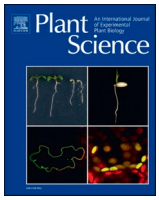
Plant Science 306 (2021) 110875

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci



Genetic variation and association analyses identify genes linked to fruit set-related traits in grapevine

Lalla Hasna Zinelabidine [a,b], Rafael Torres-Pérez [a,c], Jérôme Grimplet [a,d,e], Elisa Baroja [a], Sergio Ibáñez [a], Pablo Carbonell-Bejerano [a,f], José Miguel Martínez-Zapater [a], Javier Ibáñez [a], Javier Tello [a,*]

[a] Instituto de Ciencias de la Vid y del Vino (CSIC, UR, Gobierno de La Rioja), Logroño, 26007, Spain
[b] Laboratory of Biotechnology and Valorisation of Plant Genetic Resources, Faculty of Sciences and Techniques, University of Sultan Moulay Slimane, Beni Mellal, 23000, Morocco
[c] Servicio de Bioinformática para Genómica y Proteómica (BioinfoGP), Centro Nacional de Biotecnología (CNB-CSIC), Madrid, 28049, Spain
[d] Unidad de Hortofruticultura, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Zaragoza, 50059, Spain
[e] Instituto Agroalimentario de Aragón-IA2 (CITA-Universidad de Zaragoza), Zaragoza, 50059, Spain
[f] Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076, Tuebingen, Germany

ARTICLE INFO

Keywords:
Berry number
Flower number
Single Nucleotide Polymorphism (SNP)
Targeted sequencing
Vitis vinifera L.
Yield

ABSTRACT

Grapevine is one of the most valuable fruit crops in the world. Adverse environmental conditions reduce fruit quality and crop yield, so understanding the genetic and molecular mechanisms determining crop yield components is essential to optimize grape production. The analysis of a diverse collection of grapevine cultivars allowed us to evaluate the relationship between fruit set-related components of yield, including the incidence of reproductive disorders such as coulure and millerandage. The collection displayed a great phenotypic variation that we surveyed in a genetics association study using 15,309 single nucleotide polymorphisms (SNPs) detected in the sequence of 289 candidate genes scattered across the 19 grapevine linkage groups. After correcting statistical models for population structure and linkage disequilibrium effects, 164 SNPs from 34 of these genes were found to associate with fruit set-related traits, supporting a complex polygenic determinism. Many of them were found in the sequence of different putative MADS-box transcription factors, a gene family related with plant reproductive development control. In addition, we observed an additive effect of some of the associated SNPs on the phenotype, suggesting that advantageous alleles from different loci could be pyramided to generate superior cultivars with optimized fruit production.

1. Introduction

Grapevine is one of the most important fruit crops worldwide. According to 2018 statistics, the global area covered by vines represents 7.4 million hectares, which yielded 77.8 million tons of grapes [1]. This worldwide production is aimed to sustain wine elaboration (57 % of the annual yield), followed by fresh grape and dried grape markets (36 % and 7 %, respectively). Within the different grapevine *Vitis* spp., *V. vinifera* L. cultivars account for most of the area dedicated for grape production worldwide [2].

Differences in the grape cluster and berry at maturity between the cultivated grapevine (*V. vinifera* L. ssp. *sativa*) and its wild ancestor (*V. vinifera* L. ssp. *sylvestris*) evidence that individuals with better fruit production were selected during domestication and selection processes [3]. The most important change during grapevine domestication was the selection of hermaphrodite individuals over dioecious wild plants [3,4], which ensured a more regular grape production every season. Further development of viticulture practices likely fostered the selection of more productive vines to maximize fruit production for wine elaboration [5]. Therefore, traits like plant fertility, cluster weight and berry weight were targets of artificial selective pressures to increase grape yield [3,4]. The selection of individuals with better fruit production led to the indirect

Abbreviations: BN, Berry number; CI, Coulure index; FN, Flower number; FS, Fruit set; MI, Millerandage index; SN, Seed number.

* Corresponding author at: Instituto de Ciencias de la Vid y del Vino (ICVV), Finca La Grajera, Ctra. de Burgos Km. 6 (LO-20 – Salida 13), Autovía del Camino de Santiago, 26007, Logroño, La Rioja, Spain.
E-mail address: javier.tello@icvv.es (J. Tello).

https://doi.org/10.1016/j.plantsci.2021.110875
Received 3 February 2021; Received in revised form 2 March 2021; Accepted 6 March 2021
Available online 8 March 2021
0168-9452/© 2021 Elsevier B.V. All rights reserved.

and unconscious human-driven selection of other traits, like seed shape, related to the ability of the seed to germinate [6], and pollen viability, which optimizes ovule fertilization and so seed and berry formation [7]. That selection process boosted the development of viticulture and winemaking practices, and contributed to shape the genetic diversity that exists nowadays within the cultivated grapevine. In addition, this diversity was increased through divergent selection pressures caused by the different use of the grapes either as table- or wine-grape cultivars [8].

