



**HAL**  
open science

## VvGAI1 polymorphisms associate with variation for berry traits in grapevine

Alba M. Vargas, Loic Le Cunff, Patrice This, Javier Ibanez, M.Teresa de Andrés

► **To cite this version:**

Alba M. Vargas, Loic Le Cunff, Patrice This, Javier Ibanez, M.Teresa de Andrés. VvGAI1 polymorphisms associate with variation for berry traits in grapevine. *Euphytica*, 2013, 191, pp.85-98. 10.1007/s10681-013-0866-6 . hal-01268012

**HAL Id: hal-01268012**

**<https://hal.science/hal-01268012v1>**

Submitted on 28 May 2020

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

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

Copyright

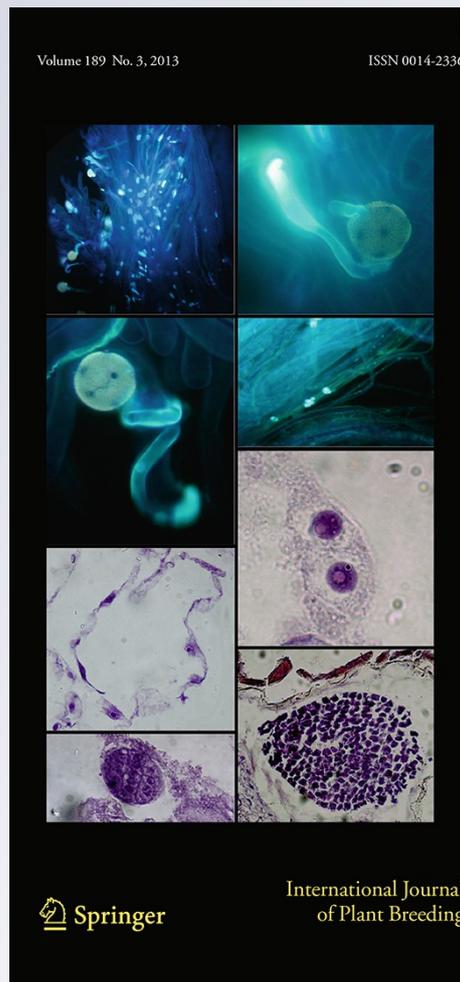
# *VvGAI1 polymorphisms associate with variation for berry traits in grapevine*

**Alba M. Vargas, Loïc Le Cunff, Patrice This, Javier Ibáñez & M. Teresa de Andrés**

**Euphytica**  
International Journal of Plant Breeding

ISSN 0014-2336

Euphytica  
DOI 10.1007/s10681-013-0866-6



**ONLINE  
FIRST**

**Euphytica**

 Springer

**Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.**

# VvGAI1 polymorphisms associate with variation for berry traits in grapevine

Alba M. Vargas · Loïc Le Cunff · Patrice This ·  
Javier Ibáñez · M. Teresa de Andrés

Received: 6 July 2012 / Accepted: 12 January 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** Gibberellins have important effects over different developmental processes in plants and the exogenous application of GA<sub>3</sub> has been extensively used in grapevine to improve certain quality traits (to control berry size, set and weight, to decrease cluster compactness and to induce the absence of seeds in grapes). In this study, we performed association analyses using a core collection of 127 table grape accessions in order to identify causal SNPs in the gene *VvGAI1*, a negative regulator of gibberellins

response. Seventeen quantitative descriptors, related with fertility, bunch and berry traits were used for the association analysis with *VvGAI1*. Fifteen polymorphisms and 15 haplotypes were identified, and some of them associated significantly with important traits for table grape: berry texture, juice yield, and bunch weight. The genetic structure found in the collection of varieties was consistent with their origins, and was included in the association test. In addition, nucleotide and haplotypic diversity, linkage disequilibrium and the possible existence of selection over the gene were also evaluated to support the association test.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10681-013-0866-6) contains supplementary material, which is available to authorized users.

**Keywords** Association analysis · Table grape · *Vitis vinifera* · *VvGAI1*

A. M. Vargas (✉) · J. Ibáñez · M. T. de Andrés  
Instituto Madrileño de Investigación y Desarrollo Rural,  
Agrario y Alimentario (IMIDRA), Finca El Encín,  
Alcalá de Henares, Spain  
e-mail: alba.vargas@madrid.org

## Introduction

L. Le Cunff · P. This  
INRA UMR 1097 Diversité et Adaptation des Plantes  
Cultivées, 2 Place Viala, 34060 Montpellier, France

The understanding of the genetic basis of important traits is one of the main targets for many crops. Identification of the implicated genes and the understanding of their interactions with other genetic or environmental factors are also important. Numerous studies have reported on genomic regions (QTL, quantitative trait loci) in which there could be a gene (or genes) involved in the expression of a trait. These studies do not allow the identification of the gene involved in the phenotypic variation studied and even less the causal polymorphism. However, they can be the starting point to reduce the pool of candidate genes

### Present Address:

L. Le Cunff  
UMT Geno-Vigne® (IFV, Supagro, INRA), 2 Place  
Viala, 34060 Montpellier, France

### Present Address:

J. Ibáñez  
CSIC-Universidad de La Rioja-Gobierno de La Rioja,  
C/Madre de Dios 51, Bajo, 26006 Logroño, La Rioja,  
Spain

Published online: 26 January 2013

 Springer

and to make easier the selection of one of them. A different approach to achieve those goals is through association studies, which involve the detection of correlation between phenotypes and molecular polymorphisms, haplotypes of genes or at the whole genome level (Gupta et al. 2005). Grapevine segregating populations have a greater diversity for identification of phenotype-associated markers due to heterozygous parental cultivars. However, their genetic background remains relatively narrow compared to diversity panels (Huang et al. 2012).

Early on, association mapping (AM) had been used in human studies to understand the genetic control of diseases. Nowadays, it has rapidly gained interest among plant scientists for studying biomass traits, yield, and disease resistance, among others. Reviews on the concept, methodology, prospects, and status of AM studies in plants have already been published (Abdurakhmonov and Abdugarimov 2008; Breseghello and Sorrells 2006; Ersoz et al. 2007; Gupta et al. 2005; Khan and Korban 2012; Myles et al. 2009; Neale and Savolainen 2004; Sorkheh et al. 2008; Zhu et al. 2008).

The use of extant natural diversity provides advantages in resolution and breadth of survey, but can also present added difficulties in accurately assessing the true cause of an association. The most serious false positives can result when unlinked markers produce a positive association because of underlying population structure. Phenotype and polymorphism frequencies can be distinct in different populations, so if one phenotype presents a high frequency in one population, any polymorphism with a high frequency in that population will be associated with that phenotype, even if they are not related (Pritchard et al. 2000b). On the other side, the resolution power of an association analysis depends on the linkage disequilibrium (LD) in the region around the candidate gene: the lower LD, the better resolution (Remington et al. 2001). Moreover, knowing the evolutionary history of the gene is interesting because if it has been subjected to selection during the domestication or the improvement process, it is probably related with an agronomical important trait (Remington et al. 2001).

