Polymorphisms in *VvPel* associate with variation in berry texture and bunch size in the grapevine

A.M. VARGAS[1], C. FAJARDO[1]*, J. BORREGO[1], M.T. DE ANDRÉS[1] and J. IBÁÑEZ[1,2]

[1] Instituto Madrileño de Investigación y Desarrollo Rural, Agrario y Alimentario (IMIDRA), Finca 'El Encín', Ctra. A2, Km 38.200, 28805 Alcalá de Henares, Madrid, Spain

[2] Instituto de Ciencias de la Vid y del Vino (CSIC, Gobierno de La Rioja, Universidad de La Rioja), Complejo Científico Tecnológico, C/ Madre de Dios 51, 26006 Logroño, Spain

* *Present address:* Facultad de Veterinaria, Departamento de Bioquímica y Biología Molecular IV, Universidad Complutense de Madrid, Avda. Puerta de Hierro s/n, 28040 Madrid, Spain

Corresponding author: Dr Javier Ibáñez, email javier.ibanez@icvv.es

Abstract

Background and Aims: The identification of the genes involved in traits of interest is one of the main aims of current plant genetic studies. Although berry texture is a quality trait of great importance in table grapes, no quantitative trait loci or related gene has been described for this trait so far.

Methods and Results: In this work, the gene *VvPel*, which codes for a pectate lyase in the grapevine, was selected as a candidate gene for an association study. This gene was sequenced in a core collection of 96 table grape accessions, which was also characterised for several bunch and berry morphological traits, including diverse texture parameters measured by means of a texture meter. The population structure, the nucleotide and haplotype diversity, the protein structure, the existence of selection, and the linkage disequilibrium (LD) were evaluated. Thirty-two single nucleotide polymorphisms (SNPs) and 15 haplotypes were identified in the *VvPel* sequence, and LD was low enough to provide a high-resolution power. Some of the polymorphisms associated significantly with texture parameters and with bunch size, explaining part of the variation found for those traits.

Conclusions: Among the polymorphisms found in the gene *VvPel*, S1027 and S405/S441 in relation to berry texture, and S48/S1030 in relation to bunch size, are the most promising and would need to be investigated further as putative causal polymorphisms or markers for the respective traits.

Significance of the Study: The present work is the first genetic study on berry texture in the grapevine, and constitutes a starting point for a deeper understanding of the genetic basis of the natural variation for berry texture and bunch size.

Keywords: *association genetics, berry firmness, bunch architecture, pectate lyase, table grape,* Vitis vinifera *L.*

Introduction

The grapevine (*Vitis vinifera* L.) is a crop with worldwide economic importance. In general, genetic research in this crop has focused on traits related to wine composition, such as berry colour and polyphenol content, as well as to disease resistance (mildew, oidium). In the last few decades, Muscat flavour and the absence of seeds have been the main traits studied in table grapes. Two approaches have been used for these studies. One is the well-known quantitative trait loci (QTL) analysis, through genotyping, mapping, and phenotyping one or several progenies, normally from a single cross. This approach allowed detecting and mapping; for example, a QTL for seedlessness, or QTLs for berry weight and for some phenology traits (Cabezas et al. 2006, Costantini et al. 2008). The second approach searches for an association between molecular polymorphisms in a candidate gene and phenotypic variation, using a collection of cultivars rather than a progeny as base material. Two association studies in grapevine have been recently published: one between DNA sequence variation at *VvmybA* genes and the colour of the grape (Fournier-Level et al. 2009), and the other between the gene *VvDXS* and Muscat flavour (Emanuelli et al. 2010). The former approach gives information about genomic regions, while the latter identifies in a candidate gene the possible polymorphism/s responsible for the phenotypic change.

Berry texture is a quality trait of great importance in table grapes for several reasons. In general, consumers prefer crisp berries, with firm flesh and thin skin, but berry texture is also important for producers, since fruit resistance during transport depends in part on it. Texture is a trait difficult to measure quantitatively. There is a descriptor for 'berry firmness' (Office International de la Vigne et du Vin 2007), but it is a sensory descriptor, and thus somewhat subjective. Several efforts have been made to achieve objective and quantitative measures by the application of rheometers and, more recently, texture meters. These instruments measure mechanical variables, most often force and deformation (distance), which have been related to sensory properties of the grape berry (Vargas et al. 2001, Sato and Yamada 2003, Rolle et al. 2011), and also to extractability of anthocyanins (Maury et al. 2009, Rolle et al. 2009, Zouid et al. 2010). Nevertheless, this trait has received little attention in the grape from a genetic perspective.

doi: 10.1111/ajgw.12029

Two growth phases occur during grape ripening: during the second phase, starting at veraison, cell expansion begins, in which cell walls become less tense because of an increase in pectin solubilisation and xyloglucan depolymerisation, while new components are synthesised. Expression studies during grape development show the involvement of distinct expansins and enzymes, such as polygalacturonase, β-galactosidase, pectin methylesterase, celulase, xyloglucan endotransglycosylase, pectin esterase, and pectate lyase (Nunan et al. 2001, Ishimaru and Kobayashi 2002, Terrier et al. 2005).

The pectate lyase (PL) is a cell wall degradation enzyme that catalyses the eliminative cleavage of de-esterified pectin, which is a major component of the primary cell walls of many higher plants. Previous studies in several species have reported a relation between this enzyme and fruit softening during ripening. A high level of expression during this developmental stage has been identified in banana (Domínguez-Puigjaner et al. 1997, Pua et al. 2001), strawberry, (Medina-Escobar et al. 1997, Jiménez-Bermúdez et al. 2002, Benítez-Burraco et al. 2003), mango (Chourasia et al. 2006), and grape (Nunan et al. 2001, Ishimaru and Kobayashi 2002), but in general, the activity, localisation, and regulation of this enzyme are little known in comparison with that of polygalacturonase (Brummell and Harpster 2001). Jiménez-Bermúdez et al. (2002) reported that the suppression of the expression of a PL gene in strawberry coincided with firmer fruit which led to an extended good postharvest life, thus supporting the commercial interest of studying this enzyme. Later, Benítez-Burraco et al. (2003) detected three PL transcripts in strawberry (*plA*, *plB*, and *plC*) that were expressed exclusively in fruit and mainly during ripening.

In the grape, the existence of at least three PL isoforms has been also published (Glissant et al. 2008). The first was described by Nunan et al. (2001) and was expressed during flowering until 28 days after bloom, and then in the period between 70 days after bloom and 96 days after bloom; expression was especially intense just after veraison (at 70 days after bloom) compared to that of the other enzymes studied (polygalacturonase, β-galactosidase, pectin methylesterase, cellulase, and xyloglucan endotransglycosylase). The second corresponds with a PL studied by Schlosser et al. (2008), the expression of which increased during veraison and ripening. This study focuses on the third PL which has been previously described by Papadakis and Kanellis (2000, GenBank accession AF243475); it is expressed in berries from pea size until veraison according to Glissant et al. (2008), but more or less uniformly during berry development according to other authors (Deluc et al. 2007, Pilati et al. 2007, Lund et al. 2008).

The aim of this work was to evaluate the role of pectate lyase in the natural variation of berry texture and of other important traits in table grapes. This study was approached through an association study using a candidate gene. This type of study, besides the statistical association itself, requires the analysis of several aspects (Whitt and Buckler 2003): the population structure of the collection, since the most serious false positives can result when unlinked markers produce a positive association because of the underlying population structure; the linkage disequilibrium (LD) extension in the candidate gene, directly responsible for the resolution power of the study; and the possible existence of selective pressure on the candidate gene, because if a gene, or part of it, has been subjected to selection during the domestication or the improvement process, probably it is involved in an important trait. In addition, knowing that the associated polymorphism produces a change in the protein structure can provide critical support, especially if functionality is presumably affected.

Materials and methods

Plant material and phenotypic data

The plant material consisted of a core collection of 96 table grape accessions (Table 1), which was derived from the collection of 322 table grape accessions maintained at the germ plasm bank of El Encín (IMIDRA, Madrid, Spain). It was built using the Maximization Strategy, implemented in the software MStrat 4.1 (Gouesnard et al. 2001), by joining three independent core sub-collections, each one containing about 90% of the total variability found for 20 microsatellite markers/simple sequence repeats (SSRs) and 55 morphological and phenological descriptors. Each core sub-collection was constructed using a reiterated procedure, incorporating in each step the accessions more frequently represented in the different collections obtained in 20 replicates. To build the second and the third sub-collections, the accessions integrated in previous sub-collections and their synonyms and sports were removed from the analysis.

The whole collection was characterised during 2004–2006 using 48 qualitative and quantitative morphological and agronomic descriptors; doubtful data were resolved in 2007. Seventeen quantitative and one qualitative descriptors, related to bunch and berry traits, were used for the association analysis (Table 2). Five bunches were described for accession (except in 2004, when only one bunch was measured), and berry dimension was measured in 15 random berries with the average then finally used for analyses.

Berry texture was measured with a TA-XT Plus texture meter (Stable Micro Systems, Surrey, England) using a 2-mm diameter punch probe in 2006 and 2007. Rolle et al. (2012) have reviewed the application of the texture meter to the grapevine. The test starts when the probe proceeds to move down onto the grape, and a rapid rise in force is observed. During this stage, the sample is deforming under the applied force, but there is no puncturing of the tissues. This stage ends abruptly when the probe punctures through the skin and begins to penetrate the flesh of the berry. A profile of force against distance is obtained, where the following parameters are measured: (i) force at 10%; and (ii) force at 20% (force in Newtons needed to deform 10 or 20% of the berry diameter, respectively); (iii) rupture force (force needed to break the berry skin); (iv) rupture distance (distance moved by the probe from the origin to the skin rupture point); (v) slope 10%; and (vi) rupture slope (slope of the curve force–distance at 10% of berry diameter deformation and at the skin rupture point, respectively); (vii) rupture area (area under the curve between the origin and the rupture distance point); and (viii) deformation rate (ratio between rupture distance and berry width). In 2005, texture was measured by means of a TA-XT2 texture meter with a 5-cm diameter probe; however, only the parameter force at 20% was measured, and the measurements are not directly comparable with the data of 2006 and 2007 because of the different probe.

Correlation analysis between traits and between years was done with SPSS 15 (IBM, New York, USA), using the Pearson correlation coefficient for quantitative traits and Kendall Tau-b correlation coefficient for qualitative traits.