The number of flowers per inflorescence is an important component of grapevine yield. In temperate climates, flower number is determined during two consecutive seasons [4], which include the formation of the inflorescence primordia in season 1, and the differentiation of individual flowers in the spring of season 2, which occurs shortly before and during budbreak time [9]. Adverse environmental conditions during both seasons may hinder flower formation in susceptible cultivars [10], which might result in low fruit production at harvest time [11]. The proportion of flower ovaries successfully becoming fruits is known as fruit set. In grapevine, the final number of berries in the cluster is definitively determined one or two weeks after anthesis [12]. A recent study indicates that fruit set is about 50 % in grapevine cultivars used for wine production, whereas this value is lower in table grape cultivars [13]. The successful conversion of flower ovaries into berries can be affected by genetically-determined abnormal reproductive dysfunctions that lead to an excessive number of seedless berries (traditionally termed "chickens") and/or live green ovaries (also known as "shot" berries) in the mature cluster [11,14]. Chicken berries are either seedless or contain seed traces, and they are formed either by parthenocarpy or by stenospermocarpy [15]. These berries present a smaller size compared to seeded (normal) berries [14]. Live green ovaries are proposed to result from successful pollination without ovule fertilization, resulting in a small, hard and green ovary with limited development [14]. Whereas the presence of live green ovaries in the cluster does not affect significantly crop yield [14–16], seedless berries can represent a relevant proportion of the berries in mature clusters of susceptible cultivars, like observed for cv. Merlot [14]. The presence of live green ovaries and seedless berries in seeded cultivars led to the definition of two abnormal reproductive phenomena in the grapevine, termed coulure (or shatter) and millerandage. According to May [17], the term coulure is used when an excessive number of flowers or very young berries shed, resulting in a mature cluster with a low number of berries. On the other hand, millerandage is the phenomenon in which the proportion of small seedless berries and/or live green ovaries is high compared to that of normal berries in the mature cluster. In this regard, two indices have been defined for the accurate and quantitative assessment of coulure and millerandage reproductive disorders, based on the relationship between the number of seeded berries, seedless berries and live green ovaries in the mature cluster (millerandage), and with respect to the number of flowers in the inflorescence (coulure) [11,16]. Coulure and millerandage reproductive disorders are increased by both high and low pre-flowering temperatures, as they interfere in the successful development of pollen, ovule and pistils structures and their functioning during the flowering process [17]. Inadequate pre-flowering water and nutritional status also impairs fruit set, increasing coulure and millerandage rates [15,17].

Grapevine fruit set and reproductive disorders have been widely analysed from a physiological point of view [18–21], and different viticulture strategies have been proposed for fruit set optimization in the vineyard [15,16,22–24]. Nevertheless, and despite their interest for long-term vineyard management strategies and for the breeding of new cultivars, little is known about the genetic basis of these traits. Berry number has been scarcely studied through the detection of Quantitative Trait Loci (QTL) in biparental populations. The analysis of several table grape progenies segregating for this trait highlighted a series of QTLs capable of explaining a low proportion of phenotypic variance [25–27]. According to a more recent report [28], some additional QTLs for berry number were found in linkage groups (LGs) 10 and 18, whose confidence interval includes 14 and 16 candidate genes, respectively. On the other hand, the analysis of this trait using phenotypic data from more than 100 cultivars and 7032 genetic polymorphisms from 182 candidate genes revealed a significant association between the number of berries in the cluster and a polymorphism in the gene sequence of a MYB-type transcription factor [29]. The low heritability values reported for the berry number trait [26,29] support the general instability of the detected QTLs over different seasons, indicating a high sensitivity of this trait to seasonal variation, what hinders consistent QTLs identification [25,26]. Furthermore, and to our knowledge, no study has addressed yet the analysis of the genetic basis of other traits related to the determination of the yield of this crop, or of reproductive disorders like coulure or millerandage.

The main objective of this work was to evaluate the relationship between the phenotypic variation for six fruit set-related traits (flower number, berry number, fruit set rate, coulure index, millerandage index and seed number) and multiple genetic variants detected in the grapevine genome through association genetics. Genetic data was obtained through the targeted sequencing of 289 candidate genes in a set of 114 grapevine cultivars. These genes were selected according to different transcriptome profiling experiments and/or their functional annotation related to reproductive development control. As a reliable phenotype input data, best linear unbiased prediction (BLUP) values were calculated from previously reported phenotypic data obtained in the same grapevine collection [13,30], and then used for association tests. This approach allowed us to obtain valuable information to reveal the complex genetic architecture of fruit set and related traits in grapevine, as well as to establish links between the phenotypic diversity and a set of Single Nucleotide Polymorphisms (SNPs) found in a reduced number of genes. The information provided in this work could be exploited to obtain grapevine cultivars with improved fruit production.

2. Materials and methods

2.1. Plant material

A collection of 114 grapevine cultivars comprising 111 *Vitis vinifera* L. genotypes and 3 *Vitis* spp. interspecific crossings (grown as hybrid direct producers) has been used (Supplementary File 1). Cultivars were maintained as previously detailed [30] in two experimental sites: "*Finca Valdegón*" (Agoncillo, La Rioja, Spain) and "*Finca La Grajera*" (Logroño, La Rioja, Spain). Phenotypic trials were carried out in 2011 and 2012 in *Finca Valdegón*, whereas plants in *Finca La Grajera* were phenotyped in 2013, 2015, 2016 and 2017.

2.2. Phenotypic descriptions and data exploration

Phenotypic trials used ten inflorescences/mature clusters per cultivar and season, describing six fruit set-related traits in different seasons, as indicated in Table 1 and as detailed elsewhere [13,30]. Briefly, ten well-developed inflorescences from ten different shoots per cultivar were tagged and bagged before flowering time (E–L 17–18 [31]). After anthesis, bags were collected and their content was spread and scanned using an EPSON Perfection V370 Photo scanner. Calyptras were counted using the image-based tool described in Ibáñez et al. [13] and used to estimate flower number (FN). At harvest time (E–L 38), the same inflorescences (now mature clusters) were collected and used to evaluate the remaining fruit set-related traits. In this process, the three berry types described by Friend et al. [14] were separately counted by hand: seeded berries of normal size (B_{SD}), which contain seeds in seeded cultivars and ripen normally; smaller seedless berries (B_{SL}), which do not contain seeds but ripen; and live green ovaries (LGOs), which show a limited development and stay hard and green at harvest. These values were used to calculate the total number of berries in the cluster (BN =

Table 1
Mean, standard deviation (SD), minimum (Min.), maximun (Max.) phenotypic values and broad-sense heritability values (H^2) obtained for fruit set and related traits evaluated in a collection of 114 grapevine cultivars during three or six seasons.