Grapevine is an ancient crop with a long history of human selection. Although wine grape is economically more important, table grape breeding is currently becoming more and more significant, and every year the number of new table grape varieties increases. In last decades, the Muscat flavor and the absence of

seeds have been some of the main interesting traits in table grape. However, nowadays there is not much information about the genetic control of other quality or production related traits, such as those related with bunch architecture, berry texture and size, or fertility. And although several QTLs have been identified for some of them (Battilana et al. 2009; Cabezas et al. 2006; Costantini et al. 2008; Doligez et al. 2002; Doligez et al. 2010; Fanizza et al. 2005; Mejía et al. 2007), there are only reports of the active utilization of their results in table grape breeding programs for seedlessness (Karaagac et al. 2012; Mejía et al. 2011). Gibberellins (GAs) are diterpenoid compounds which regulate plants growth and control different developmental processes, including germination, leaf expansion, stem elongation or flower and fruit development (Olszewski et al. 2002). Gibberellic acid (GA<sub>3</sub>) has been used extensively to control berry size, set and weight, to decrease cluster compactness and to induce the absence of seeds (seedlessness) in grapes (Mullins et al. 1992; Winkler et al. 1974). A slight increase in flesh firmness has also been detected after GA<sub>3</sub> application (Sato et al. 2004; Singh et al. 1978). Therefore, the genes involved in the regulation of the GAs content can be considered as candidate genes for the regulation of these phenotypic traits. One of them is *VvGAI1* (DELLA subfamily, GRAS family). A mutant for this gene was observed in Pinot Meunier, in which a polymorphism in the DELLA domain gave place to a dwarf phenotype with a high fertility and loss of juvenility, consequence of the development of certain primordia into inflorescences instead of tendrils (Boss and Thomas 2002). Although in that work the authors did not detect expression in immature or ripening berries, several authors observed later that *VvGAI1* is expressed, even at high levels, in this tissue during berry development, declining transiently at veraison, and with higher levels of expression in seeds (Deluc et al. 2007; Grimplet et al. 2007).

In grapevine, association studies have been carried out only with genes with a previously known relationship with the studied trait [the color of the grape skin (Fournier-Level et al. 2011; Lijavetzky et al. 2006; This et al. 2007), fleshless berry locus (Houel et al. 2010) or co-locating with the Muscat flavor QTL (Emanuelli et al. 2010), and grape berry proanthocyanidin composition (Huang et al. 2012)]. The goal of this work was to evaluate the existence of associations between the natural phenotypic variation found for

different important quality traits in table grape and sequence polymorphisms at a gene (*VvGAI1*), which could have a pleiotropic effect and be involved in their genetic regulation, since it is a negative regulator of GAs response. In this way, this work with the gene *VvGAI1* may be considered a case study for performing an association study using a standard methodology (Hirschhorn et al. 2002). Firstly, we selected a collection of varieties of *Vitis vinifera* representing a high variability for these traits and we studied its genetic structure. Secondly, the traits were scored and the candidate gene was sequenced in the collection. Once the sequence polymorphisms were identified, haplotypes were inferred to do an in-depth diversity and LD analysis. Finally, it was evaluated if the significant correlations obtained could be related with possible structural and functional changes in the protein caused by the polymorphisms identified.

## Materials and methods

### Plant material and genetic and phenotypic data

The plant material derived from an initial collection of 352 table grape accessions maintained at the germplasm bank of El Encín (IMIDRA), which was characterized for 48 qualitative and quantitative descriptors during the 2005 and 2006 seasons. Twenty nuclear microsatellites were studied in two independent multiplex PCRs (Ibáñez et al. 2009).

For phenotypic characterizations five bunches harvested with a ripening index [ $^{\circ}$ Brix/tartaric acid (g/l)] of 2.5 were described for each accession. Berry weight and volume were evaluated using 100 berries and berry average dimensions were measured from 15 random berries. Juice yield was estimated using the volume of liquid obtained from squeezing the same 100 berries in a blender. Seventeen quantitative descriptors, related with fertility, bunch and berry traits were used for the association analysis (supplementary Table 1).

Berry texture was defined in 2006 using a TA-XTPlus texture-meter (Stable Micro Systems) with a 2 mm diameter punch probe. 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 into the sample flesh. A force against distance profile is obtained. The measured properties were *force at 10* and *20* % (force in Newtons made by the texture-meter when the berry has been deformed 10 and 20 % of its diameter, respectively), *rupture force* (force made by the texture-meter in the skin rupture point), *rupture distance* (distance from the origin point to the rupture point), *slope 10* % and *rupture slope* (slope force–distance at 10 % of berry diameter deformation and at the rupture point, respectively), *turgidity* (rupture slope–slope 10 %), *rupture energy* (area under the curve between the origin point and the rupture distance) and *deformation rate* (rupture distance/berry width).

Correlation analyses between traits and between years were done with SPSS15 (SPSS Inc. 2006), using Pearson correlation coefficient.

### Core collection: population structure and kinship matrix

The association tests involved a core collection of 127 table grape accessions (supplementary Table 2) selected through the Maximization Strategy implemented in the software MStrat 4.1 (Gouesnard et al. 2001), which contained 100 % of the total phenotypic and genotypic variability detected in the original 352 table grape accessions.

Structure analysis for the core collection was carried out by using the software Structure 2.0 (Pritchard et al. 2000a). Eleven non-linked nuclear microsatellites were used (VMC1B11, VMC4F3-1, VVIH54, VVIN16, VVIV37, VVMD7, VVMD21, VVMD24, VVMD28, VVMD32 and VVS2). A model, with 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 1,500,000, with 500,000 burn-in steps, and 10 iterations for each number of putative populations were used. Evanno criterion was used to decide the populations number (Evanno et al. 2005).

A kinship matrix was constructed using Tassel software (Bradbury et al. 2007) and 12 additional and non-linked microsatellite markers (VVMD5, VVMD25, VVMD27, VVIB01, VVIN73, VVIP31, VVIQ52, VVIV67, *ssrVrZAG62*, *ssrVrZAG67*, *ssrVrZAG83*, and *ssrVrZAG112* (Sefc et al. 1999), amplified according to Ibáñez et al. (2009)).

Finally, for the evaluation of the detected structure five chloroplast microsatellites were also analyzed (Ccmp3, Ccmp5, Ccmp10 (Weising and Gardner 1999), ccSSR9, and ccSSR14 (Chung and Staub 2003)) using a multiplex PCR (Ibáñez et al. 2009). The detected chlorotypes were named according to Arroyo-García et al. (2006).

### Gene amplification and sequencing

Primers for the amplification of the gene *VvGAI1* (GenBank library accession number AF378125.1) were designed using Primer3 v0.2 (Rozen and Skaltsky 2000). Four fragments (a–d) were sequenced and the primers for each one were GAI-a Fw (5'-CC ACCCTCTGCAACTCCTC-3'), GAI-a Rv (5'-TGCG AGTCAACAAGGACAAC-3'), GAI-b Fw (5'-CAA TTCCAGGTAAGGCTCT-3'), GAI-b Rv (5'-GCCA CGAAACCTCTGTATT-3'), GAI-c Fw (5'- ACTT CACCGCCAATCAAGC-3'), GAI-c Rv (5'-CGAAT AGAGCCAGCAGCATA-3'), GAI-d Fw (5'-AGTG GAGCGGCACGAAAC-3') and GAI-d Rv (5'-TCAC AAGGC(AT)CGGTCATACA-3'). PCR amplifications were performed using 1.05 U of a Hot Start Taq polymerase (Hotmaster, Eppendorf), 0.2 mM dNTPs, buffer 1× (stock 10×: 25 mM TrisHCl pH 8, 35 mM KCl, 0.1 mM EDTA, 1 mM DTT, 50 % glycerol, 0.5 % Tween 20, 0.5 % Igepal CA-630 and stabilizers) with 2.5 mM MgCl<sub>2</sub>, 0.2 μM primers and 5 ng of ADN, in a final volume of 21 μl. The program of PCR consisted of an initial denaturing step of 94 °C for 2 min followed by 35 cycles of (94 °C for 1 min, 60 °C for 1 min and 65 °C for 2 min), and a final extension step of 65 °C for 30 min. The PCR products were purified using ExoSap-IT (usb) and the Agen-court AMPure method (Beckman Coulter) and directly sequenced in the two ways at the Genomic Unit of the Parque Científico de Madrid (Spain) and at INRA (Montpellier, France). Sequencing analyses and alignments were performed with SeqScape 2.5 (Applied Biosystem) and BioEdit 7.0.0 (Hall 1999).