Population structure and kinship matrix

The core collection was subjected to structure analysis with the software Structure 2.0 (Pritchard et al. 2000). Eleven non-linked nuclear SSRs were used: VMC1B11 (Zyprian 2005, GenBank accession BV681754) VMC4F3-1 (*Vitis* Microsatellite Consortium), VVIH54, VVIN16, VVIP60, VVIV37 (Merdinoglu

Table 1. Core collection of 96 cultivars, sorted according to the inferred genetic structure with two sub-populations (Q1 and Q2). The table includes the individual's estimated membership fraction in each of the two sub-populations, and the geographic origin, breeder, and pedigree, if known (according to VIVC, http://www.vivc.de).

Population	Prime name	Q1	Q2	Origin	Breeder	Pedigree
Q1	Helvany	0.984	0.016	Lebanon		
	De cilindro	0.983	0.017	Spain		
	Ragol	0.979	0.021	Spain		
	Verico	0.978	0.022	Cyprus		
	Dattier noir	0.974	0.026			
	Jerónimo	0.974	0.026	Spain		
	Dabouki	0.973	0.027	Armenia		
	Imperial roja	0.973	0.027	Spain	Unknown	Ohanes × Ragol
	Zeini Abiad	0.973	0.027	Lebanon		
	Canner seedless	0.971	0.029	USA	H.P. Olmo	Hunisa × Sultanina
	Negra tardía	0.969	0.031	Spain		
	Ferral	0.966	0.034	Portugal		
	De cuerno	0.964	0.036	Spain		
	Doroni Maceron	0.963	0.037			
	Talismano	0.962	0.038	Spain	Unknown	Ohanes × Italia
	Ahmeur bou Ahmeur	0.955	0.045	Algeria		
	Korinthiaki	0.954	0.046	Greece		
	Luglienga bianca	0.953	0.047	Spain		
	Tebrizi	0.953	0.047	Azerbeijan		
	El Farryali	0.952	0.048	Morocco		
	Nehelescol	0.951	0.049	Israel		
	Conca D'Oro	0.947	0.053	Italy	F. Paulsen	Nota S.P. × Chasselas rose
	Negra dorada	0.947	0.053	Spain		
	Pirovano 18	0.947	0.053	Italy	A. Pirovano	Chasselas musqué × Sultanina
	Oscari rose	0.946	0.054	Egypt		
	Castellano morado	0.940	0.060	Spain		
	Rosaki noir de semis	0.936	0.064			
	Moscato di Terracina	0.935	0.065	Italy		
	Chaouch blanc	0.932	0.068	Turkey		
	Pizzutello nero	0.927	0.073	Spain	Unknown	Cornichon blanc × Prune de Cazouls
	Schiras	0.927	0.073			
	Bruni 1	0.927	0.073	Italy	B. Bruni	
	Aledo	0.926	0.074	Spain		
	Corinto bianco	0.924	0.076	Spain		
	Agostenga	0.921	0.079	Italy		
	Tempranilla blanca	0.913	0.087	Spain		
	Aledo Real	0.905	0.095	Spain		
	Cornichon violet	0.901	0.099			
	Khalili Belyi	0.900	0.100	Azerbeijan		
	Turki	0.899	0.101	Tunisia		
	Ophtalmo	0.896	0.104	Cyprus		
	Kishmish Chernyi	0.886	0.114	Uzbekistan		
	Sultanina	0.879	0.121	Turkey		
	Chaouch rozovyi	0.859	0.141	Turkey		
	Beba	0.847	0.153	Spain		
	Albillo Real	0.845	0.155	Spain		
	Foster's white seedless	0.840	0.160			
	Planta mula	0.771	0.229	Spain		
	Ohanes red	0.704	0.296	Portugal		
	Corniola	0.647	0.353	Italy		
	Beauty seedless	0.610	0.390	USA	H.P. Olmo	Koenigin der Weingaerten × Black seedless
	Danlas	0.597	0.403	France	P. Truel	Dabouki × Chasselas blanc

Table 1. (continued)

Population	Prime name	Q1	Q2	Origin	Breeder	Pedigree
Q2	Cyperntraube	0.484	0.516	Cyprus		
	Lasina	0.478	0.522	Croatia		
	Dimyat	0.473	0.527	Bulgaria		P1: Heunisch Weiss
	Emerald seedless	0.408	0.592	USA	H.P. Olmo	Emperor × Sultana moscata
	Pirovano 166A	0.368	0.632	Italy	A. Pirovano	Delizia di Vaprio × Black Monukka
	Trentham black	0.351	0.649	United Kingdom		
	Cape Currant	0.323	0.677	South Africa		
	Barbableu	0.321	0.679	Italy	Unknown	Olivetta nera × Muscat Hamburg
	Clotilde Prosperi	0.280	0.720	Italy	V. Prosperi	(Regina × Sabalkanskoi) × Italia
	Rodi	0.257	0.743	Italy	A. Pirovano	Muscat of Alexandria × Sultanina
	Kover Szoeloe	0.219	0.781	Hungary	Unknown	Kadarka × Munkatsy Jozsef
	Graziella I	0.196	0.804	Italy	Unknown	Pirovano 89 × Maria Pirovano
	Black Alicante	0.169	0.831	Spain		
	Barlinka	0.167	0.833	South Africa		
	Queen	0.155	0.845	USA	H.P. Olmo	Muscat Hamburg × Sultanina
	Pizzutello moscato biondo	0.106	0.894	Italy	A. Pirovano	Pirovano 22 × Aurora
	Malingre precoce	0.101	0.899	France	Malingre	Bicane × Pinot
	Bruni 415	0.094	0.906	Italy	B. Bruni	Muscat of Alexandria × Bruni 12
	Italia x Sultanina V-6	0.094	0.906	Bulgaria		
	Vivona 378	0.087	0.913	Italy	A. Vivona	Muscat of Alexandria × Bruni 415
	Maria Rosa	0.076	0.924	Italy	A. Pirovano	
	Attilio Ragionieri	0.069	0.931	Italy	Unknown	Moscato d'Adda × Pirovano 61
	Latina	0.069	0.931	Italy	A. Pirovano	Italia × David
	Roi des precoces	0.069	0.931	France	Unknown	Cinsaut × Pinot
	Marocain noir	0.068	0.932	France		
	Alba Magna	0.063	0.937	Italy	A. Pirovano	Moscato d'Adda × Foster's white seedling
	Bogni 8	0.060	0.940	Italy	Bogni	Madeleine Angevine × Angelo Pirovano
	Kharistvala Kolkhuri	0.051	0.949	Georgia		
	Galletta rosa	0.048	0.952	Italy	A. Pirovano	Pirovano 122 × Aurora
	Bruni 125	0.047	0.953	Italy	B. Bruni	Alphonse Lavallé × Agostenga rosa
	Lady Downe's seedling	0.044	0.956	United Kingdom	Unknown	Marocain noir × Muscat of Alexandria
	Mistress Hall	0.043	0.957	United Kingdom	Unknown	Marocain noir × Muscat of Alexandria
	Viola	0.035	0.965	Italy	G. Dalmasso	Muscat Hamburg × Pirovano 62
	Primiera	0.031	0.969	Italy	A. Pirovano	Delizia di Vaprio × Madeleine Angevine
	Chasselas violet	0.030	0.970	France		
	General de la Marmora	0.029	0.971	France	Moreau Robert	
	Leopold III	0.029	0.971	Belgium		Alphonse Lavallée 4N
	Schiava grossa	0.027	0.973	Italy		
	Thalloczy Lajos	0.025	0.975	Hungary	J. Mathiasz	Sicilien × Muscat of Alexandria
	Muscat of Alexandria	0.024	0.976	Italy	Unknown	Muscat blanc à petits grains × Axina de tres bias
	Trieste	0.024	0.976	Italy	A. Pirovano	Moscato d'Adda × Pirovano 61
	Fusca	0.022	0.978	Italy	A. Pirovano	Garganega × Prunella moscato
	Pirovano 771	0.020	0.980	Italy	A. Pirovano	
	Pirovano 671	0.019	0.981	Italy	A. Pirovano	

Table 2. Traits used for the association study with *VvPel* polymorphisms.

Organ / trait	Descriptors
Fertility	Number of inflorescences per shoot
Bunch	Bunch width
	Bunch length
	Bunch weight
	Length of peduncle
Berry	Berry width
	Berry length
	Berry weight
	Berry volume
	CIRG
Yield	Juice yield
Berry texture	Flesh firmness (sensory)
	Force at 10%
	Force at 20%
	Rupture force
	Rupture slope
	Rupture area
	Deformation rate

CIRG, colour index for red grapes.

et al. 2005), VVMD7, VVMD21, VVMD28, VVMD32 (Bowers et al. 1996, 1999), and VVS2 (Thomas and Scott 1993). A model, in which a putative number between one and seven populations and correlated allele frequencies (Falush et al. 2003), was assumed. Monte Carlo Markov Chain run length period of 100 000, with 100 000 burn-in steps, and 20 iterations for each number of putative populations, was used. The populations number was selected with the Evanno criterion (Evanno et al. 2005).

A kinship matrix was constructed with TASSEL 2.1 (Bradbury et al. 2007) using another set of 12 non-linked microsatellite markers: ssrVrZAG29, ssrVrZAG62, ssrVrZAG67, ssrVrZAG83 (Sefc et al. 1999), VVMD5, VVMD25, VVMD27 (Bowers et al. 1996, 1999; VVIB01, VVIN73, VVIP31, VVIQ52 and VVIV67 (Merdinoglu et al. 2005).

DNA extraction, polymerase chain reaction (PCR), separation of fragments, and data analysis are described by Ibáñez et al. (2009).