Trait	Acronym	Season	Mean	SD	Min.	Max.	H^2
Berry number	BN	2011	136.3	47.5	42.5	272	0.48
		2012	108.9	39.0	37.8	210.2	
		2013	123.4	50.4	42.7	285.9	
		2015	147.2	65.3	20.7	376.8	
		2016	202.0	87.8	60.6	454.0	
		2017	155.3	69.6	51.4	356.6	
Coulure index	CI	2015	5.1	2.0	−0.1	8.6	0.72
		2016	4.9	2.6	−2.4	8.6	
		2017	5.5	2.1	0.1	9.1	
Flower number	FN	2015	405.3	209.7	85.3	1149.1	0.75
		2016	532.3	272.0	140.6	1255.9	
		2017	463.3	252.0	133.2	1288.7	
Fruit set (%)	FS	2015	43.2	0.2	12	101.0	0.72
		2016	47.5	0.2	13	114.0	
		2017	42.1	0.2	9	96.0	
Millerandage index	MI	2015	1.8	1.4	0	5.5	0.39
		2016	1.3	1.0	0	5.3	
		2017	1.3	1.3	0	10.0	
Seed number	SN	2011	2.0	0.5	0	3.1	0.64
		2012	2.2	0.6	0	3.8	
		2013	1.9	0.5	0	3.5	

$B_{SD} + B_{SL}$), the fruit set rate (FS= $\left(\frac{B_{SD}+B_{SL}}{FN}\right) \times 100$) and the coulure (CI= $10 - \left\{\frac{(B_{SD}+B_{SL}+LGOs)\times 10}{FN}\right\}$) and millerandage (MI= $10 - \left\{\frac{B_{SD} \times 10}{B_{SD}+B_{SL}+LGOs}\right\}$) indices in 2015, 2016 and 2017, after Collins et al. [16]. Data for the number of seeds per berry (SN) and the total number of berries per cluster (BN) for 2011, 2012 and 2013 were retrieved from Tello et al. [30]. Phenotypic data is available at https://github.com/jvtello/FruitSet_pheno.

Broad-sense heritability (H^2) was estimated for each trait as described elsewhere [29], using the variance components obtained by the MINQUE method by means of SPSS v.22.0 (IBM, Chicago, USA). From the phenotype data (mean values of 10 inflorescences/mature clusters), best linear unbiased prediction (BLUP) values were estimated using the *lmer* function of the *lme4* package [32] for R v. 3.6.2 (http://www.r-project.org/) to fit a linear mixed-effects model to the experimental data, using genotypes as random variables, and season and plot data as model covariates [33]. These BLUP values have a positive correlation with phenotypic raw values, and they were used for statistical analyses and genetic association tests. Phenotypic distributions and correlation between seasonal and BLUP values were explored using the *ggplot2* and *corrplot* R packages, respectively, using Pearson coefficients ($p < 0.05$) for correlation analyses. The correlation network was constructed using the *corrr* R package. Scripts are available at https://github.com/jvtello/FruitSet_scripts.

2.3. DNA extraction and SSR analysis: cultivar identification, population structure and kinship analyses

Young and fresh leaves were collected for each cultivar and frozen at −80 °C. DNA was extracted as previously detailed [34]. After DNA quality and concentration assessment by visual comparison with lambda DNA on ethidium bromide-stained agarose gels (0.8 %), and by means of a NanoDrop spectrophotometer (Thermo Scientific, Wilmington, USA), DNAs were screened at 25 nuclear SSR loci [33] for cultivar identification. To this aim, genetic profiles were pair-wise compared with the grapevine genetic profiles stored in the ICVV SSR-database.

This set of SSR markers was used for population structure estimation, using the Bayesian clustering method implemented in STRUCTURE v.2.3.4 (Pritchard et al., 2000) and assuming an admixture model with independent segregation of alleles. This method was run with a hypothetical number of genetic groups (K) ranging from 1 to 15, each one tested in five independent repetitions with a burn-in period of 100,000 followed by 150,000 MCMC iterations. The most probable levels of genetic stratification were determined according to the ΔK method [35] (as implemented in STRUCTURE HARVESTER [36]) and additional corrections for ΔK artefacts derived from markedly low likelihoods for $K = 1$ [37] (Supplementary File 2), analyzing the cultivars assigned to each genetic group. A cultivar was assigned to a certain group considering a critical membership coefficient (q-value) threshold of 0.75; otherwise it was considered as admixed. Then, the formed groups in each level of stratification under evaluation were examined in the light of the available information on cultivar origin and main use in the *V*IVC database (www.vivc.de). Once the optimum level of genetic stratification was set, the five repetitions were aligned in a single matrix using CLUMPP v.1.1 (Jakobsson and Rosenberg, 2007), and ultimately used as population structure correcting factor (Q) in association tests. On the other hand, this set of markers was also used to evaluate the pairwise relatedness between cultivars using the method indicated in Wang [38] and implemented in the *related* package for R [39], as previously detailed [34]. The resulting matrix was used as an additional correcting factor (K) in association tests.

2.4. Candidate genes selection and DNA re-sequencing

In this work, we used 289 candidate genes to test their association with fruit set-related traits. This set of genes includes the core set of 183 genes previously used in Tello et al. [29], which was derived from transcriptomic analyses related to cluster architecture and development. Here, this set was extended with 106 additional candidate genes selected from (i) diverse transcriptomic experiments aimed at the study of grapevine traits like seedlessness, berry size, flower sex and bud fertility ([40–42] and unpublished data), and (ii) their putative molecular function, including a series of MIKC-type genes of the MADS-box type II subfamily, suggested to play crucial roles in flowering and flower development [43,44]. The whole list of candidate genes analysed in this work can be found in the Supplementary File 3.

The annotated sequences of these 289 genes in the *V. vinifera* PN40024 grapevine reference genome (12X v1) were retrieved from the CRIBI server to delimit the regions for targeted sequencing. Update of the gene nomenclature according to the latest release of the genome assembly from the same scaffolds was performed *a posteriori* [45] (https://urgi.versailles.inra.fr/Species/Vitis/Annotations). As a general rule, we included up to 1 kb at the 5′ region for each selected gene to

sequence potential regulatory regions, unless another gene was predicted in these ranges. Introns were eliminated from gene sequences larger than 10 kb. These targeted regions were sequenced in the 114 grapevine genotypes by the Beijing Genomics Institute (BGI, Shenzhen, People's Republic of China) using a Illumina HiSeq 2000 platform as previously detailed [29,34]. Sequencing reads (with an average length of 90 nucleotides) were aligned to the whole PN40024 reference genome (12X v1) with the software Bowtie 2 [46] using the following command settings line:–phred64 –end-to-end -N 0 -L 25 –gbar 2 –np 6 –rdg 6,4 -X 400 –fr –no-unal. Single Nucleotide Polymorphisms (SNPs) between genotypes were detected by means of the SAMtools package v.0.1.19 [47]. Insertion/Deletion events were not considered in this work. Results were filtered to generate a consensus genotype per variety by means of an ad hoc Perl script available at https://github.com/ratope/VcfFilter. SNPs with a frequency of the minor allele (MAF) below 0.05 and those not mapped in at least 100 cultivars were discarded for association tests. As a result, 15,309 SNPs have been used for association analyses (Supplementary File 1). The position of these SNPs in the PN40024 grapevine reference genome were plotted using the R package *qtl* 1.42–8 [48].