### Statistical tests and association study

Haplotypes were inferred with the software Phase 2.1 (Stephens et al. 2001). The diversity parameters  $\pi$  (Nei 1987) and  $\theta$  (Watterson 1975), the Neutrality Test of Fu and Li (1993), the recombination rate (Hudson 1987) and the prediction of possible recombination

events (Hudson and Kaplan 1985) were estimated using DnaSP 4 (Rozas and Rozas 1999). LD analyses were carried out with DnaSP 4 and Tassel 2.1 (Bradbury et al. 2007). Association test consisted of Mixed Linear Model (MLM), implemented in Tassel, and combined phenotypic data from the two years by calculating best linear unbiased predictors (BLUPs).

Mixed Linear Model allows to incorporate in the model random effects, such as information about relationships among individuals by means a kinship matrix and the population structure. Two association analyses were done, one with individual SNPs and the other with the haplotypes inferred with Phase. Only haplotypes with frequencies higher than 1 % were used for the association tests.

## Results

Different analyses were carried out to complement the association test: in the population, the existence of a structure; and in the gene *VvGAI1*, the nucleotide and haplotype diversity, the protein structure, and the evaluation of the selection existence and LD.

### Population structure

The results obtained using Structure 2.0, suggest the existence of three populations in the core collection (Q1, Q2 and Q3) (supplementary Table 2). The population Q1 is integrated by 40 accessions, with a high representation of Sultanina related cultivars, 55 % of which have an Oriental origin. The population Q2 is formed by 46 accessions, 46 % of which are ancient cultivars from the Iberian Peninsula and 41 % are of mixed use (table and wine). The population Q3, includes mostly recent bred varieties, obtained from crosses involving Muscat cultivars as progenitors (41 %), and is represented by 41 accessions (66 % of Italian and 20 % of French origin (Ibáñez et al. 2009; Vargas et al. 2009)). Five polymorphic chloroplast microsatellite loci were also studied in order to test the consistency of structure detected. The four different chlorotypes identified in the core collection were found in all three populations detected (Q1, Q2 and Q3) although, with different frequencies (Table 1). The most frequent chlorotype in the core collection was chlorotype C, which was present in nearly 50 % of varieties included in population Q3. Chlorotypes D and

**Table 1** Chloroplast haplotypes found in the core collection (C127) and frequency distribution of the four chlorotypes in the populations obtained (Q1, Q2 and Q3)

Chlorotype	C127	Q1 (%)	Q2 (%)	Q3 (%)
A	27	15.38	<b>37.78</b>	9.30
B	14	15.38	4.44	13.95
C	<b>50</b>	33.33	33.33	<b>51.16</b>
D	36	<b>35.90</b>	24.44	25.58

Chlorotypes named according to Arroyo-García et al. (2006)

The bold values indicate frequency of the majority haplotype in each population

A showed a slightly majority in populations Q1 and Q2 respectively. It is also interesting emphasize that the 63 % of varieties carrying chlorotype A grouped in Q2.

### Nucleotide and haplotype diversity and protein structure prediction

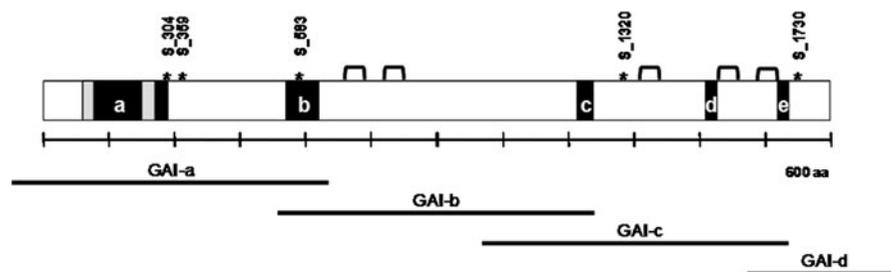
The complete coding region of the gene *VvGAI1* was sequenced (1,841 bp; 1,770 coding bp and 71 non coding bp at 3'UTR) (Fig. 1). Fifteen SNPs were detected in *VvGAI1* (one each 123 bp on average), five of which produced an amino acid change in the protein (Table 2). The nucleotide diversity was  $\pi = 0.0016$  and  $\theta = 0.0013$  ( $\theta$  standard deviation = 0.0004). SNP allele frequency found in *VvGAI1*, mayor and minor alleles detected by position are show in Table 3. Fifteen haplotypes were inferred (supplementary Table 3). Two of them were majorities, showing frequencies of 0.30 and 0.34 in the core collection analyzed, and six were unique, appearing only once in one accession. Seven haplotypes, including the six unique, were discarded for the association analysis, having frequencies lower than 0.01.

Of five SNPs showing amino acidic substitutions, two could lead to changes in the physical properties of the protein. SNP in the position 304 (S304) gives place to a change between proline and serine amino acids in position 102 of the protein (Pro102/Ser102), and S583 produces a change between Pro195 and Ala195. Proline is a hydrophobic polar amino acid, while serine is polar and alanine is hydrophobic. S359 gives place to a change between two polar amino acids (Thr120/Asn120), and, as Pro102/Ser102, the polymorphism is located near the DELLA region (Fig. 1). The amino acid Pro195/Ala 195 is located near a conserved motive (VVLVD).

The S1730 does not produce a change in the protein physical properties but the corresponding amino acid (Ala577/Val577) is localized near the SAW motif (Fig. 1), which is a conserved motif characteristic of the GRAS proteins. The S1320 produces an amino acidic change in the site Ile440/Met440, localized in a partially conserved region.

### Selection and linkage disequilibrium

The results of Neutrality tests were not significant for the core collection. Nevertheless, taking into account population structure and analyzing 100 bp windows, negative values with  $p < 0.1$  were obtained with Fu and Li tests for positions in which there were amino acid polymorphisms. The  $D^*$  test statistic evaluates the deviation of neutrality and is based on the differences between the number of singletons (mutations appearing only once among the sequences) and the total number of mutations. In the population Q1,  $D^*$  was  $-1.9$  in three regions: between 226 and 400 bp (includes S359), between 951 and 1,125 bp and



**Fig. 1** Scheme of the functional structure of the *VvGAI1* protein: **a** DELLA region (grey bands are DELLA and TVHYNP motifs), **b** Valine, serine and threonine rich region, **c** LXXLL motif, implicated in protein–protein interaction,

**d** RVER motif, **e** SAW motif; brackets denote leucine heptad repeats; asterisks denote amino acid polymorphisms; GAI-a, GAI-b, GAI-c and GAI-d correspond to the DNA sequenced fragments

**Table 2** Description of polymorphisms detected in the core collection (C127) and in each population (Q1, Q2 and Q3)

Diversity	C127	Q1	Q2	Q3
Accessions	127	40	46	41
Total SNPs	15	12	13	14
Unique SNPs	1	4	2	1
SNPs in non-coding region	1	1	1	1
SNPs in coding region				
No amino acid change	9	8	8	8
Amino acid change	5	3	4	5
Haplotypes	15	10	9	9
$\pi$	0.0016	0.0013	0.0015	0.0018
$\theta$	0.0013	0.0013	0.0014	0.0015
Standard deviation $\theta$	0.0004	0.0004	0.0004	0.0004

between 1,651 and 1,800 bp (includes S1730). In the population Q2 D\* was  $-2$  between 651 and 800 bp and between 1,226 and 1,400 bp (includes S1320); in the population Q3, D\* was  $-1.9$  between 501 and 625 bp (includes S583) (Fig. 2).