Gene amplification and sequencing

Primers for the gene *VvPel* (AF243475.1) were designed using Primer3 0.2 (Rozen and Skaletsky 2000). Three fragments (a–c) were sequenced with the following primers: Pel-a Fw (5′-CCCTCGTTTGCCAGTTTATG-3′), Pel-a Rv (5′-GCAGAGAA TGCCCAGGTAAG-3′), Pel-b Fw (5′-ACTTCCACCAATGGC ATACA-3′), Pel-b Rv (5′-GGCGGAAAGAATGGTAGAAT-3′), Pel-c Fw (5′-TGGCGTAATGGATGGTAATG-3′), and Pel-c Rv (5′-AAAATAGGCGACGGAAAAGG-3′). PCR amplifications were performed using 3.5 U of a 3′→5′ exonuclease activity Taq-polymerase (Expand Long Range, Roche, Indianapolis, IN, USA), 0.5 mM dNTPs, buffer 1× (stock 5×: 20 mM Tris-HCl pH 7.5, 100 mM KCl, 0.1 mM EDTA, 1 mM dithiothreitol, 50% glycerol, 0.5% Tween 20, 0.5% Nonidet P40) with 2.5 mM $MgCl_2$, 0.35 μM primers and 5 ng of DNA, in a final volume of 50 μL. The PCR program consisted of an initial denaturing step of 92°C for 4 min followed by 10 cycles of 92°C for 1 min, 50°C for 1 min, and 68°C for 2.5 min, 30 cycles of 92°C for 1 min, 50°C for 1 min, and 68°C for 2.5 min+20 s/cycle, and a final extension step of 68°C for 30 min. The amplified DNA fragments were sequenced at the Genomic Unit of the Parque Científico de Madrid (Spain). Sequencing analysis and alignment were performed with SeqScape 2.5 (Applied Biosystem, Foster City, CA, USA) and BioEdit 7.0.0. (Ibis Biosciences, Carlsbad, CA, USA; Hall 1999). Sequences match the GenBank accession numbers JQ743924-JQ744019.

Prediction of secondary structure

The amino acid sequence for each haplotype was obtained, and secondary structure was predicted by means of the software PSSFinder, both implemented in Softberry, Inc. (NY, USA; http://linux1.softberry.com/berry.phtml).

Statistical tests and association study

Haplotypes were inferred with the software PHASE 2.1 (Stephens et al. 2001). The diversity parameters π (Nei 1987) and θ_w (Watterson 1975) and the Neutrality Test of Fu and Li (Fu and Li 1993) were estimated with DnaSP 4 (Rozas and Rozas 1999). Linkage disequilibrium (LD) analyses were carried out with DnaSP 4 and TASSEL 2.1. The association test was done using a mixed linear model (MLM) implemented in TASSEL, with individual SNPs and with the inferred haplotypes. Some variables were transformed to better fit assumptions of normality.

Results

Core collection

The core collection was built with the objective of having at least a triple representation for most of the classes/alleles present. For this purpose, three sub-collections were independently constructed and then joined. The first one contained 36 accessions (11% of the global collection) and represented 96% of the total (allelic and morphological) variability, the second one presented 32 accessions and represented 93% of the remaining variability, and the third one contained 28 accessions and 88% of the remaining variability (Table 3). The final core collection contained 96 accessions and 98% of the total phenotypic and genotypic variability detected in the original collection.

Population structure

The structure analysis suggested the existence of two populations (Q1 and Q2) (Table 1). The population Q1 is comprised of 52 accessions, mainly ancient cultivars, from the Iberian Peninsula (Spain or Portugal, 38%) or with an Oriental origin (35%). The population Q2 comprises 44 accessions, most of which originated relatively recently from crosses with common parents such as Muscat of Alexandria or Chasselas by European breeders like Pirovano (Ibáñez et al. 2009, Vargas et al. 2009). This structure, obtained from microsatellite data, was examined for important morphological traits: bunch length and width and berry firmness. *T*-tests showed significant differences ($P < 0.05$) between the two populations for bunch length (2006 and 2007) and berry firmness (2004, 2006, and 2007).

Nucleotide and haplotype diversity

The complete coding region of the gene *VvPel* could not be sequenced because of the presence of a poly T giving rise to an indel placed at 186 bp from the initiation codon. The sequence

Table 3. Core sub-collections of 36, 32, and 28 accessions obtained from a global collection of 322 table grape accessions, which altogether constitute the core collection of 96 accessions used for the association analysis.

Sub-collection 1		Sub-collection 2		Sub-collection 3	
Accession name	**No accession**	**Accession name**	**No accession**	**Accession name**	**No accession**
Ahmeur bou Ahmeur	BGVCAM1375	Pirovano 620	BGVCAM2588	Agostenga	BGVCAM0923
Aledo	BGVCAM1191	Barlinka	BGVCAM1427	Alba Magna	BGVCAM1216
Aledo Real	BGVCAM1402	Black Alicante	BGVCAM1477	Albillo	BGVCAM1105
Attilio Ragionieri	BGVCAM1474	Bogni 8	BGVCAM1377	Beauty seedless	BGVCAM0797
Barbableu	BGVCAM0941	Chaouch blanc	BGVCAM1760	Bruni 125	BGVCAM1480
Beba dorada de Jaén	BGVCAM2341	Corinto bianco	BGVCAM1447	Bruni 415	BGVCAM1723
Korinthiaki	BGVCAM1212	Cornichon violet	BGVCAM1448	Clotilde Prosperi	BGVCAM1355
Kishmish Chernyi	BGVCAM1722	Corniola	BGVCAM1220	Dabouki	BGVCAM1384
Canner seedless	BGVCAM1378	Viola	BGVCAM1440	Dattier noir	BGVCAM1491
Cape Currant	BGVCAM1726	Fusca	BGVCAM1221	De cilindro	BGVCAM1610
Castellano morado	BGVCAM2246	General de la Marmora	BGVCAM1388	Foster's white seedless	BGVCAM2355
Chaouch rozovyi	BGVCAM1727	Helvany	BGVCAM1496	Frankenthal	BGVCAM1356
Chasselas violet	BGVCAM1798	Jerónimo	BGVCAM0825	Graziella I	BGVCAM1325
Conca d'Oro	BGVCAM1434	Lady Downe's seedling	BGVCAM1497	Kharistvala Kolkhuri	BGVCAM0984
Danlas	BGVCAM2349	Lasina	BGVCAM1238	Imperial roja	BGVCAM1615
De cuerno	BGVCAM2078	Latina	BGVCAM1326	Kover Szoeloe	BGVCAM2255
Doroni Maceron	BGVCAM2351	Mistress Hall	BGVCAM1498	Cyperntraube	BGVCAM1213
El Farryali	BGVCAM1494	Moscato di Terracina	BGVCAM1357	Muscat of Alexandria	BGVCAM1997
Emerald seedless	BGVCAM1472	Negra dorada	BGVCAM2270	Pirovano 166A	BGVCAM?003
Ferral	BGVCAM0815	Negra tardía	BGVCAM1410	Pirovano 771	BGVCAM1328
Marocain noir	BGVCAM1271	Ophtalmo	BGVCAM1214	Pizzutello moscato biondo	BGVCAM1329
Italia X Sultanina V-6	BGVCAM1331	Pirovano 18	BGVCAM0857	Pizzutello nero	BGVCAM1225
Jouanenc	BGVCAM0826	Pirovano 671	BGVCAM1504	Ragol	BGVCAM1426
Khalili Belyi	BGVCAM2362	Malingre precoce	BGVCAM1746	Red Ohanes	BGVCAM2470
Leopold III	BGVCAM1338	Primiera	BGVCAM0859	Rodi	BGVCAM1265
Maria rosa	BGVCAM1327	Queen	BGVCAM1397	Bruni 1	BGVCAM0866
Nehelescol	BGVCAM1394	Sultanina	BGVCAM1077	Dymiat	BGVCAM1230
Oscari rose	BGVCAM1272	Talismano	BGVCAM1330	Thalloczy Lajos	BGVCAM1361
Planta mula	BGVCAM1158	Tempranilla blanca	BGVCAM2281		
Roi des precoces	BGVCAM1506	Trentham black	BGVCAM1510		
Rosaki noir de semis	BGVCAM1507	Verico	BGVCAM1215		
Shiradzouli belyi	BGVCAM1065	Vivona 378	BGVCAM1437		
Schiras	BGVCAM1066				
Trieste	BGVCAM1755				
Turki	BGVCAM1511				
Zeini Abiad	BGVCAM1513				

finally analysed was 1400 bp long (1054 coding bp and 346 noncoding bp: 252 bp at 3′UTR and 94 intronic bp) (Figure 1). Thirty-two SNPs were detected in *VvPel* (each 44 bp on average), five of which were non-synonymous or replacement polymorphisms (Table 4). The nucleotide diversity was $\pi = 0.0042$ and $\theta_w = 0.0039$ (θ_w standard deviation = 0.0011), where π is the average number of the nucleotide differences by site between two sequences (for the total of sequences is the average of all possible comparisons between each pair of sequences), and θ_w is the number of segregating sites in genetic models without recombination.

Fifteen haplotypes were inferred. Two were in a majority, showing frequencies of 0.24 (H1) and 0.21 (H10) in the collection analysed, and two were unique (H5 and H15), appearing only once in one accession (Table 5).

Prediction of the protein structure

Five SNPs showed polymorphisms that lead to non-synonymous amino acid substitutions, and for two of them, the software predicts changes in the physical properties of the protein. The SNP in the position 1027 (S1027) gives place to a change between glycine and alanine amino acids in position 359 of the protein (Gly359/Ala359), and S1122 produces a change between Gly391 and Asp391. Glycine is a polar amino acid, while alanine is hydrophobic, and asparagine has negative charge. The remaining amino acid replacements are predicted to

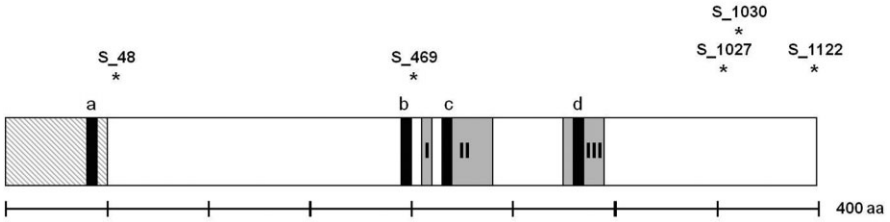

Figure 1. Scheme of the functional structure of *VvPel* protein (Domínguez-Puigjaner et al. 1997, Pua et al. 2001, Chourasia et al. 2006, Xiao et al. 2008): (a) glycosylation site; (b), (c) Ca++ binding sites; (d) catalytic site; grey bands are the conserved motifs I, II, and III in pectate lyase B; striped zone represents the non-analysed sequence; asterisks denote non-synonymous polymorphisms.

produce no physical changes in the protein but are localised near conserved motifs. The amino acid aa63 (Asp/Glu, affected by the SNP S48) is located near the possible glycosylation site, aa204 (Gly/Ser, by S469) is near the motif I, possibly involved in the secondary structure, and aa360 (Ala/Val, by S1030), as the above mentioned aa359, is located near to a conserved region in different species (Figure 1).

Table 4. Polymorphisms detected in the table grape core collection (CN96) and in each population (Q1 and Q2).