2.5. Model testing and association tests

Marker-trait association tests between genotypic and phenotypic data were performed between 15,309 SNPs and the BLUP values obtained for six fruit set-related traits using TASSEL v.3.0 [49]. Within the four association models available, the Mixed Linear Model (MLM) correcting for both kinship (*K*) and structure (*Q*) effects provided the best control of type-I errors (false positives) for most of the traits evaluated (see QQ plots in the Supplementary File 4). Therefore, only the results obtained for the MLM are shown and discussed.

Given the increased risk of type I errors (false positives) when making multiple comparisons and aware of the high linkage disequilibrium (LD) present between the SNPs tested in this work [29], we used three LD-corrected thresholds, which correspond to the "suggestive", "significant" and "highly significant" thresholds proposed by Duggal et al. [50]. Following this method, we estimated the number of effective SNPs (or independent SNPs) present in our data as the sum of the number of haplotype blocks (haplo-blocks) plus all inter-block (unlinked) SNPs detected in each chromosome. Haplo-blocks were estimated using the solid spine of LD algorithm implemented in Haploview v.4.2 with a critical D' value of 0.80. As shown in the Supplementary File 5, results indicated the existence of 1461 effective SNPs. This number was used to calculate the three abovementioned LD-corrected *p*-value thresholds.

Besides, the presence of SNPs and genes simultaneously associated with more than one trait was explored by means of the UpSetR interactive graphical tool [51].

2.6. Evaluation of the individual and combined phenotypic effect of associated SNPs

Following the procedure suggested by Su et al. [52], we tested the combined effect of highly-associated SNPs on each trait. To simplify, we first selected the four SNPs from four different genes associated with each trait with the lowest *p*-value (according to TASSEL results). Then, we identified which of the two alleles of each selected SNP is predicted to increase the BLUP phenotypic value of each trait (advantageous allele). Then, we determined the number of advantageous alleles present in each cultivar (for each trait). This information was used as a dependent variable in a series of linear regression models aimed to predict BLUP phenotypic values. These models were calculated by means of SPSS, being considered significant at $P < 0.001$. Lastly, a one-way ANOVA followed by a Tukey post-hoc tests were calculated with SPSS to assess significant differences ($P < 0.05$) in BLUP values between grapevine cultivars varying in the number of these alleles. Boxplots showing the effect of the variable number of advantageous alleles on the distribution of BLUP values were obtained with the *ggplot2* package for R.

3. Results

3.1. Phenotypic diversity

The grapevine collection explored to analyze the genetic determinism of fruit set-related traits showed a high phenotypic variation (Table 1 and Supplementary File 6), supporting its suitability for this study. To cite some examples, BN varied by an 18.2-fold factor in 2015 (from 20.7 to 376.8 berries per cluster), FN by a 13.5-fold factor in 2015 (from 85.3 to 1149.1 flowers per inflorescence) and FS by a 10.7-fold factor in 2017 (from 9 to 96 %) (Table 1). Variation in these six traits was highly correlated among different years (Supplementary File 7), and a high significant correlation was observed between seasonal data and BLUP values for each trait too (Supplementary File 7). Broad-sense heritability values (H^2) were high to moderate, especially for FN (0.75), CI (0.72), FS (0.72) and SN (0.64) (Table 1). Altogether, these results suggest the existence of a strong genetic component for the traits analyzed in this work, supporting the interest of exploring their genetic basis through association genetics.

Attending to correlation coefficients between BLUP values (Fig. 1a, b and Supplementary File 7), a very high significant correlation between CI and FS was obtained (-0.97, $p < 0.05$). High significant correlation coefficients were also found between CI and FN (0.65, $p < 0.05$) and between FN and FS (-0.63, $p < 0.05$). The rest of significant pair-wise correlations between the traits evaluated in this work had lower coefficients (Fig. 1a and b). Interestingly, we found that SN only correlates significantly with MI (-0.22, $p < 0.05$).

3.2. Population structure

STRUCTURE analysis and ΔK criterion suggested $K = 2$ as the optimal level of genetic structure for this set of cultivars, with a likely second level of genetic structure at $K = 3$ (Supplementary File 2). At $K = 2$, 82 genotypes were assigned to two genetic groups (K2−1 and K2−2) considering a critical q-value threshold of 0.75 for group assignation (the remaining 32 genotypes were considered as admixed). The first genetic group (K2−1) was formed by 54 wine or multi-purpose grape cultivars mainly from Spain, and a small group of table grape cultivars from different origins, including Afus Ali, Cardinal, Cornichon, Delight and Dominga. The second group (K2−2) included 28 wine (or multi-purpose) grape cultivars, with some of the oldest West European cultivars included in this work (Savagnin, Pinot Noir, Cabernet Franc) and their descendants (Alfrocheiro, Molar, Silvaner Gruen, and Verdejo Blanco descend from Savagnin; Chardonnay Blanc and Gamay Noir from Pinot Noir; and Cabernet Sauvignon and Merlot Noir from Cabernet Franc), mainly from France and Portugal (Supplementary File 1). At $K = 3$, and considering the same q-value threshold for group assignment, 91 cultivars were successfully assigned to one out of the three genetic groups, with the 23 remaining cultivars considered as admixed. The first genetic group (K3−1) included 13 cultivars from different regions, but most of them with a clear aptitude for table grape production (like Cardinal, Dominga or Italia). The second group (K3−2) included 29 wine (or multi-purpose) grape cultivars, including 27 previously found in K2−2. The third genetic group (K3−3) contained 49 cultivars, most of them wine (or multi-purpose) Iberian grape cultivars like Airén, Cayetana, Palomino Fino, Parellada or Tempranillo Tinto (Supplementary File 1). Considering the agreement of this level of stratification with current knowledge of grapevine genetic structure at a species level [53–55], and that the ΔK criterion tends to artificially favor $K = 2$ due to markedly low likelihoods for $K = 1$ [37], the *q*-matrix obtained for $K = 3$ was considered as more appropriate and therefore used as population structure correcting factor in association tests.