The parameter  $r^2$  declined below 0.1 over 1,600 bp in the C127 (Fig. 3). Considering each population, differences were scarcely appreciated, but in Q3  $r^2$  also declined below 0.1 over 1,600 bp, while in Q1

and Q2 remained over 0.1. Certain polymorphisms were linked in different grade, being the highest linkage ( $r^2 > 0.7$ ,  $p < 0.0001$ ) between the pairs S30–S96, S369–S630, S359–S1026, S720–S1320 and S879–S1730 (supplementary Fig. 1).

The recombination rate without selection (R) was 0.0159 for the core collection, while for the Q1 population it was 0.0074, for Q2 0.0201 and for Q3 0.0222. A unique possible recombination event was detected between the sites S369 and S630 in the C127, while taking into account the different populations, only another one possible event was detected in the population Q2 between the sites S369 and S912.

#### Association test

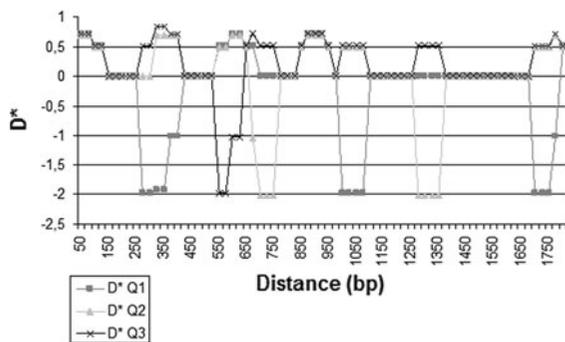
Significant associations at the 0.01 level were obtained for different traits: *bunch weight*, *juice yield* and several *texture parameters* (Table 4). The polymorphism S912 could be associated with *bunch weight* and *juice yield* variables explaining 4 and 6 % of their total phenotypic variance, respectively. The data of two years (2005 and 2006) for these descriptors showed a distribution in which the lowest values corresponded to the genotype C:C and the highest values were represented by the cultivars with genotype T:T (Fig. 4).

**Table 3** Allele frequencies found for each SNP observed in *VvGAI1*, and amino acid position where the polymorphism cause a replacement

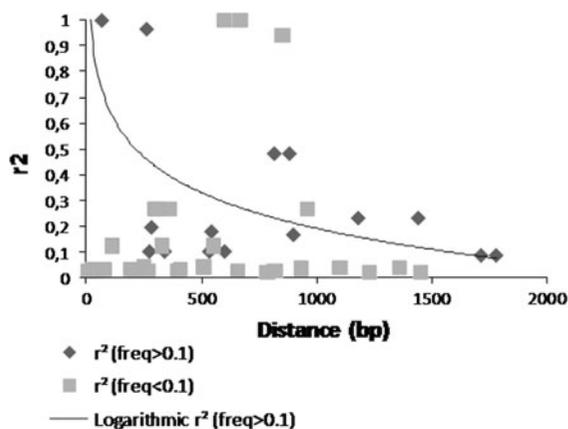
SNP locus	Mayor allele frequency	Minor allele frequency	Amino acid position	Amino acid replacement
S30	0.840	0.160		
S96	0.840	0.160		
<b>S304</b>	0.992	0.008	aa 102	Pro → Ser
S309	0.996	0.004		
<b>S359</b>	0.960	0.040	aa 120	Thr → Asn
S369	0.660	0.340		
<b>S583</b>	0.960	0.040	aa 195	Pro → Ala
S630	0.660	0.340		
S720	0.988	0.012		
S879	0.930	0.070		
S912	0.730	0.270		
S1026	0.960	0.040		
<b>S1320</b>	0.988	0.012	aa 440	Ile → Met
<b>S1730</b>	0.930	0.070	aa 577	Ala → Val
S1810 <sup>a</sup>	0.690	0.310		

<sup>a</sup> Non-coding region

The bold values indicate amino acid changing polymorphisms



**Fig. 2** Distribution observed for Fu and Li  $D^*$  in the three populations



**Fig. 3** Box plots showing linkage disequilibrium in the gene *VvGAI1*, measured with  $r^2$  index from inferred haplotypes using DnaSP

Regarding texture parameters, the significant associations obtained for S369 and S630 with the *deformation rate* and the *rupture slope* could explain the 6–7 % of their total phenotypic variance. Accessions with genotype T:T have lower *deformation rate* at the rupture point and higher *rupture slope* values (see S630 in Fig. 4). The genotype T:T appeared in 15 accessions, representing a frequency of 0.12. The *rupture slope* also could be associated with S879 and S1730, which explained the 6 % of the total phenotypic variance. The heterozygous genotype C:T (frequency of 0.12) and the homozygous genotype T:T (only in one accession) would seem associated with low values of *rupture slope* (see S1730 in Fig. 4). The *rupture force* and the *force at 20 %* associated with S359 and S1026, which explained the 4–5 % of their total phenotypic variance. The genotype A:C shows

lower values for those parameters and is present in 11 accessions (9 %) (see S359 in Fig. 4). In general, the variance explained by the markers was much lower (4–7 %) than the variance explained by the models (45–52 %).

Only one significant association at 0.01 level was obtained for the haplotypes: with the *juice yield* ( $p = 0.004$ ) (Table 5). Nevertheless, at 0.05 significance level, the *berry width* ( $p = 0.019$ ) and the *force at 10 %* ( $p = 0.024$ ) could be also associated. Though the variance explained by the model was similar to that obtained for each SNP marker, the variance explained by the marker was higher in this case (35 % for the *juice yield*, 23 % for the *berry width* and 33 % for the *force at 10 %*).

### Correlation within and between traits

We have observed distinct correlation rates between the different traits studied. Correlations between berry size descriptors (*berry width*, *length*, *weight* and *volume*) and *juice yield* were high ( $r > 0.6$ ) in 2005 and 2006. Correlations between bunch size descriptors (*bunch width*, *length* and *weight*) were high too, except between the *bunch length* and *width* of 2006 ( $r = 0.5$ ).

High correlations were also obtained between most of the texture parameters in the same year. The highest values were between the *force at 10 %*, the *force at 20 %* and the *rupture slope*, and between the *rupture force* and the *force at 20 %* ( $r > 0.7$ ). The *rupture force* and the *rupture slope* also showed a high value ( $r = 0.57$ ), as *deformation rate* and *force at 10 %* but in opposite direction ( $r = -0.65$ ). The correlation between the *deformation rate* and the *rupture slope* was also negative ( $r = -0.5$ ).

Regarding the correlation between both years data, only correlations for the *number of inflorescence per shoot* and the *length of peduncle* were very low ( $r = 0.4$  and  $0.3$ , respectively). All the mentioned correlations were significant ( $p < 0.01$ ).