Diversity	CN96	Q1	Q2
Accessions	96	52	44
Total SNPs	32	32	21
SNPs in non-coding region	10	10	6
Synonymous SNPs	17	17	12
Nonsynonymous SNPs	5	5	3
Haplotypes	15	13	13
π	0.0042	0.0041	0.0042
θ_w	0.0039	0.0043	0.0030
Standard deviation θ_w	0.0011	0.0013	0.0010

SNP, single nucleotide polymorphism.

Regarding the secondary structure, the software predicts changes only in the protein resulting from haplotype H6. This protein would show a structure in α-helix between the amino acids aa334 and aa345 and between aa348 and aa359, while the resting haplotypes would give place to proteins with β-sheet structure between aa348 and aa350 and between aa353 and aa357, without α-helix between aa334 and aa345 (Figure 2). In haplotype H6, S1027 would give rise to alanine, while for the rest of haplotypes, it would give rise to glycine.

Selection and linkage disequilibrium

Selection was evaluated by means of the Fu and Li test. Values obtained for D* were 2.05 ($P < 0.02$) for the core collection, 1.98 ($P < 0.02$) for the population Q1, and 1.76 ($P < 0.02$) for the population Q2. Nevertheless, Tajima test results were not significant.

The parameter r^2 declined below 0.1 over 1200 bp taking into account the polymorphisms with a frequency above 0.1 (Figure 3). Certain polymorphisms were stronger linked , being the higher linkage ($1 \geq r^2 \geq 0.8$, $P < 0.0001$) between S18, S48, S441, S833, and S1030, between S582, S713, S727, and S1174, and between S469, S1122, S1274, S1295, and S1370 (Figure 4).

Association test

Significant associations at the 0.05 level for two or more years were obtained only between some SNPs and traits related to bunch size and berry texture. Table 6 shows the significant

Table 5. Haplotype frequencies observed in the table grape core collection (CN96) and in each population (Q1 and Q2) for the 15 haplotypes inferred from the sequence of *VvPel*.

H	Haplotype	Frequency at CN96	Frequency at Q1	Frequency at Q2
H1	CCCCTGGCCAAGCTCCCCGATGCGCGAATCGA	0.240	0.317	0.182
H2	CCCCTGGCCAAGCTCCCCGATGCGCAAACTGT	0.010	0.019	–
H3	CCCCTGGCCAAGCCCCCCGATGCGCGAATCGA	0.130	0.144	0.114
H4	CCCCTTGCCAAACTCCCCGATGCGCAAACTGT	0.063	0.058	0.068
H5	CCCCCGGCCAAGCTCCCAGATGCGCGGATCGA	0.005	–	0.011
H6	CCCCCGGCCAAGCTCCCAGATGCCCGAATCGA	0.208	0.183	0.239
H7	CCCCCTGACTAGCTCCCCGATGCGCGAATCAA	0.010	0.019	–
H8	CCCCCTGACTAGCTCCCCGATGCGCGGATCGA	0.094	0.048	0.148
H9	CCCCCTAACTAGCTCCCCGATGCGCGAATCGA	0.078	0.087	0.068
H10	CTCCTTGACTAGCTCCCCGATGCGCGAATCGA	0.021	–	0.011
H11	CTCTTGGACTAGTTCTTCGACACGCGAATCGA	0.094	0.038	0.159
H12	TCGCTGGCCTAGCTCTTCGACATGTAAATCGA	0.016	0.029	–
H13	TCGCTGGCCTTGCTTTTCAGTATGTAATTCGA	0.016	0.029	–
H14	TCGCTTGCTTTACTCTTCGACATGTAAATCGA	0.010	0.019	–
H15	TCGCTTGCTTTACTCTTCGACATGTAAACTGT	0.005	0.010	–

Haplotype H6

```
                  310       320       330       340       350
Structure                 βββββ ααααααααα     ααααααααααααα  ααα
Aa sequence       GGSADPTINSQGNRFLAPNDRFKKAVTKHEDAPESEWRHWNWRSEGDLML

                  360       370       380       390
Structure         αααααααααα                βββββββββ
Aa sequence       NGAFFLQSAAGASSSYARASSLSARPSSLVGSITLGSGALGCRKGSRC
```

Resting haplotypes

```
                  310       320       330       340       350
Structure                 βββββ ααααααααα                   βββ
Aa sequence       GGSADPTINSQGNRFLAPNDRFKKAVTKHEDAPESEWRHWNWRSEGDLML

                  360       370       380       390
Structure         βββββ                     βββββββββ
Aa sequence       NGAFFLQSGAGASSSYARASSLSARPSSLVGSITLGSGALGCRKGSRC
```

Figure 2. Protein secondary structures predicted for the haplotype 6 and for the resting haplotypes inferred for *VvPel*.

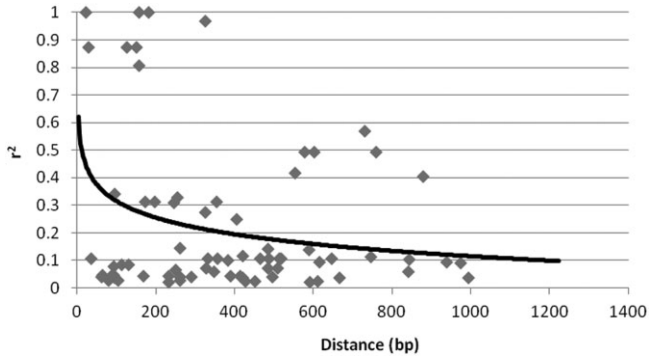

Figure 3. Linkage disequilibrium in the gene *VvPel*, measured with r^2 index (◆) from the inferred haplotypes using DnaSP software, and considering only the single nucleotide polymorphisms (SNPs) with frequency above 0.1. Logarithmic r^2 (—).

associations with non-synonymous polymorphisms and one interesting synonymous polymorphism (S405), while the remainder appear in Table 7. The replacement and tightly linked polymorphisms S48 and S1030 associated with bunch width explained 8% and the 11% of the total phenotypic variance for the data measured in 2004 and 2006, respectively. These SNPs also associated with bunch length for 3 years (2004, 2006 and 2007), explaining a variance proportion between 5 and 12%. The phenotypic distribution against each S48 allele was similar for the two traits, showing the heterozygous genotype with the highest value and the G:G genotype with the lowest, even in 2005, when data did not associate significantly (Figure 5). Polymorphisms linked to S48 and/or S1030 (S18, S582, S713, S727, S833, and S1174) also associated with these traits, with *P* values below 0.001 in many cases, and explaining a similar variance proportion (Table 7).

Polymorphisms S48 and S1030 were again associated with berry texture parameters, flesh firmness, and force at 10%, explaining 7% of the variance of the 2006 data (sensory and texture meter) and 10% of the variance of the 2004 data (sensory). The SNP S441, linked to them, associated with flesh firmness in 2005 too (Table 7). As it can be observed in Figure 6, the heterozygous genotype A:T showed a value of flesh firmness higher than that of the homozygous genotype A:A, but the minor allele frequency was low (0.03). S405 associated with force at 10% in 2006 and with force at 20% in 2005, so as with flesh firmness in 2005 and 2006. The lowest value for these parameters corresponded to genotype C:C (Figure 6), but the frequency of the genotype C:T was 0.03.

Polymorphism S469 also associated with force at 10%, explained 6% of the variance, while S1027 associated with rupture slope explained 8% of the variance. The distribution of the data of rupture slope against that of S1027 showed a pattern in which the genotype C:C presented the highest value against genotypes G:C and G:G (Figure 6). These polymorphisms presented a frequency of 0.08 (S469) and 0.21 (S1027) for the minority allele.

Finally, among the synonymous polymorphisms, S279 and S438 stand out because of the low *P*-values obtained in their associations with rupture area, which explained the 9% of the variance (Table 7). In this case, the heterozygous genotype (A:C and A:T, respectively) also showed a value of rupture area higher than that of the homozygous genotypes (Figure 6), and the frequency of the minority allele was high for both SNPs (around 0.3). In general, the variance explained by the marker was much lower (2–14%) than the variance explained by the model (12–59%).

Correlation within and between traits

Correlations between bunch architecture descriptors (bunch width, length, and mass) were high in 2005 and 2006, but in 2004 and 2007, bunch length showed a lower correlation value with the other two traits ($r = 0.5$).

A high correlation was also obtained between most of the texture parameters in the same year. The highest r values were obtained between force at 10%, force at 20%, slope at 10% and rupture slope, and between rupture force, rupture area, and rupture distance ($r > 0.7$). Deformation rate and rupture distance also showed a high value ($r = 0.76$). Further strong correlation was obtained between deformation rate and force at 10% but in the opposite direction ($r = -0.61$). The parameter force at 20% in 2005 correlated in high degree with force at 10% ($r = 0.64$), force at 20% ($r = 0.60$), and slope at 10% ($r = 0.66$) in 2006. The sensory descriptor for flesh firmness mid-level correlated with the force at 20% parameter measured in 2005 ($r = 0.5$) and to a lesser degree with the parameters measured in 2006: force at 10%, force at 20%, slope at 10%, and rupture slope (r around 0.4).