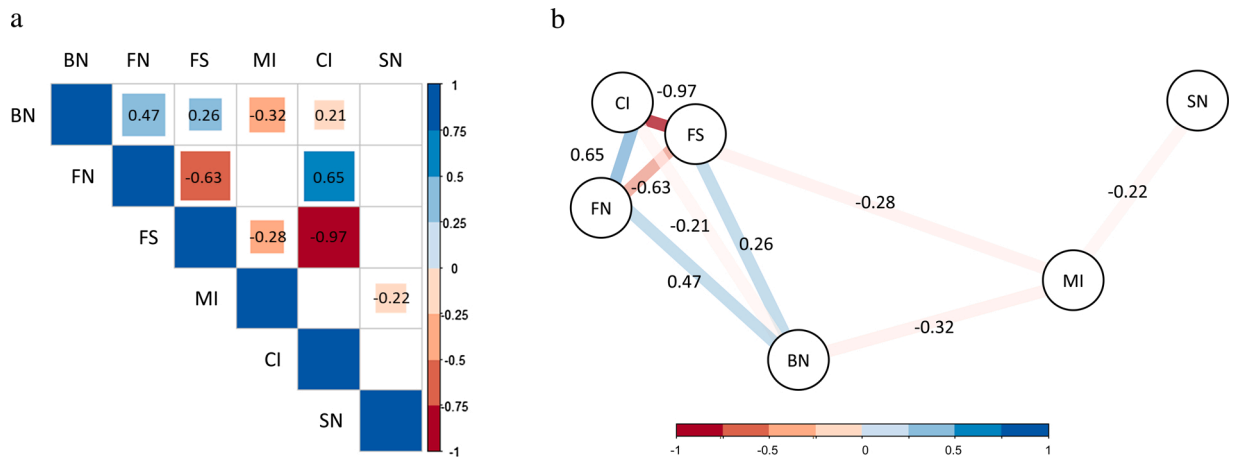


Fig. 1. Correlogram (a) and correlation network (b) based on the pairwise Pearson's correlation coefficients obtained between BLUP values for fruit set and related traits. Only significant correlation coefficients (p-value <0.05) are shown. In a, squares size and colour vary according to correlation coefficients (blue-to-red scale), which are indicated. In b, the colour of the connecting lines varies according to correlation coefficients (see blue-to-red scale). BN: Berry number; CI: Coulure index; FN: Flower number; FS: Fruit set; MI: Millerandage index; SN: Seed number (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

3.3. Targeted sequencing, SNP detection and linkage disequilibrium evaluation

A total of 289 candidate genes accounting for 993.9 kbp were targeted for re-sequencing, which represents 0.2 % of the PN40024 grapevine reference genome assembly sequence. Data can be accessed at the NCBI's BioProject PRJNA625274 (sequence read runs from SRR11547919 to SRR11548032). These genes were scattered throughout the 19 LGs of the grapevine reference genome, being LG18 the most represented one (169,221 bp, 41 genes) followed by LG5 (133,718 bp, 44 genes) and LG11 (79,844 bp, 22 genes) (Supplementary File 3). The analysis of such sequenced regions in 114 grapevine cultivars allowed us to identify 15,309 SNPs with a MAF > 0.05 (Supplementary File 1), which correspond to an average of one SNP every 65 bp. These genetic variants were found across all the LGs of the grapevine genome, with a high number of them located in LG18 (2659 SNPs), LG5 (1929 SNPs), LG12 (1475 SNPs) and LG11 (1366 SNPs) (Fig. 2 and Supplementary File 5). The LGs that showed less variants were LG6, LG14 and LG13, with 125, 229 and 282 SNPs, respectively.

The Haploview-generated LD-plot obtained from genetic data on 15,309 SNPs detected 1247 haplo-blocks linking 15,095 SNPs (98.6 %) (Supplementary File 5). These haplo-blocks linked a variable number of SNPs (from 2 to 101 SNPs, data not shown), and they were found well scattered across the reference genome. Thus, the LGs with a higher number of haplo-blocks were LG18 (207 haplo-blocks linking 2622 SNPs), LG5 (151 haplo-blocks linking 1915 SNPs) and LG12 (130 haplo-blocks linking 1453 SNPs). On the other hand, only 214 unlinked inter-block SNPs were found, many of them present in LG18, LG11 and LG12 (37, 22 and 22 SNPs, respectively). No unlinked SNPs were found in LG17. Following the method of Duggal et al. [50], we used this information to estimate the effective number of independent SNPs (effective SNPs). This approach led to the detection of 1461 effective SNPs, representing 9.5 % of the 15,309 SNPs initially detected (Supplementary File 5).

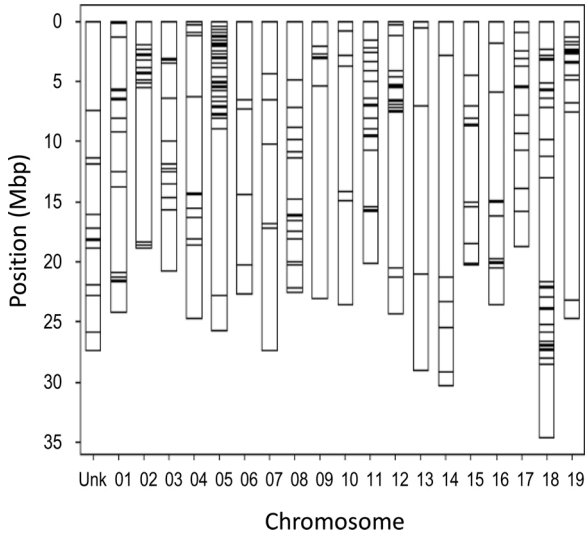


Fig. 2. Position of the SNPs detected in this work on the PN40024 grapevine reference genome (12X V1).