## Discussion

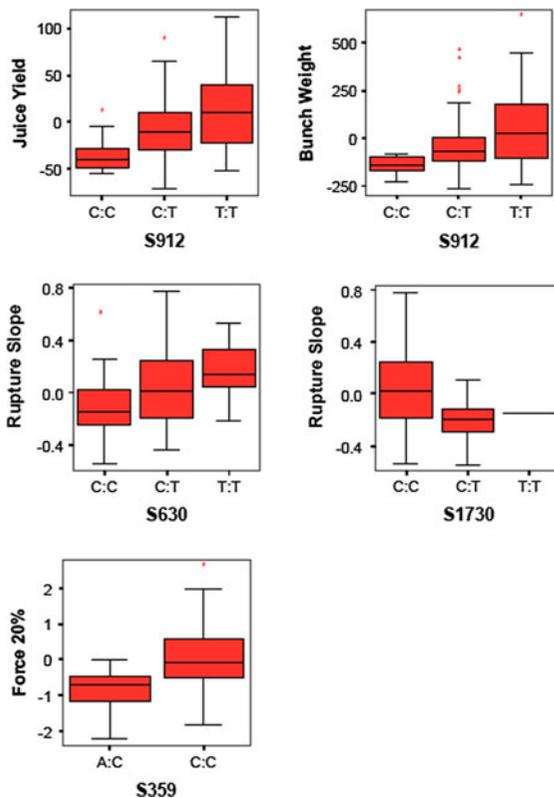
### Genetic structure

The data indicate that the structure obtained for the selected core collection is consistent to use it for the association analysis. Our results corroborate the

**Table 4** Significant associations found between trait descriptors and SNP markers

Trait	Descriptor	Locus	F_marker	<i>p</i> value	Rsq_model	Rsq_marker
Bunch weight	<i>Bunch weight</i>	S912	5.0483	0.0079	0.4818	0.0440
Juice yield	<i>Juice yield</i>	S912	7.0892	0.0012	0.4791	0.0621
Berry texture	<i>Deformation rate</i>	S630	5.8102	0.0042	0.4788	0.0651
		S369	5.7714	0.0043	0.4784	0.0647
	<i>Force at 20 %</i>	S1026	7.8162	0.0064	0.4756	0.0494
		S359	7.8162	0.0064	0.4756	0.0494
		S1026	7.8401	0.0062	0.4633	0.0448
	<i>Rupture force</i>	S359	7.8401	0.0062	0.4633	0.0448
		S630	6.9104	0.0016	0.5184	0.0716
	<i>Rupture slope</i>	S369	6.4465	0.0024	0.5142	0.0674
		S879	5.6534	0.0048	0.5068	0.0600
		S1730	5.4926	0.0056	0.5052	0.0584

Data calculated from two years of measures. F\_marker is the *F* test calculated as MS\_marker/MS\_error. *p* value is the probability of a larger *F* based on the *F* distribution. Rsq Model is the fraction of the total variation explained by the full model. Rsq\_Marker is the fraction of the total variation explained by the marker after fitting the other model effects

**Fig. 4** Genotypic effect of *VvGAI1* SNPs on the main associations found

grouping obtained by Aradhya et al. (2003) using nuclear microsatellites, that confirmed the existence of one group of Oriental cultivars, including Sultanina and derived varieties (Q1); another group of Muscat varieties (Q3); and a third group of Central European table and mixed use varieties (Q2). The latter one is represented in our work mainly by cultivars from the Iberian Peninsula, but that could be caused by differences on sampling. As Aradhya et al. (2003) suggested, this population could have originated by the fusion between the native wine European varieties and Oriental varieties that would have arrived through South Europe. The population Q2 has 13 % of Moroccan and Oriental cultivars, evidence that supports this hypothesis.

Organelle DNA information has proven quite useful in a variety of evolutionary studies. Chloroplast microsatellites have been used to investigate the origin of grapevine (Arroyo-García et al. 2006; Imazio et al. 2006). Such studies suggest the usefulness of chloroplast genome markers to provide information on haplotype distributions, which could help to identify further geographical areas for grapevine varietal evolution. We found different chlorotype distributions in the three populations detected in our sample. Chlorotype A (with a high prevalence in Q2) groups

**Table 5** Significant associations between phenotypic trait descriptors and haplotypes at 0.05 significance level

Trait/descriptor	F_marker	p_marker	Rsq_model	Rsq_marker
Berry width	1.8325	0.0198	0.5358	0.2286
Juice yield	2.1676	0.0041	0.3900	0.3555
Berry texture/force at 10 %	1.9461	0.0244	0.5418	0.3285

Only haplotypes with frequencies higher than 1 % were used for the association tests. F\_marker: F-test calculated as MS marker/MS error; p\_marker: probability of a larger F based on the F distribution; Rsq\_marker: fraction of the total variance explained

most of cultivars from the Iberian Peninsula, in agreement with Arroyo-García et al. (2006) and Cunha et al. (2009). Chlorotype C (prevalent in Q3) is the major one among samples of Central European origin (Arroyo-García et al. 2006). And chlorotype D followed by C (the same as in Q1) are very common among Near East and Middle East Europe samples, and less frequent among Western Europe cultivars (Arroyo-García et al. 2006). Differences in distributions of chlorotype frequencies between the different populations were consistent with the genetic structure detected using nuclear microsatellites.

#### Diversity and molecular evolution

The gene *VvGAI1* showed low nucleotide diversity when compared with other studies in grapevine, where a frequency of one SNP each 47–64 bp was detected on average (Emanuelli et al. 2010; Le Cunff et al. 2008; Lijavetzky et al. 2007; Salmaso et al. 2004; This et al. 2007). This could be related with the fact that the sequenced non coding region was very short (4 %), but also with a probably low mutation rate because of the important biological function of *VvGAI1*. The diversity detected in the gene *Dwarf8*, which in maize affect the quantitative variation of flowering time and plant height (and whose wheat orthologs (*Rht-B1/Rht-D1*) contributed to the increased yields seen in the 'Green Revolution' varieties (Peng et al. 1999)), was also low in comparison with other genes in maize (Thornberry et al. 2001). In fact, *VvGAI1* has been used as a marker in the *Vitaceae* family (Wen et al. 2007). Focusing only on the coding region, our results are similar to those obtained by Le Cunff et al. (2008) for three candidate genes (one SNP each 127 bp, on average). Nevertheless, the haplotype diversity was higher than twice the obtained by Lijavetzky et al. (2007) and Salmaso et al. (2004) in grapevine. In other plants, the haplotype diversity detected is also lower:

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 for 28 genes analyzed within 19 accessions. Tenaillon et al. (2004) also observed a significant haplotype excess in *Dwarf8*, explaining it as a possible consequence of the accrual of singletons after a severe selective sweep. More recently, Fournier-Level et al. (2009) estimated between 24 and 42 haplotypes for three *Myb* genes in grapevine.

Finding a low diversity is not the only sign of a possible important biological function of the gene. Genes involved in important physiological processes, such as *VvGAI1*, could constitute selection targets. So, it is interesting to study its evolutionary history to unmask possible traces of selection occurring along the domestication and/or in recent selection processes. These traces of selection could support possible associations detected between SNPs in the gene of interest and traits probably selected by the human being. Neutrality tests are very restrictive, especially Tajima Test. So, we used Fu and Li Test and considered significance at the 0.1 level. The difference between these tests is caused by the susceptibility to the singleton polymorphisms; Tajima's D is negative if there is an excess of rare polymorphisms, while Fu and Li's D\* is negative if an excess of singletons is present. Traces of positive selection were detected over the gene *VvGAI1* when it was analyzed separately in each population over 100 bp windows. In the population Q1, the regions that showed significant values for D\* included the amino acid changing SNPs S359 and S1730, and the silent SNP S879, which is linked to S1730. In the population Q2, the regions with significant values for D\* matched with the non-synonymous SNP S1320 and its linked marker S720. In the population Q3, the significant region included another amino acid changing marker (S583). It is difficult to know the underlying cause for such

significant deviations from neutrality. They could be because of the existence of selection, but also they could be a consequence of possible bottlenecks in the grapevine evolutionary history, for example during the phylloxera invasion, but could also be caused by LD with other genes of the genome subjected to selection, what is frequent in cultured species.