The most significant correlation between the 4-year data for the traits with significant associations was obtained between 2005 and 2006 ($r > 0.7$). Correlations between the other years

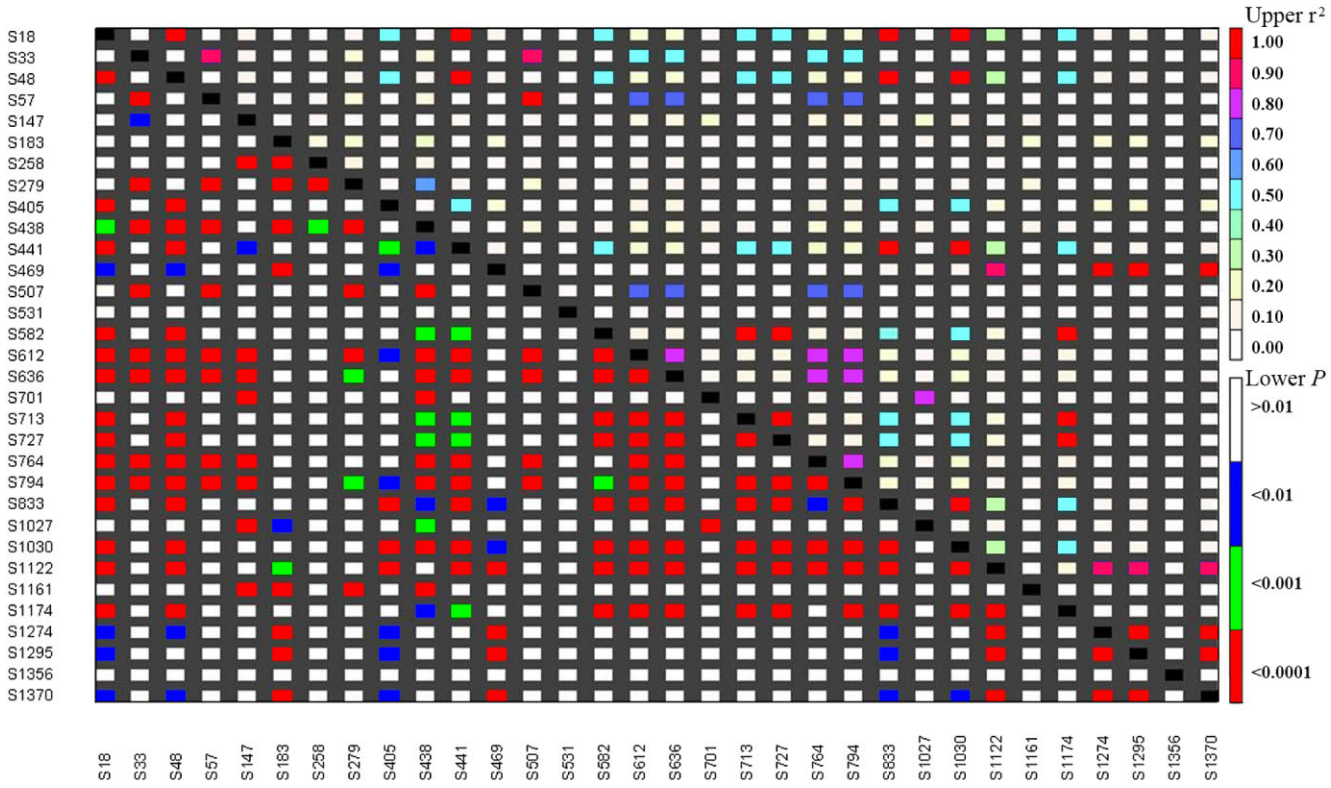

Figure 4. Linkage disequilibrium between individual SNPs in the gene *VvPel*, measured with r^2 index from genotypes using TASSEL software. The upper triangle shows the linkage (r^2), and the lower triangle shows the significance (*P*).

showed $r > 0.4$, except between bunch length in 2004 and 2006 ($r = 0.38$) and bunch width in 2004 and 2007 (not significant).

Discussion

Pectate lyase is an enzyme involved in cell wall disassembly, needed for cell enlargement and division. Its role in fruit softening during the ripening process has been demonstrated in several species, such as banana, strawberry, and mango (Domínguez-Puigjaner et al. 1997, Medina-Escobar et al. 1997, Pua et al. 2001, Jiménez-Bermúdez et al. 2002, Benítez-Burraco et al. 2003, Chourasia et al. 2006). During the past few years, diverse gene expression studies have provided insights into an understanding of pectate lyase function during this process. This work offers a new approach, attempting to identify, through an association study, where the variation in the gene sequence may have a consequence in the natural phenotypic variation in grape traits. Of the genes coding for pectate lyase in grapevine, only the one studied here was completely sequenced at the beginning this study (GenBank accession AF243475). It corresponds to VIT_14s0219g00230 in the present 12X version of the grapevine genome (http://genomes.cribi.unipd.it/). The selection was supported by the fact that the three PLs described in strawberry were all expressed in fruit during ripening. Later studies in the grapevine showed that the PL selected here was the one less expressed after veraison: according to Glissant et al. (2008), it expresses in berries from pea size (first stage studied) until veraison, showing the highest level of expression at veraison and the lowest just before it, while the other two studied PL transcripts are mostly abundant at veraison and after veraison (Nunan et al. 2001, Schlosser et al. 2008). According to the data published in PLEXdb (http://plexdb.org) by three different authors, the gene studied here expresses more or less uniformly during berry development (Deluc et al. 2007, Pilati et al. 2007, Lund et al. 2008).

A published phylogenetic analysis of diverse plant and bacterial PLs showed a cluster of PLs expressed in fruits and another cluster of microbial PLs. The PL studied here, with a *Zinnia elegans* (Fam. *Asteraceae*) PL, was between those two groups (Chourasia et al. 2006). *Zinnia elegans* PL expresses during vascular differentiation in phloem and xylem and in cells that constitute the recent product of meristematic divisions (Domingo et al. 1998).

Core collection

Spurious associations are frequent in association analyses as a result of data bias. In wide germplasm collections, the cost of certain molecular approaches can be high, so it is necessary to reduce the number of the samples selected for analysis, while retaining the maximum amount of variability. For this purpose, the building of representative core collections is useful. Nevertheless, for association studies, not only should the collection be representative of the existing diversity for the traits under study, it is important that every phenotypic class is represented several times to avoid spurious associations because of a low frequency of data. The core collection used in this study fulfilled that purpose, because it was built from the joining of three independent sub-collections (each one containing between 88 and 96% of the total variability), thus ensuring at least a representation in triplicate of each class for most descriptors and markers. Thus, any spurious associations because of unique phenotypic classes were avoided.

Table 6. Significant associations ($P < 0.05$) found between traits and SNP markers. It includes associations with non-synonymous markers and with S405. The rest are in Table 7. Only associations for 2 years or more are included, except in the case of texture-meter data, because measurements were taken only for 1 year. F: F-test calculated as MS marker/MS error; *P*: Probability of a larger F based on the F distribution; Rsq: Fraction of the total variance explained. In bold, *P* values below 0.01.

Trait	Descriptor	SNP	Year	Marker			Model
				F	P	Rsq	Rsq
Bunch architecture	Bunch width	S48	2004	3.24	0.0448	0.08	0.15
			2006	8.29	**0.0005**	0.10	0.50
		S1030	2004	3.24	0.0448	0.08	0.15
			2006	8.44	**0.0005**	0.11	0.48
	Ln Bunch length	S48	2004	7.25	**0.0014**	0.11	0.45
			2006	3.24	0.0441	0.05	0.31
			2007	10.03	**0.0026**	0.12	0.38
		S1030	2004	7.24	**0.0014**	0.11	0.45
			2006	3.25	0.0438	0.05	0.30
			2007	10.04	**0.0026**	0.12	0.38
Berry texture	Flesh firmness	S48	2004	3.79	0.0272	0.07	0.34
			2006	4.85	0.0102	0.10	0.15
		S405	2005	5.56	0.0205	0.04	0.42
			2006	8.27	**0.0051**	0.08	0.14
		S1030	2004	3.79	0.0272	0.07	0.34
			2006	4.85	0.0102	0.10	0.15
	Force 10%	S48	2006	3.36	0.0404	0.07	0.24
		S405	2006	6.77	0.0112	0.07	0.24
		S469	2006	5.26	0.0247	0.06	0.22
		S1030	2006	3.36	0.0404	0.07	0.24
	Force 20%	S405	2005	4.21	0.0432	0.02	0.59
	Rupture slope	S1027	2006	3.25	0.0446	0.08	0.14

MS, mean squares; SNP, single nucleotide polymorphism.

Genetic structure

Another main cause of spurious associations is when population structure is ignored during sample analysis. Therefore, population structure needs to be evaluated and included in the statistical model. The structure obtained here is consistent and adequate for the association analysis, since it appears to cluster the cultivars according to their geographic origin, and to selection and breeding practices. Besides, significant differences between the mean values for the traits involved in associations in the two populations support the inclusion of the structure in the association analysis. When results obtained with three models (general linear model (GLM) without structure, GLM with structure, and MLM with structure and kinship) are compared, it can be observed that MLM and GLM with structure gave similar results for texture, while MLM is more conservative for bunch dimensions (data not shown).

Diversity and molecular evolution

The gene *VvPel* showed a similar nucleotide diversity to that obtained by other authors for diverse genes studied in the grapevine, where one SNP each 47–64 bp was detected on average (Salmaso et al. 2004, Lijavetzky et al. 2007, This et al. 2007, Le Cunff et al. 2008, Emanuelli et al. 2010). Considering that the collection includes table grape cultivars exclusively, it is a gene with a high mutation rate. Nevertheless, the haplotype diversity was low in comparison with that obtained by Lijavetzky et al. (2007) and Salmaso et al. (2004) in the grapevine. The ratio haplotypes/SNPs was 0.7 according to Salmaso et al. (2004), and we obtained a ratio of 0.47. In other plants, the haplotype diversity detected is lower too: in maize, Ching et al. (2002) detected on average six haplotypes for 18 genes within 36 inbred lines and, in sunflower, Fusari et al. (2008) detected between one and nine haplotypes in 28 genes analysed in 19 accessions. Nevertheless, in grapevine, Fournier-Level et al. (2009) estimated between 24 and 42 haplotypes for three *Myb* genes and Emanuelli et al. (2010) detected 96 haplotypes for the gene *VvDXS*.

Traces of balanced selection were detected over the gene by means of the Fu and Li Test, in each population and in the whole collection. Nevertheless, it is difficult to know the underlying cause for such significant deviations from neutrality. The deviation could be because of the existence of selection but could also be a consequence of possible bottlenecks that happened during grapevine evolutionary history, for example during the phylloxera invasion. The deviation could also be caused by LD with other genes of the genome subjected to selection, what is frequent in cultured species.

Linkage Disequilibrium can also provide useful information for the detection of selection, since it increases with selective pressure, though it is also sensitive to factors like bottlenecks.

Table 7. Significant associations ($P < 0.05$) found between traits and synonymous SNP markers. Only those associations detected for at least two years were included, except in the case of texture-meter data, because measurements were taken only for 1 year. F: F-test calculated as MS marker/MS error; *P*: Probability of a larger F based on the F distribution; Rsq: Fraction of the total variance explained. In bold, *P* values below 0.01.