As indicated in materials and methods, the number of effective SNPs (1461) was ultimately used to calculate three different LD-corrected p-value thresholds: "suggestive" (6.84×10^{-4}), "significant" (3.42×10^{-5}) and "highly significant" (6.84×10^{-6}). They were used to control for the presence of type-I errors (false positives).

3.4. Association tests' results

In this work, we have evaluated a total of 91,854 marker-trait MLM association models (15,309 SNPs for six traits). Among them, 237 associations (0.24 %) were found to be significant under the "suggestive" threshold of 6.84×10^{-4} (Supplementary File 8). Attending to these results, we found four SNPs significantly associated with BN, 78 with CI, 20 with FN, 76 with FS, four with MI and 55 with SN (Table 2). Only 18 SNPs (0.02 %) were found to be significantly associated under the "significant" threshold of 3.42×10^{-5}, and eight of them (0.009 %) under the "highly significant" threshold of 6.84×10^{-6} (Table 2 and Supplementary File 8).

The four SNPs associated with BN were located in three genes coding for transcription factors of different families (NAC, MYB and MADS-domains transcription factors) located on LG1, LG7 and LG16. Regarding FS, the 76 associated SNPs were identified in seven genes located on LG18 (4 genes), LG11 (2 genes) and LG12 (1 gene), four of them coding for MADS-box transcription factors. We found 78 SNPs associated with CI, 73 of them also associated with FS (Table 2), indicating the intimate relationship between these two traits. SNPs associated with FN were found in 9 genes (on LG1, LG3, LG6, LG8, LG10, LG11, LG15 and LG18) coding for proteins with different functions, like a NtPRp27 secretory protein (Vitvi03g01597), an iron regulated

Table 2
Genes and SNPs significantly associated with fruit set and other fruit set-related traits. SNPs N indicate the number of SNPs in the gene significantly associated with a certain trait. The name of the SNP with the strongest association (lowest *p*-value in the MLM model) per gene/trait combination is indicated (Representative SNP). If two (or more) perfectly linked SNPs in the same gene obtained the same *p*-value, only one aleatory SNP is indicated.

Trait[1]	Gene ID[2]	Gene Functional Annotation (VCost.v3)	SNPs N	Representative SNP[3]	Model *p*-value[4]	Trait variance (%)[5]
	Vitvi01g01038	NAC domain-containing protein 29 (*NAP*, *NAC029*)	2	**SNP_07567**	4.09×10^{-4}*	11.05
BN	Vitvi07g00455	MYB domain protein 108b (*MYB108b*)	1	**SNP_2113**	1.78×10^{-4}*	12.57
	Vitvi16g00894	MADS-box AGAMOUS-LIKE 6 (*AGL6*)	1	**SNP_12846**	1.31×10^{-4}*	16.16
	Vitvi03g01320	MADS-box protein SHORT VEGETATIVE PHASE 4 (*SVP4*)	1	SNP_08605	6.50×10^{-4}*	13.02
	Vitvi11g00228	MYB domain protein 68 (*MYB068*)	1	SNP_3000	4.30×10^{-4}*	13.81
	Vitvi11g01393	Unknown	3	SNP_3044	1.55×10^{-4}*	15.78
	Vitvi12g00019	MADS-box protein SHATTERPROOF 2 (*SHP2*)	6	SNP_11181	2.98×10^{-4}*	14.51
CI	Vitvi16g00894	MADS-box AGAMOUS-LIKE 6 (*AGL6*)	1	SNP_12849	2.63×10^{-4}*	11.79
	Vitvi18g00517	MADS-box protein SHORT VEGETATIVE PHASE 2 (*SVP2*)	6	**SNP_13547**	1.02×10^{-4}*	16.60
	Vitvi18g00553	MADS-box AGAMOUS-LIKE 21 (*AGL21*)	51	**SNP_13714**	3.83×10^{-5}*	18.53
	Vitvi18g02145	MADS-box AGAMOUS-LIKE 12 (*AGL12*)	4	**SNP_14764**	1.38×10^{-4}*	16.01
	Vitvi18g02631	Unknown	5	**SNP_6049**	1.47×10^{-4}*	15.88
	Vitvi01g00008	MADS-box APETALA 1 (*AP1*)	2	**SNP_07474**	1.94×10^{-4}*	14.66
	Vitvi03g01597	NtPRp27 secretory protein	10	**SNP_08233**	4.23×10^{-5}*	17.52
	Vitvi03g01763	No hit	1	SNP_1031	5.67×10^{-4}*	16.98
	Vitvi06g01106	Cyclin CYCB1_2	1	**SNP_09966**	1.08×10^{-4}*	15.74
FN	Vitvi08g00630	U-box domain-containing protein	1	SNP_2646	6.14×10^{-4}*	12.55
	Vitvi10g01358	Iron regulated transporter	1	**SNP_2941**	7.54×10^{-5}*	16.43
	Vitvi11g01072	Potassium-sodium symporter HKT2	1	SNP_3672	3.14×10^{-4}*	13.77
	Vitvi15g00225	MADS-box protein SHORT VEGETATIVE PHASE 3 (*SVP3*)	1	SNP_12609	2.40×10^{-4}*	14.26
	Vitvi18g01884	Far-red impaired responsive family protein	2	SNP_15204	4.85×10^{-4}*	12.97
	Vitvi11g00228	MYB domain protein 68 (*MYB068*)	3	**SNP_3000**	1.36×10^{-4}*	15.96
	Vitvi11g01393	Unknown	2	SNP_3036	2.14×10^{-4}*	15.09
	Vitvi12g00019	MADS-box protein SHATTERPROOF 2 (*SHP2*)	5	SNP_11181	2.64×10^{-4}*	14.68
FS	Vitvi18g00517	MADS-box protein SHORT VEGETATIVE PHASE 2 (*SVP2*)	6	SNP_13547	1.65×10^{-4}*	15.59
	Vitvi18g00553	MADS-box AGAMOUS-LIKE 21 (*AGL21*)	51	**SNP_13714**	3.55×10^{-5}*	18.61
	Vitvi18g02145	MADS-box AGAMOUS-LIKE 12 (*AGL12*)	4	**SNP_14764**	8.38×10^{-5}*	16.91
	Vitvi18g02631	Unknown	5	**SNP_6049**	1.16×10^{-4}*	16.28
	Vitvi07g01441	MADS-box protein SHORT VEGETATIVE PHASE 1 (*SVP1*)	1	**SNP_07337**	5.55×10^{-4}*	11.31
MI	Vitvi11g00492	Axial regulator YABBY5	1	**SNP_3360**	2.22×10^{-4}*	13.04
	Vitvi17g00229	TERMINAL FLOWER 1 (*TFL1*)	1	**SNP_13301**	1.63×10^{-4}*	16.92
	Vitvi18g02631	Unknown	1	**SNP_6037**	3.20×10^{-4}*	12.35
	Vitvi02g01270	Receptor protein kinase	7	**SNP_0927**	4.99×10^{-6}***	24.60
	Vitvi02g01288	9-cis-epoxycarotenoid dioxygenase	1	SNP_08133	1.06×10^{-4}*	20.72
	Vitvi04g00573	TOPLESS-RELATED 1 (*TPR1*)	16	SNP_1081	4.22×10^{-5}*	19.90
	Vitvi05g00281	Binding	3	SNP_1413	5.14×10^{-5}*	19.47
	Vitvi05g00523	GCN5 N-acetyltransferase	1	SNP_1818	5.20×10^{-5}*	19.45
SN	Vitvi08g01710	BZIP transcription factor	2	SNP_2469	2.74×10^{-4}*	15.92
	Vitvi11g00355	Sulfate transporter 3.4	10	SNP_3091	3.47×10^{-5}*	20.32
	Vitvi12g00574	E-beta-ocimene synthase	3	**SNP_5240**	1.67×10^{-5}**	21.92
	Vitvi14g01341	MADS-box APETALA 1 (*AP1*)	10	**SNP_11736**	4.11×10^{-6}***	25.04
	Vitvi16g00120	Transcription factor jumonji (jmj)	1	**SNP_13044**	2.76×10^{-5}**	20.82
	Vitvi18g03021	Laccase	1	SNP_6431	5.41×10^{-4}*	14.50