Nevertheless, two of those SNPs involved (S359 and S1730) were also associated with different texture parameters, pointing to a possible real selection process. The polymorphism S583 associated with *force at 20 %* too, though with a lower significant level ( $0.01 < p < 0.05$ ; not shown in table). These data could support the reliability of the associations obtained.

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. LD along the gene was low enough to provide a good resolution in the association analyses ( $r^2 < 0.1$  over 1,600 bp). This low LD could indicate absence of selection, but the  $r^2$  value is similar to another gene thought to be the object of artificial selection, *VvmybA1*, the major gene responsible for berry colour, for which  $r^2$  falls below 0.1 in around 700 bp (This et al. 2007). In *Dwarf8*, a supposed highly selected gene,  $r^2$  falls below 0.2 in about 1,500 bp and below 0.1 in 2,500 bp (Remington et al. 2001). So, the existence of selection in *VvGAI1* cannot be discarded.

#### Association analysis

Variations in the gene *VvGAI1* could have a pleiotropic effect, since it works as a negative regulator of gibberellin action, and these hormones are involved in many developmental steps. For this reason, different traits were included in the association study, and associations were obtained with very diverse traits. Nevertheless, the effect of the individual gene polymorphisms on the traits studied is relatively low. This fact could be due precisely to the involvement of this gene in important metabolic ways, limiting the possible variation to that of low or null influence in its function.

Our results suggest that the gene *VvGAI1* could be involved in berry firmness, since texture parameters associated with different SNPs. These data agree with those reported by Singh et al. (1978) and

Sato et al. (2004), who detected a slight increase in flesh firmness after GA<sub>3</sub> application. Berry texture is a very complex trait, and texture-meter helps to dissect it into simpler components.

*Force at 20 %* is highly correlated with *rupture force* ( $r = 0.76$ ), and both traits are associated with the SNP S359 (and with S1026, highly linked to S359), which has a minor allele frequency of 0.04. This polymorphism explained the 5 % of the *force at 20 %* total phenotypic variance and produces a change between threonine and asparagine (both polar amino acids) in the aa120, which is located near the DELLA region. The accessions with genotype A:C showed lower values for those parameters, which correspond to less firm berries (Fig. 4). No genotype A:A was detected for this SNP, maybe because it could contribute to form soft berries, which is not a desirable character in table grapes. Individuals with the homozygous genotype A:A could have suffered an artificial negative selection, making the a allele in S359 equivalent to a deleterious allele.

On the other hand, *rupture slope* associated with S1730 (and S879, highly linked) and explained 6 % of its total phenotypic variance. This polymorphism has a frequency of the minority nucleotide of 0.07 and produces a replacement between alanine and valine in the aa577, near the SAW motif, which is highly conserved in most GRAS proteins. The genotypes C:T and T:T associated with the lowest values of *rupture slope* (Fig. 4), that match up with softer berries.

S359 and S1730 are amino acid changing polymorphisms and present at low frequencies, which could be expected if a negative selection operated over these polymorphisms. The low frequencies make the associations less reliable, since they could be a random consequence. That is not the case of two linked silent polymorphisms (S369 and S630), which showed a higher frequency for the minority nucleotide (0.34). These SNPs associated with *rupture slope* and *deformation rate* explaining 6–7 % of the total phenotypic variance. The correlation between these texture parameters is negative ( $r = -0.5$ ,  $p < 0.01$ ) and their distribution against the S630 shows that accessions with genotype T:T have lower *deformation rate* in the rupture point and higher *rupture slope* values (Fig. 4). Together, these attributes appear in crunchy berries, which is a very important trait in table grapes. Given that the polymorphisms associated here are silent, it is likely there is another linked polymorphism not

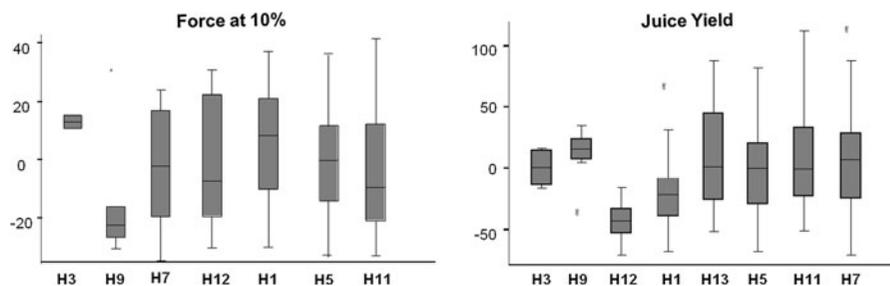
detected in this work, perhaps located in the promoter region, which has not been studied in this work. Moreover, the fact that SNPs explained low percentages of the total phenotypic variance in comparison with haplotypes could indicate that some of the SNPs analyzed here have not an important influence in gene function. Since in grapevine the LD decays slower when studied using haplotypes (Fournier-Level et al. 2009; Lijavetzky et al. 2007), the larger percentages explained when used for the association analyses could be pointing towards the existence of another SNP linked that could be the actual responsible for the variation. Other works have showed associations between polymorphisms in the regulatory region and traits such as color (Fournier-Level et al. 2009; Huang et al. 2012), while Fournier-Level et al. (2011) detected silent SNPs associated with the methylation level of anthocyanins, arguing that at least one SNP must be in LD with the causal variation. On the other hand, silent SNPs in the coding region of *VvDXS* associated to Muscat flavor due to the high LD with a non-synonymous polymorphism (Emanuelli et al. 2010).

The correlation between *force at 20 %* and *rupture slope* was very high, explaining their associations with the same polymorphisms. The high total phenotypic variance for the *force at 10 %* was explained by the haplotype variation. Interestingly, this could be explained by an additive effect of different SNPs on this texture parameter. Several haplotypes (H2, H4, H8 and H15) show very high values of *force at 10 %*, but they are represented only by one accession, which is not significant. However, the haplotype H9, which appeared in 11 accessions, associated with the lowest values of *force at 10 %* (Fig. 5). That points to this haplotype as an interesting target for future studies on berry texture as a representative of soft berry flesh (data not shown).

*Juice yield* and *bunch weight* associated with the silent polymorphism S912, and explained 6 and 4 % of total phenotypic variance, respectively. The data of 2 years (2005 and 2006) for these descriptors showed a distribution in which the lower values corresponded to the genotype C:C and the higher values were represented by the cultivars with genotype T:T (Fig. 4). The frequency of the minority nucleotide is high (0.27), supporting the credibility of the association, but it does not produce any amino acid replacement. So, it could be linked to another polymorphism not detected in this work that could be responsible for the phenotypic variation. The 2 year data for these traits were moderately correlated ( $r = 0.6$  in 2005 and  $r = 0.4$  in 2006;  $p < 0.01$ ), and indicates that other factors such as environment may affect these traits too.