Trait	Descriptor	Locus	Year	Marker			Model
				F	P	Rsq	Rsq
Bunch architecture	Bunch width	S18	2004	3.24	0.0448	0.08	0.15
			2006	8.27	**0.0005**	0.10	0.50
		S582	2004	6.38	0.0137	0.07	0.15
			2006	13.47	**0.0004**	0.08	0.47
		S713	2004	6.38	0.0137	0.07	0.15
			2006	13.58	**0.0004**	0.09	0.47
		S727	2004	6.38	0.0137	0.07	0.15
			2006	13.63	**0.0004**	0.09	0.46
		S833	2004	3.24	0.0448	0.08	0.15
			2006	8.43	**0.0005**	0.11	0.48
		S1174	2004	6.38	0.0137	0.07	0.15
			2006	13.75	**0.0004**	0.09	0.46
	Ln bunch length	S18	2004	7.25	**0.0014**	0.11	0.45
			2006	3.24	0.0441	0.05	0.31
			2007	10.03	**0.0026**	0.12	0.38
		S582	2006	5.35	0.0232	0.04	0.30
			2004	12.59	**0.0007**	0.10	0.44
		S713	2006	5.36	0.0231	0.04	0.30
			2004	12.59	**0.0007**	0.10	0.44
		S727	2006	5.36	0.0231	0.04	0.30
			2004	12.59	**0.0007**	0.10	0.44
		S833	2004	7.24	**0.0014**	0.11	0.45
			2006	3.25	0.0438	0.05	0.31
			2007	10.04	**0.0026**	0.12	0.38
		S1174	2006	5.37	0.0229	0.05	0.29
			2004	12.59	**0.0007**	0.10	0.44
Berry texture	Deformation rate	S279	2006	3.76	0.0282	0.09	0.16
		S1161	2006	7.33	**0.0085**	0.09	0.16
	Flesh firmness	S18	2004	3.79	0.0273	0.07	0.34
			2006	4.85	0.0102	0.10	0.15
		S258	2004	4.11	0.0203	0.07	0.35
			2006	3.22	0.0449	0.07	0.12
		S441	2004	7.64	**0.0072**	0.07	0.35
			2005	4.97	0.0283	0.03	0.42
			2006	7.96	**0.0060**	0.08	0.14
		S833	2004	3.79	0.0272	0.07	0.34
			2006	4.85	0.0102	0.10	0.15
	Force 10%	S18	2006	3.36	0.0404	0.07	0.24
		S833	2006	3.36	0.0404	0.07	0.24
		S1274	2006	5.27	0.0247	0.06	0.22
		S1295	2006	5.27	0.0246	0.06	0.22
		S1370	2006	5.27	0.0246	0.06	0.22
	Force 20%	S531	2006	3.33	0.0439	0.11	0.20
		S701	2006	3.46	0.0391	0.11	0.21
	Rupture area	S279	2006	5.68	**0.0052**	0.09	0.42
		S438	2006	5.21	**0.0077**	0.09	0.42
		S612	2006	3.49	0.0359	0.06	0.39
		S636	2006	3.49	0.0358	0.06	0.39
		S794	2006	3.50	0.0357	0.06	0.39
		S1161	2006	5.56	0.0211	0.05	0.37
	Rupture force	S147	2006	3.17	0.0479	0.04	0.51
		S1161	2006	4.13	0.0458	0.03	0.48
			2007	4.53	0.0381	0.04	0.59
	Rupture slope	S701	2006	3.67	0.0305	0.09	0.15

MS, mean squares; SNP, single nucleotide polymorphism.

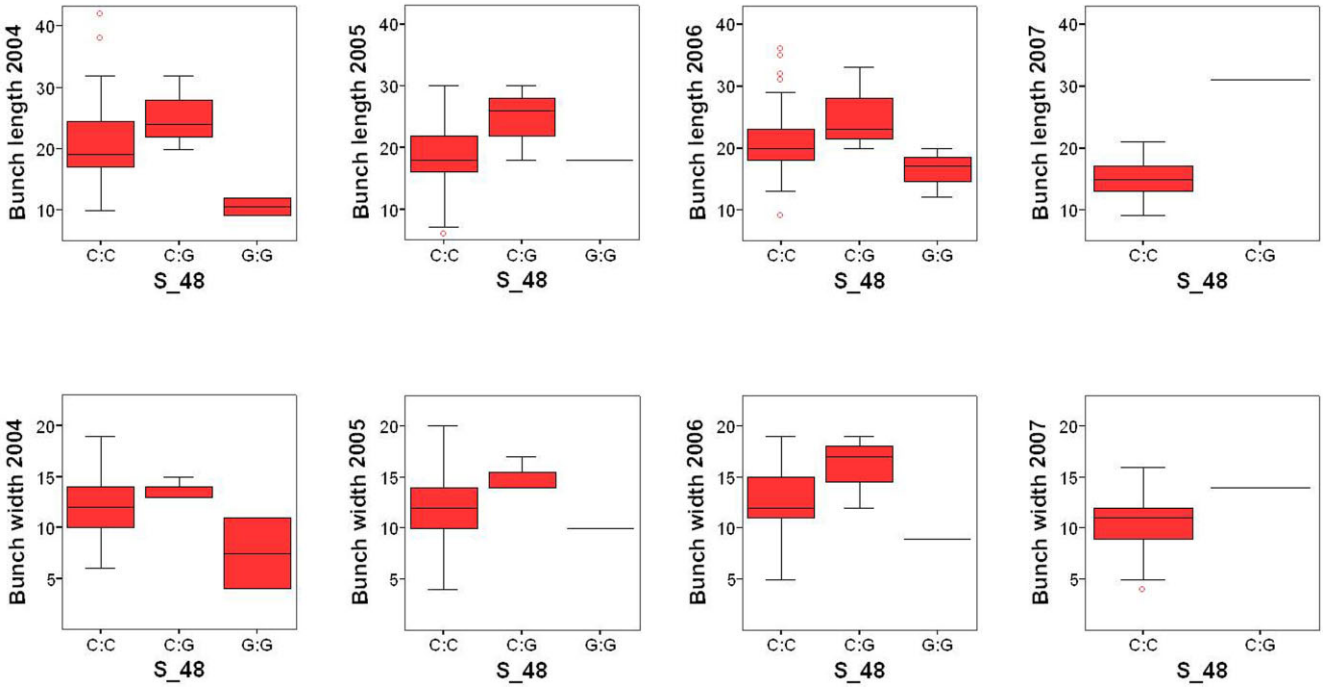

Figure 5. Box plots of the phenotypic distribution against SNP genotype for the main associations with bunch size traits in 2005 and 2006. Bunch length and bunch width measured in cm. A box plot shows a five-number summary: the smallest observation (sample minimum), lower quartile, median, upper quartile, and largest observation (sample maximum); it also indicates which observations, if any, might be considered outliers.

Along the gene, LD was low enough to provide a good resolution in the association analyses ($r^2 < 0.1$ over 1200 bp). A higher LD was obtained by Emanuelli et al. (2010), showing a value of $r^2 < 0.6$ over 4500 bp in *VvDXS*, a gene involved in Muscat flavour. They obtained significant positive values with the Fu and Li Test, pointing to a possible selective pressure over this gene. The low LD observed here could indicate absence of selection, but it falls in a similar way to other genes such as *VvmybA1*, a major gene responsible for berry colour in the grapevine, that may have undergone artificial selection, for which r^2 falls below 0.1 in around 700 bp (This et al. 2007). In *Dwarf8*, a supposed highly selected gene in maize, r^2 falls below 0.2 in about 1500 bp and below 0.1 in 2500 bp (Remington et al. 2001). So, the existence of selection in *VvPel* cannot be discarded after these data.

Association analysis

The results obtained in this work indicate that the *VvPel* could be involved in berry texture, as it was initially hypothesised, but also in bunch size. Data for those traits are mostly correlated over different years, showing the coherence of the observations. In both cases, numerous linked SNPs associated with those traits, most of them explaining around 10% of the phenotypic variance. Most of the associated SNPs with a high minor allele frequency are synonymous, and thus cannot cause directly the phenotypic changes but may be linked to other polymorphisms located in the regulatory region or in the non-analysed coding sequence which might actually be responsible for the variation.

S1027 is the only polymorphism predicted to produce conformational changes in the secondary structure of the PL protein. It is interesting because the minor allele has a high frequency (0.21), and it associated with the texture parameter rupture slope. This parameter was considered a flesh firmness index by Vargas et al. (2001), who found a high correlation ($r = 0.9$) between the firmness of the intact berry and the gradient or elasticity coefficient (rupture slope in this work). S1027 produces a replacement in the position aa359 of glycine (polar) by alanine (hydrophobic), and it could be responsible for the secondary structure change detected in the protein derived from the haplotype H6, which is different to the other. This region (from aa337 to aa355) is highly conserved in different species, so it could be involved in an important function of the enzyme that could be affected by a conformational change.

Another non-synonymous polymorphism, S469, is placed beside a Ca^{++} binding site and the conserved region corresponding to the motif I (Figure 1). A protein conformational change caused by this polymorphism could affect the pectate lyase function. This SNP associated with force at 10% in 2006, explaining 6% of the total phenotypic variance for this parameter.

Other interesting associations were obtained between the synonymous polymorphisms S279 and S438 and the texture parameter rupture area. It is calculated as the area under the force-deformation curve from the origin to the rupture distance, and measures the mechanical work done during a test. So, it has been related to the work necessary to break the berry, an indicator of toughness in the force-deformation curve (Sato et al. 1997). The polymorphisms explained a variance proportion of 9–14% for this parameter, showing a strong significance.