[1] BN: Berry number; CI: Coulure index; FN: Flower number; FS: Fruit set; MI: Millerandage index; SN: Seed number.
[2] According to the VCost.v3 gene annotation version.
[3] SNPs indicated in bold were used to test their additive effect on BLUP variation.
[4] *: *p*-value $< 6.84 \times 10^{-4}$; **: *p*-value $< 3.42 \times 10^{-5}$, *** *p*-value $< 6.84 \times 10^{-6}$.
[5] According to TASSEL results.

transporter (Vitvi10g01358) and an *APETALA 1* MADS-box transcription factor (Vitvi01g00008). For MI, we found significant associations with SNPs located in the gene sequence of a *SHORT VEGETATIVE PHASE* (*SVP*) MADS-box transcription factor (Vitvi07g01441), a *YABBY* axial regulator (Vitvi11g00492), a homolog to *TERMINAL FLOWER 1* (*TFL1*) (Vitvi17g00229) and a protein of unknown function (Vitvi18g02631). Lastly, SNPs associated with SN were found in 11 genes on LG2, LG4, LG5, LG8, LG11, LG12, LG14, LG16 and LG18. Within these SNPs, we found those significantly associated even when considering the "highly significant" threshold used in this work (6.76×10^{-6}). They were found in two different genes: a homolog to *APETALA 1* (*AP1*) (Vitvi14g01341) and a gene coding for a receptor protein kinase (Vitvi02g01270). More detailed information is provided in Table 2 and in the Supplementary File 8.

UpSetR revealed the presence of 73 SNPs significantly associated with CI and FS (Fig. 3a). No other intersecting SNPs were found involving the other traits analyzed in this work. Regarding genes, we found six genes associated with CI and FS, one with BN and CI, and one with CI, FS and MI (Fig. 3b). Genes associated with CI and FS code for four MADS-box transcription factors (Vitvi12g00019, Vitvi18g00517, Vitvi18g00553 and Vitvi18g02145), one MYB domain protein (Vitvi11g00228), and one protein of unknown function (Vitvi11g01393). An *AGAMOUS 6* (*AGL6*) -like MADS-box gene (Vitvi16g00894) was found to associate with BN and CI, and we detected a gene coding for a protein of unknown function (Vitvi18g02631) associated with CI, FS and MI.

3.5. Analysis of the combined effect of associated SNPs

Lastly, we explored the combined effect of the associated SNPs on phenotypic variation, focusing on four SNPs (from four different genes) per trait (Supplementary File 9). For BN, only three SNPs were selected, as the significantly associated SNPs with this trait belong to only three genes (Table 2). The detailed evaluation of the grapevine accessions indicated a great variability in the number of alleles increasing BLUP phenotypic values. Thus, for a certain trait, we found cultivars with none advantageous alleles to others with up to eight advantageous alleles, as observed for FS for example (Fig. 4 and Supplementary File 10). Linear

