molecular cloning, biochemical properties and developmentally regulated expression. Plant Mol Biol 48:401–411

You Q, Chen F, Wang X, Sharp JL, You YR (2012) Analysis of phenolic composition of Noble muscadine (*Vitis rotundifolia*) by HPLC-MS and the relationship to its antioxidant capacity. J Food Sci 77:C1115–C1123

Zhao Q, Duan CQ, Wang J (2010) Anthocyanins profile of grape berries of *Vitis amurensis*, its hybrids and their wines. Int J Mol Sci 11:2212–2228

Zhu L, Zhang YL, Lu J (2012) Phenolic contents and compositions in skins of red wine grape cultivars among various genetic backgrounds and originations. Int J Mol Sci 13:3492–3510