The SNP S405 associated with force at 20% in 2005, force at 10% in 2006, and flesh firmness in 2005 and 2006. These variables correlated at mid and high level. The fact that sensory and texture-meter measured data for berry texture associated for 2 years supports the results. The heterozygous genotype is present in the cultivars with firmer berries (Figure 6). The polymorphism S405, as well as S441, which associated for three years with flesh firmness, in spite of their low minor allele frequencies, could be worth testing as markers for flesh firmness in marker-assisted breeding in table grapes. The fact that SNP S441 associated with flesh firmness in 2004, 2005, and 2006

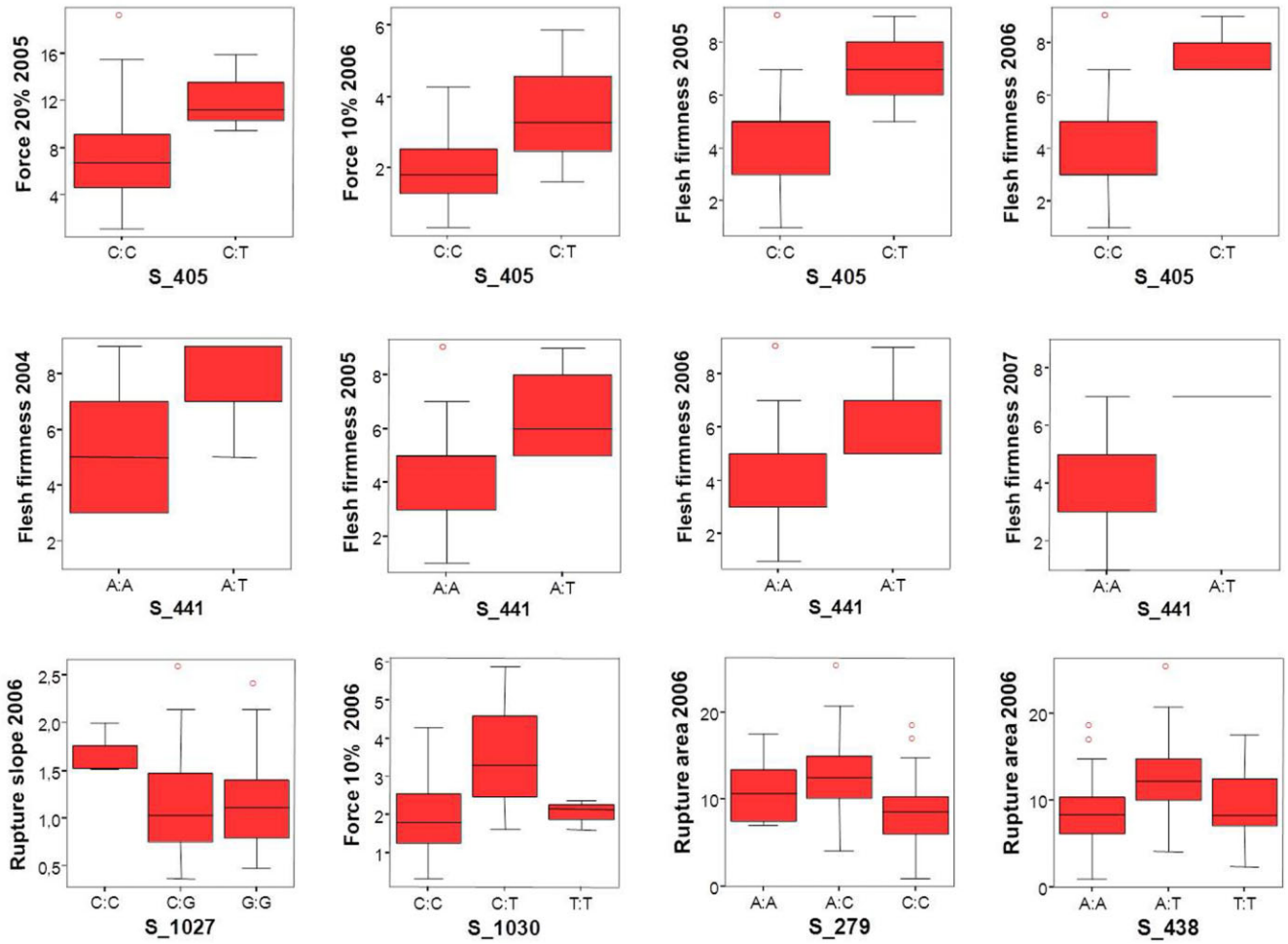

Figure 6. Box plots of the phenotypic distribution against single nucleotide polymorphism (SNP) genotype for the main associations with berry texture. Flesh firmness is classified from 1 (very soft) to 9 (very firm); forces measured in N; rupture area measured in N/mm; slope in N/mm. A box plot shows a five-number summary: the smallest observation (sample minimum), lower quartile, median, upper quartile, and largest observation (sample maximum); it also indicates which observations, if any, might be considered outliers.

and did not associate with 2007 data was because of the existence of missing data in the latter year.

Although associations with bunch size were not initially expected, two non-synonymous SNPs (S48 and S1030) associated at least in 2 years with bunch width and length explaining as much as 12% of the variation present. In the case of S48, the polymorphism produces an amino acid replacement which might affect an adjacent glycosylation site (Figure 1). Although, given the low frequencies obtained for the minor alleles (0.05), these associations need to be confirmed; they are especially interesting since that possible function had not been previously described for this gene. *VvPel* could be involved in an early stage of bunch development, affecting the growth of the inflorescence. Two processes are responsible for the development of the inflorescence: cell division and cell enlargement. Pectate lyase could be involved in the latter, by contributing to cell wall degradation. *VvPel* expression has not been studied in this stage in the grapevine, but this enzyme isoform is similar to *Z. elegans* (Asteraceae) pectate lyase, which shows an expression related to vascular differentiation in phloem and xylem and to cells that constitute recent product of meristematic divisions (Domingo et al. 1998).

The relatively low proportion of variance explained by the different SNPs could be because of the existence of three genes coding for this enzyme, collaborating moreover with others in the cell wall degradation. Taking into account the structure and the relationships between cultivars, the model explains up to 59% of the phenotypic variance, which is a high proportion, in the range of values published for QTLs in the grapevine (Marguerit et al. 2009).

Although the gene *VvPel* was not completely sequenced, the results obtained in this work could be of value for the present table grape breeding programs, given the importance of berry texture for the table grape market. Bunch architecture is also of interest, because it has major implications for quality and production, including disease control in the vineyard.

A recent transcriptomic study (Fasoli et al. 2012) showed a high level of expression of the gene studied in the rachis during fruit set and after fruit set, significantly higher than the expression found in seeds at mid-ripening stage (Seed-MR). Also it showed a higher level of expression, compared to Seed-MR, in young inflorescences, in berry pericarp at fruit set, and in berry flesh after fruit set. These data indicate that the *VvPel* studied here is expressed in organs and stages which may be relevant for the associations found.

Conclusions

The present work is a starting point for the study of the genetic basis of the natural variation in berry texture and inflorescence architecture in the grapevine. Berry texture is a complex

trait where many different enzymatic activities and regulation processes are involved. Several polymorphisms in the gene *VvPel*, which codes for pectate lyase, have associated significantly with parameters related to berry texture and bunch size. Among them, polymorphisms S1027 and S405/S441 in relation to berry texture, and S48/S1030 in relation to bunch size, appear to be the most promising ones.

Acknowledgements

This study was made possible with the funding from the GrapeGen project (joint venture between Genome Canada and Genoma España) and the AGL2010-15694 from the MICINN (Spain). A.M. Vargas was funded by a pre-doctoral fellowship from IMIDRA. We thank Carlos González Guillén, Nuria Rodríguez Jiménez, Concepción López Rivas, M. Dolores Vélez, Silvia Hernáiz, and Paz Fernández for their technical assistance in the morphological descriptions, and Loïc Le Cunff and Patrice This for their help with the association analysis, and Jérôme Grimplet and Pablo Carbonell for their help with gene expression data.

References

Benítez-Burraco, A., Blanco-Portales, R., Redondo-Nevado, J., Bellido, M.L., Moyano, E., Caballero, J.L. and Muñoz-Blanco, J. (2003) Cloning and characterization of two ripening-related strawberry (*Fragaria x ananassa* cv. Chandler) pectate lyase genes. Journal of Experimental Botany **54**, 633–645.

Bowers, J.E., Dangl, G.S., Vignani, R. and Meredith, C.P. (1996) Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). Genome **39**, 628–633.

Bowers, J.E., Dangl, G.S. and Meredith, C.P. (1999) Development and characterization of additional microsatellite DNA markers for grape. American Journal of Enology and Viticulture **50**, 243–246.

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics **23**, 2633–2635. doi: 10.1093/bioinformatics/btm308.

Brummell, D.A. and Harpster, M.H. (2001) Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. Plant Molecular Biology **47**, 311–340. doi: 10.1023/a:1010656104304.

Cabezas, J.A., Cervera, M.T., Ruiz-García, L., Carreño, J. and Martínez-Zapater, J.M. (2006) A genetic analysis of seed and berry weight in grapevine. Genome **49**, 1572–1585.

Ching, A., Caldwell, K.S., Jung, M., Dolan, M., Smith, O.S., Tingey, S., Morgante, M. and Rafalski, A.J. (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. Bmc Genetics **3**, 19. doi: 10.1186/1471-2156-3-19.

Chourasia, A., Sane, V.A. and Nath, P. (2006) Differential expression of pectate lyase during ethylene-induced postharvest softening of mango (*Mangifera indica* var. Dashehari). Physiologia Plantarum **128**, 546–555.

Costantini, L., Battilana, J., Lamaj, F., Fanizza, G. and Grando, M.S. (2008) Berry and phenology-related traits in grapevine (*Vitis vinifera* L.): from Quantitative Trait Loci to underlying genes. BMC Plant Biology **8**, 38. doi: 10.1186/1471-2229-8-38.

Deluc, L., Grimplet, J., Wheatley, M., Tillett, R., Quilici, D., Osborne, C., Schooley, D., Schlauch, K., Cushman, J. and Cramer, G. (2007) Transcriptomic and metabolite analyses of Cabernet Sauvignon grape berry development. Bmc Genomics **8**, 429. doi: 10.1186/1471-2164-8-429.

Domingo, C., Roberts, K., Stacey, N.J., Connerton, I., Ruiz-Teran, F. and McCann, M.C. (1998) A pectate lyase from *Zinnia elegans* is auxin inducible. Plant Journal **13**, 17–28.

Domínguez-Puigjaner, E., Llop, I., Vendrell, M. and Prat, S. (1997) A cDNA clone highly expressed in ripe banana fruits shows homology to pectate lyases. Plant Physiology **114**, 1071–1076.

Emanuelli, F., Battilana, J., Costantini, L., Le Cunff, L., Boursiquot, J.M., This, P. and Grando, M.S. (2010) A candidate gene association study on muscat flavor in grapevine (*Vitis vinifera* L.). BMC Plant Biology **10**, 241. doi: 10.1186/1471-2229-10-241.

Evanno, G., Regnaut, S. and Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology **14**, 2611–2620.

Falush, D., Stephens, M. and Pritchard, J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics **164**, 1567–1587.

Fasoli, M., Dal Santo, S., Zenoni, S., Tornielli, G.B., Farina, L., Zamboni, A., Porceddu, A., Venturini, L., Bicego, M., Murino, V., Ferrarini, A., Delledonne, M. and Pezzotti, M. (2012) The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. The Plant Cell **24**, 3489–3505. doi: 10.1105/tpc.112.100230.