*Juice yield* and *berry width* variation associated with the haplotypes had high values of their total phenotypic variance. These descriptors were highly correlated ( $r = 0.7$  for 2005 data and  $r = 0.8$  for 2006 data). The fact that SNPs explained low percentages of the total phenotypic variance in comparison with haplotypes could indicate that the SNPs analyzed here have not an important influence in the gene function. Another SNP linked to the interesting haplotypes could be the actual responsible for the variation. There are two interesting genes in the region: a putative pectinesterase (at a distance of 114,208 bp from *VvGAI1*), which could be involved in texture variation, and a major latex-like protein (at 53,012 bp from *VvGAI1*), associated with fruit and flower development in addition to plant pathogenesis responses and with a possible developmental role such as cell and tissue expansion (Ruperti et al. 2002). The LD obtained for the sequence analyzed is very low, pointing to no linkage with these genes, but the analysis of a larger region would be necessary to discard it.

**Fig. 5** Haplotypic effect of *VvGAI1* gene for force at 10 % and for juice yield. (Only haplotypes with frequencies higher than 1 % are shown). Haplotype numbering corresponds with haplotypes in Table S3



Though Boss and Thomas (2002) showed that *VvGAI1* was involved in the fertility in grapevine, the descriptor analyzed here (number of inflorescences per shoot) did not associate with any *VvGAI1* SNP or haplotype, from 2 years of measures. The cause could be related to environmental factors, since the correlation between the 2 years data was low for this trait, in agreement with Doligez et al. (2010). On the other hand, the mutation described by Boss and Thomas (2002) affecting DELLA box has not been detected in this work. In general, most of the described DELLA mutations give place to drastic phenotypic changes (Boss and Thomas 2002; Peng et al. 1999). However, Thornsberry et al. (2001) observed that *Dwarf8* was associated with minor quantitative variations in the blooming time, comparable to the associations detected here. The results found in this work should be taken into account when using the recently proposed model plant for grapevine studies: grape microvine (Chaïb et al. 2010), as this plant originated from a *VvGAI1* mutant, and several traits, in addition to those described by Boss and Thomas (2002), could be affected, slightly, but significantly.

## Conclusions

In this study we have searched for association between seventeen quantitative descriptors, related with fertility, bunch and berry traits and the *VvGAI1* gene. The use of both, phenotypic and genetic data, for the selection of the material to be studied, allowed us to build a robust core collection that included 100 % of genetic and phenotypic variability. Significant associations at 0.01 level for 2 years, were obtained with traits related to bunch weight, juice yield and berry texture. Although their stability must be further studied across years and environments, these associations could be an interesting starting point for breeding programs, given the importance of texture and berry and bunch dimensions for the table grape market.

**Acknowledgments** 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 predoctoral fellowship from IMIDRA. We thank M. Carmen Fajardo, 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.

## References

- Abdurakhmonov IY, Abdurakimov A (2008) Application of association mapping to understanding the genetic diversity of plant germplasm resources. *Int J Plant Genomics* 2008:574–927
- Aradhya MK, Dangl GS, Prins BH, Boursiquot JM, Walker MA, Meredith CP, Simon CJ (2003) Genetic structure and differentiation in cultivated grape, *Vitis vinifera* L. *Genet Res* 81:179–192
- Arroyo-García R, Ruiz-García L, Bolling L, Ocete R, López MA, Arnold C, Ergul A, Soyomezoglu G, Uzun HI, Cabello F, Ibáñez J, Aradhya MK, Atanassov A, Atanassov I, Balint S, Cenis JL, Costantini L, Goris-Lavets S, Grando MS, Klein BY, McGovern PE, Merdinoglu D, Pejic I, Pelsy F, Primikiris N, Risovannaya V, Roubelakis-Angelakis KA, Snoussi H, Sotiri P, Tamhankar S, This P, Troshin L, Malpica JM, Lefort F, Martínez-Zapater JM (2006) Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp sativa) based on chloroplast DNA polymorphisms. *Mol Ecol* 15:3707–3714
- Battilana J, Costantini L, Emanuelli F, Sevini F, Segala C, Moser S, Velasco R, Versini G, Grando MS (2009) The 1-deoxy-D-xylulose 5-phosphate synthase gene co-localizes with a major QTL affecting monoterpene content in grapevine. *Theor Appl Genet* 118:653–669
- Boss PK, Thomas MR (2002) Association of dwarfism and floral induction with a grape ‘green revolution’ mutation. *Nature* 416:847–850
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635
- Breseghele F, Sorrells ME (2006) Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Sci* 46:1323–1330
- Cabezas JA, Cervera MT, Ruiz-García L, Carreno J, Martínez-Zapater JM (2006) A genetic analysis of seed and berry weight in grapevine. *Genome* 49:1572–1585
- Chaïb J, Torregrosa L, Mackenzie D, Corena P, Bouquet A, Thomas MR (2010) The grape microvine—a model system for rapid forward and reverse genetics of grapevines. *Plant J* 62(6):1083–1092
- Ching A, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski AJ (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet*. doi:10.1186/1471-2156-3-19
- Chung SM, Staub JE (2003) The development and evaluation of consensus chloroplast SSRs for chloroplast genetic analysis. *Theor Appl Genet* 107:757–767
- Costantini L, Battilana J, Lamaj F, Fanizza G, Grando MS (2008) Berry and phenology-related traits in grapevine (*Vitis vinifera* L.): from quantitative trait loci to underlying genes. *BMC Plant Biol*. doi:10.1186/1471-2229-8-38