Fournier-Level, A., Le Cunff, L., Gomez, C., Doligez, A., Ageorges, A., Roux, C., Bertrand, Y., Souquet, J.-M., Cheynier, V. and This, P. (2009) Quantitative genetic bases of anthocyanin variation in grape (*Vitis vinifera* L. ssp. sativa) berry: a quantitative trait locus to quantitative trait nucleotide integrated study. Genetics **183**, 1127–1139. doi: 10.1534/genetics.109.103929.

Fu, Y.X. and Li, W.H. (1993) Statistical tests of neutrality of mutations. Genetics **133**, 693–709.

Fusari, C.M., Lia, V.V., Hopp, H.E., Heinz, R.A. and Paniego, N.B. (2008) Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach. BMC Plant Biology **8**, 7. doi: 10.1186/1471-2229-8-7.

Glissant, D., Dedaldechamp, F. and Delrot, S. (2008) Transcriptomic analysis of grape berry softening during ripening. Journal International des Sciences de la Vigne et du Vin **42**, 1–13.

Gouesnard, B., Bataillon, T.M., Decoux, G., Rozale, C., Schoen, D.J. and David, J.L. (2001) MSTRAT: an algorithm for building germplasm core collections by maximizing allelic or phenotypic richness. Journal of Heredity **92**, 93–94.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alingmnet editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series **41**, 95–98.

Ibáñez, J., Vargas, A.M., Palancar, M., Borrego, J. and de Andrés, M.T. (2009) Genetic relationships among table-grape varieties. American Journal of Enology and Viticulture **60**, 35–42.

Ishimaru, M. and Kobayashi, S. (2002) Expression of a xyloglucan endotransglycosylase gene is closely related to grape berry softening. Plant Science **162**, 621–628.

Jiménez-Bermúdez, S., Redondo-Nevado, J., Muñoz-Blanco, J., Caballero, J.L., López-Aranda, J.M., Valpuesta, V., Pliego-Alfaro, F., Quesada, M.A. and Mercado, J.A. (2002) Manipulation of strawberry fruit softening by antisense expression of a pectate lyase gene. Plant Physiology **128**, 751–759.

Le Cunff, L., Fournier-Level, A., Laucou, V., Vezzulli, S., Lacombe, T., Adam-Blondon, A.-F., Boursiquot, J.-M. and This, P. (2008) Construction of nested genetic core collections to optimize the exploitation of natural diversity in *Vitis vinifera* L. subsp. sativa. BMC Plant Biology **8**, 31.

Lijavetzky, D., Cabezas, J.A., Ibáñez, A., Rodríguez, V. and Martínez-Zapater, J.M. (2007) High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. Bmc Genomics **8**, 424.

Lund, S., Peng, F., Nayar, T., Reid, K. and Schlosser, J. (2008) Gene expression analyses in individual grape (*Vitis vinifera* L.) berries during ripening initiation reveal that pigmentation intensity is a valid indicator of developmental staging within the cluster. Plant Molecular Biology **68**, 301–315. doi: 10.1007/s11103-008-9371-z.

Marguerit, E., Boury, C., Manicki, A., Donnart, M., Butterlin, G., Nemorin, A., Wiedemann-Merdinoglu, S., Merdinoglu, D., Ollat, N. and Decroocq, S. (2009) Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. Theoretical and Applied Genetics **118**, 1261–1278. doi: 10.1007/s00122-009-0979-4.

Maury, C., Madieta, E., Le Moigne, M., Mehinagic, E., Siret, R. and Jourjon, F. (2009) Development of a mechanical texture test to evaluate the ripening process of Cabernet Franc grapes. Journal of Texture Studies **40**, 511–535. doi: 10.1111/j.1745-4603.2009.00195.x.

Medina-Escobar, N., Cárdenas, J., Moyano, E., Caballero, J. and Muñoz-Blanco, J. (1997) Cloning, molecular characterization and expression pattern of a strawberry ripening-specific cDNA with sequence homology to pectate lyase from higher plants. Plant Molecular Biology **34**, 867–877.

Merdinoglu, D., Butterlin, G., Bevilacqua, L., Chiquet, V., Adam-Blondon, A.F. and Decroocq, S. (2005) Development and characterization of a large set of microsatellite markers in grapevine (*Vitis vinifera* L.) suitable for multiplex PCR. Molecular Breeding **15**, 349–366. doi: 10.1007/s11032-004-7651-0.

Nei, M. (1987) Molecular evolutionary genetics (Columbia University Press: New York).

Nunan, K.J., Davies, C., Robinson, S.P. and Fincher, G.B. (2001) Expression patterns of cell wall-modifying enzymes during grape berry development. Planta **214**, 257–264.

Office International de la Vigne et du Vin (2007) 2nd edition of the OIV descriptor list for grape varieties and Vitis species. http://www.oiv.int/oiv/info/enplubicationoiv#grape

Pilati, S., Perazzolli, M., Malossini, A., Cestaro, A., Dematte, L., Fontana, P., Dal Ri, A., Viola, R., Velasco, R. and Moser, C. (2007) Genome-wide transcriptional analysis of grapevine berry ripening reveals a set of genes similarly modulated during three seasons and the occurrence of an oxidative burst at veraison. BMC Genomics **8**, 428. doi: 10.1186/1471-2164-8-428.

Pritchard, J.K., Stephens, M. and Donnely, P. (2000) Inference of population structure using multilocus genotype data. Genetics **155**, 945–959.

Pua, E.-C., Ong, C.-K., Liu, P. and Liu, J.-Z. (2001) Isolation and expression of two pectate lyase genes during fruit ripening of banana (*Musa acuminata*). Physiologia Plantarum **113**, 92–99.

Remington, D.L., Thornsberry, J.M., Matsuoka, Y., Wilson, L.M., Whitt, S.R., Doeblay, J., Kresovich, S., Goodman, M.M. and Buckler, E.S. (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proceedings of the National Academy of Sciences of the United States of America **98**, 11479–11484.

Rolle, L., Torchio, F., Zeppa, G. and Gerbi, V. (2009) Relationship between skin break force and anthocyanin extractability at different ripening stages. American Journal of Enology and Viticulture **60**, 93–97.

Rolle, L., Giacosa, S., Gerbi, V. and Novello, V. (2011) Comparative study of texture properties, color characteristics and chemical composition of ten white table grape varieties. American Journal of Enology and Viticulture **62**, 49–56. doi: 10.5344/ajev.2010.10029.

Rolle, L., Siret, R., Rio Segade, S., Maury, C., Gerbi, V. and Jourjon, F. (2012) Instrumental texture analysis parameters as markers of table-grape and winegrape quality: a review. American Journal of Enology and Viticulture **63**, 11–28. doi: 10.5344/ajev.2011.11059.

Rozas, J. and Rozas, R. (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15**, 174–175.

Rozen, S. and Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. Krawetz, S. and Misener, S., eds. Methods in Molecular Biology (Humana Press: Totowa, NJ, USA) pp. 365–386.

Salmaso, M., Faes, G., Segala, C., Stefanini, M., Salakhutdinov, L., Zyprian, E., Toepfer, R., Grando, M.S. and Velasco, R. (2004) Genome diversity and gene haplotypes in the grapevine (*Vitis vinifera* L.), as revealed by single nucleotide polymorphisms. Molecular Breeding **14**, 385–395.

Sato, A. and Yamada, M. (2003) Berry texture of table, wine, and dual-purpose grape cultivars quantified. HortSscience **38**, 578–581.

Sato, A., Yamane, H., Hirakawa, N., Otobe, K. and Yamada, M. (1997) Varietal differences in the texture of grape berries measured by penetration tests. Vitis **36**, 7–10.

Schlosser, J., Olsson, N., Weis, M., Reid, K., Peng, F., Lund, S. and Bowen, P. (2008) Cellular expansion and gene expression in the developing grape (*Vitis vinifera* L.). Protoplasma **232**, 255–265. doi: 10.1007/s00709-008-0280-9.

Sefc, K.M., Regner, F., Turetschek, E., Glossl, J. and Steinkellner, H. (1999) Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. Genome **42**, 367–373.

Stephens, M., Smith, N.J. and Donnelly, P. (2001) A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics **68**, 978–989.

Terrier, N., Glissant, D., Grimplet, J., Barrieu, F., Abbal, P., Couture, C., Ageorges, A., Atanassova, R., Leon, C., Renaudin, J., Dedaldechamp, F., Romieu, C., Delrot, S. and Hamdi, S. (2005) Isogene specific oligo arrays reveal multifaceted changes in gene expression during grape berry (*Vitis vinifera* L.) development. Planta **222**, 832–847. doi: 10.1007/s00425-005-0017-y.

This, P., Lacombe, T., Cadle-Davidson, M. and Owens, C.L. (2007) Wine grape (*Vitis vinifera* L.) color associates with allelic variation in the domestication gene VvmybA1. Theoretical and Applied Genetics **114**, 723–730. doi: 10.1007/s00122-006-0472-2.

Thomas, M.R. and Scott, N.S. (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). Theoretical and Applied Genetics **86**, 985–990.

Vargas, A., Pérez, J., Zoffoli, J.P. and Pérez, A. (2001) Comparacion de variables de textura en la medicion de firmeza de bayas de uva Thompson seedless. Ciencia e Investigacion Agraria **28**, 37–42.

Vargas, A.M., de Andrés, M.T., Borrego, J. and Ibáñez, J. (2009) Pedigrees of fifty table grape cultivars. American Journal of Enology and Viticulture **60**, 525–532.

Watterson, G.A. (1975) Number of segregating sites in genetic models without recombination. Theoretical Population Biology **7**, 256–276.

Whitt, S.R. and Buckler, E.S. (2003) Using natural allelic diversity to evaluate gene function. Methods in Molecular Biology **236**, 123–140.

Xiao, Z.H., Bergeron, H., Grosse, S., Beauchemin, M., Garron, M.L., Shaya, D., Sulea, T., Cygler, M. and Lau, P.C.K. (2008) Improvement of the thermostability and activity of a pectate lyase by single amino acid substitutions, using a strategy based on meltingtemperature-guided sequence alignment. Applied and Environmental Microbiology **74**, 1183–1189. doi: 10.1128/aem.02220-07.

Zouid, I., Siret, R., Mehinagic, E., Maury, C., Chevalier, M. and Jourjon, F. (2010) Evolution of grape berries during ripening: investigations into the links between their mechanical properties and the extractability of their skin anthocyanins. Journal International des Sciences de la Vigne et du Vin **44**, 87–99.

Manuscript received: 28 May 2012
Revised manuscript received: 21 December 2012
Accepted: 21 February 2013