- Cunha J, Teixeira Santos M, Carneiro LC, Feveireiro P, Eiras-Dias JE (2009) Portuguese traditional grapevine cultivars and wild vines (*Vitis vinifera* L.) share morphological and genetic traits. *Genet Resour Crop Ev* 56:975–989
- Deluc LG, Grimplet J, Wheatley MD, Tillett RL, Quilici DR, Osborne C, Schooley DA, Schlauch KA, Cushman JC, Cramer GR (2007) Transcriptomic and metabolite analyses of Cabernet Sauvignon grape berry development. *BMC Genomics*. doi:10.1186/1471-2164-8-429
- Doligez A, Bouquet A, Danglot Y, Lahogue F, Riaz S, Meredith CP, Edwards KJ, This P (2002) Genetic mapping of grapevine (*Vitis vinifera* L.) applied to the detection of QTLs for seedlessness and berry weight. *Theor Appl Genet* 105: 780–795
- Doligez A, Bertrand Y, Dias S, Grolier M, Ballester JF, Bouquet A, This P (2010) QTLs for fertility in table grape. *Tree Genet Genomes* 6(3):413–422
- Emanuelli F, Battilana J, Costantini L, Le Cunff L, Boursiquot JM, This P, Grando MS (2010) A candidate gene association study on muscat flavor in grapevine (*Vitis vinifera* L.). *BMC Plant Biol*. doi:10.1186/1471-2229-10-241
- Ersoz ES, Yu J, Buckler ES (2007) Applications of linkage disequilibrium and association mapping in crop plants. In: Varshney RK, Tuberosa R (eds) *Genomics-assisted crop improvement: vol. 1 genomics approaches and platforms*. Springer, The Netherlands, pp 97–119
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14:2611–2620
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587
- Fanizza G, Lamaj F, Costantini L, Chaabane R, Grando MS (2005) QTL analysis for fruit yield components in table grapes (*Vitis vinifera*). *Theor Appl Genet* 111:658–664
- Fournier-Level A, Le Cunff L, Gomez C, Doligez A, Ageorges A, Roux C, Bertrand Y, Souquet JM, Cheynier V, 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
- Fournier-Level A, Huguency P, Verriès C, This P, Ageorges A (2011) Genetic mechanisms underlying the methylation level of anthocyanins in grape (*Vitis vinifera* L.). *BMC Plant Biol*. doi:10.1186/1471-2229-11-179
- Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics* 133:693–709
- Fusari CM, Lia VV, Hopp HE, Heinz RA, Paniego NB (2008) Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach. *BMC Plant Biol*. doi:10.1186/1471-2229-8-7
- Gouesnard B, Bataillon TM, Decoux G, Rozale C, Schoen DJ, David JL (2001) MSTRAT: an algorithm for building germ plasm core collections by maximizing allelic or phenotypic richness. *J Hered* 92:93–94
- Grimplet J, Deluc LG, Tillett RL, Wheatley MD, Schlauch KA, Cramer GR, Cushman JC (2007) Tissue-specific mRNA expression profiling in grape berry tissues. *BMC Genomics*. doi:10.1186/1471-2164-8-187
- Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Mol Biol* 57:461–485
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41:95–98
- Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K (2002) A comprehensive review of genetic association studies. *Genet Med* 4:45–61
- Houel C, Bounon R, Chaïb J, Guichard C, Péros JP, Bacilieri R, Dereeper A, Canaguier A, Lacombe T, N'Diaye A, Le Paslier MC, Vernerey MS, Coriton O, Brunel D, This P, Torregrosa L, Adam-Blondon AF (2010) Patterns of sequence polymorphism in the fleshless berry locus in cultivated and wild *Vitis vinifera* accessions. *BMC Plant Biol* 10:284–299
- Huang YF, Doligez A, Fournier-Level A, Le Cunff L, Bertrand Y, Canaguier A, Morel C, Miralles V, Veran F, Souquet JM, Cheynier V, Terrier N, This P (2012) Dissecting genetic architecture of grape proanthocyanidin composition through quantitative trait locus mapping. *BMC Plant Biol* 12:30–59
- Hudson RR (1987) Estimating the recombination parameter of a finite population-model without selection. *Genet Res* 50:245–250
- Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA-sequences. *Genetics* 111:147–164
- Ibáñez J, Vargas AM, Palancar M, Borrego J, de Andrés MT (2009) Genetic relationships among table-grape varieties. *Am J Enol Vitic* 60:35–42
- Imazio S, Labra M, Grassi F, Scienza A, Failla O (2006) Chloroplast microsatellites to investigate the origin of grapevine. *Genet Resour Crop Ev* 53:1003–1011
- Karaagac E, Vargas AM, de Andrés MT, Carreño I, Ibáñez I, Carreño J, Martínez-Zapater JM, Cabezas JA (2012) Marker assisted selection for seedlessness in table grape breeding. *Tree Genet Genomes*. doi:10.1007/s11295-012-0480-0
- Khan MA, Korban SS (2012) Association mapping in forest trees and fruit crops. *J Ex Bot*. doi:10.1093/jxb/ers105
- Le Cunff L, Fournier-Level A, Laucou V, Vezzuli S, Lacombe T, Adam-Blondon A, Boursiquot JM, This P (2008) Construction of nested genetic core collections to optimize the exploitation of natural diversity in *Vitis vinifera* L. subsp *sativa*. *BMC Plant Biol*. doi:10.1186/1471-2229-8-31
- Lijavetzky D, Ruiz-García L, Cabezas JA, De Andrés MT, Bravo G, Ibáñez A, Carreño J, Cabello F, Ibáñez J, Martínez-Zapater JM (2006) Molecular genetics of berry colour variation in table grape. *Mol Genet Genomics* 276:427–435
- Lijavetzky D, Cabezas JA, Ibáñez A, Rodriguez V, Martínez-Zapater JM (2007) High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. *BMC Genomics*. doi:10.1186/1471-2164-8-424
- Mejía N, Gebauer M, Muñoz L, Hewstone N, Muñoz C, Hinrichsen P (2007) Identification of QTLs for seedlessness, berry size, and ripening date in a seedless x seedless table grape progeny. *Am J Enol Vitic* 58(4):499–507

- Mejía N, Soto B, Guerrero M, Casanueva X, Houel C, Miccono MA, Ramos R, Le Cunff L, Boursiquot JM, Hinrichsen P, Adam-Blondon AF (2011) Molecular, genetic and transcriptional evidence for a role of VvAGL11 in stenospermocarpic seedlessness in grapevine. *BMC Plant Biol*. doi:10.1186/1471-2229-11-57
- Mullins MG, Bouquet A, Williams LE (1992) *Biology of grapevine*. Cambridge University Press, Cambridge
- Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES (2009) Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* 21:2194–2202
- Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. *Trends Plant Sci* 9:325–330
- Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York
- Olszewski N, Sun TP, Gubler F (2002) Gibberellin signaling: biosynthesis, catabolism, and response pathways. *Plant Cell* 14:S61–S80
- Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP (1999) 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400:256–261
- Pritchard JK, Stephens M, Donnelly P (2000a) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000b) Association mapping in structured populations. *Am J Hum Genet* 67:170–181
- Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doeblay J, Kresovich S, Goodman MM, Buckler ES (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. *P Natl Acad Sci USA* 98:11479–11484
- Rozas J, Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15:174–175
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132:365–386
- Ruperti B, Bonghi C, Ziliotto F, Pagni S, Rasori A, Varotto S, Tunutti P, Giovannoni JJ, Ramina A (2002) Characterization of a major latex protein (mlp) gene down-regulated by ethylene during peach fruitlet abscission. *Plant Sci* 163:265–272
- Salmaso M, Faes G, Segala C, Stefanini M, Salakhutdinov L, Zyprian E, Toepfer R, Grando MS, Velasco R (2004) Genome diversity and gene haplotypes in the grapevine (*Vitis vinifera* L.), as revealed by single nucleotide polymorphisms. *Mol Breeding* 14:385–395
- Sato A, Yamada M, Iwanami H, Mitani N (2004) Quantitative and instrumental measurements of grape flesh texture as affected by gibberellic acid application. *J Jpn Soc Hortic Sci* 73:7–11
- Sefc KM, Regner F, Turetschek E, Glössl J, Steinkellner H (1999) Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. *Genome* 42:367–373
- Singh K, Weaver RJ, Johnson JO (1978) Effect of applications of gibberellic-acid on berry size, shatter, and texture of Thompson seedless grapes. *Am J Enol Vitic* 29:258–262
- Sorkheh K, Malysheva-Otto LV, Wirthensohn MG, Tarkesh-Esfahani S, Martínez-Gómez P (2008) Linkage disequilibrium, genetic association mapping and gene localization in crop plants. *Genet Mol Biol* 31:805–814
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978–989
- Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS (2004) Selection versus demography: a multilocus investigation of the domestication process in maize. *Mol Biol Evol* 21:1214–1225
- This P, Lacombe T, Cadle-Davidson M, Owens CL (2007) Wine grape (*Vitis vinifera* L.) color associates with allelic variation in the domestication gene VvmybA1. *Theor Appl Genet* 114:723–730
- Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES (2001) Dwarf8 polymorphisms associate with variation in flowering time. *Nat Genet* 28:286–289
- Vargas AM, de Andrés MT, Borrego J, Ibáñez J (2009) Pedigrees of fifty table-grape cultivars. *Am J Enol Vitic* 60(4):525–532
- Watterson GA (1975) Number of segregating sites in genetic models without recombination. *Theor Popul Biol* 7:256–276
- Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* 42:9–19
- Wen J, Nie Z-L, Soejima A, Meng Y (2007) Phylogeny of Vitaceae based on the nuclear GAI1 gene sequences. *Can J Bot* 85:731–745
- Winkler AJ, Cook JA, Kliewer WM, Lider LA (1974) *General viticulture*. University of California Press, Berkeley
- Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. *Plant Genome* 1:5–20