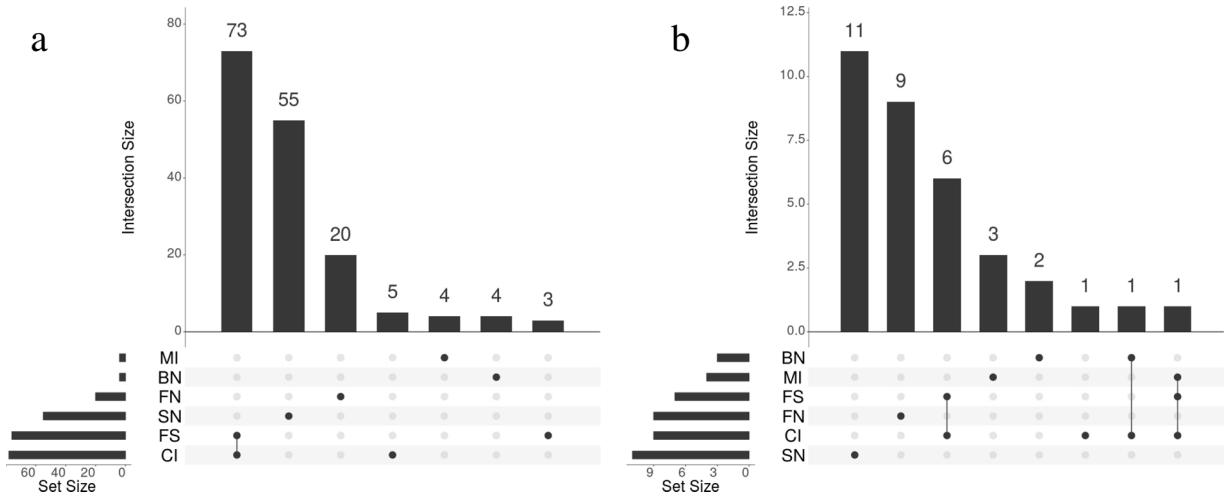


Fig. 3. UpSetR plots showing the number of SNPs (a) or genes (b) significantly (p-value $< 6.84 \times 10^{-4}$) associated with one or more of the traits analyzed in this work. The total number of SNPs/genes associated with a trait are shown on the bottom left corner of each plot. SNPs/genes associated with more than one trait are shown by links connecting filled circles. Vertical bars represent the number of specific or overlapping SNPs/genes between sets. BN: Berry number; CI: Coulure index; FN: Flower number; FS: Fruit set; MI: Millerandage index; SN: Seed number.

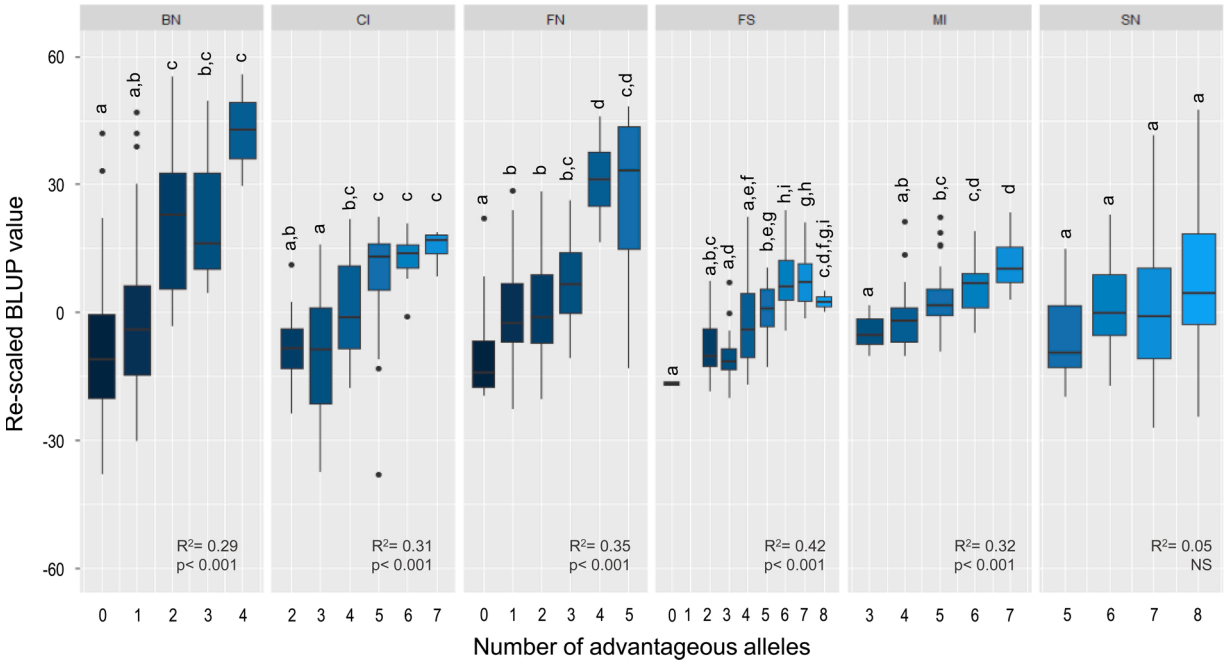


Fig. 4. Additive effect of alleles of associated SNPs on BLUP phenotypic values. Linear regression modelling results (R^2 and p-value) are shown in the lower corner of each boxplot. Different lowercase letters indicate significant differences between groups of grapevine cultivars differing in the number of alleles contributing to increase BLUP values ($p < 0.05$). Outliers are indicated as black dots. BN: Berry number; CI: Coulure index; FN: Flower number; FS: Fruit set; MI: Millerandage index; SN: Seed number.

modelling results indicated a significant additive effect of the number of these SNP alleles in all the traits except in SN, with R^2 values ranging from 0.42 to 0.29 (for FS and BN, respectively) (Fig. 4). For example, the mean BLUP value for FS of the grapevine cultivars with eight of these SNP alleles (i.e.: homozygous for the advantageous alleles in the four selected SNPs) was 6.8 times higher than that of those cultivars with none of these alleles (i.e.: homozygous for the non-advantageous alleles). Similarly, the mean BLUP value for FN was 2.5 higher in the cultivars with five advantageous alleles compared to those without them. One-way ANOVA and Tukey post-hoc tests confirmed these additive effects in BN, CI, FN, FS, and MI traits, at least between the groups of cultivars with the most different number of advantageous alleles (Fig. 4).

4. Discussion

Fruit set is one of the main determinants of grapevine yield [56]. Nevertheless, its genetic determinism has been scarcely explored, which contrasts with the wide knowledge available for other yield components like berry weight, cluster number or, to a lesser extent, berry number (see Vezzulli et al. [57] and references therein). This lack of information could be partially explained by the effect of environmental factors (hot and cold temperature, solar radiation, rainfalls) on fruit set [17], which makes difficult to separate genetic from environment effects on

phenotypic variation. For berry number, low phenotypic correlations between years due to high genotype × environment interaction have been previously reported [26,27], agreeing with the moderate value of broad-sense heritability found for this trait in this work ($H^2 = 0.48$). On the contrary, we found that traits like flower number, fruit set rate and coulure have a strong genetic basis, as inferred from their high H^2 values (Table 1). These results are in line with those of Dry et al. [11], who observed that environmental conditions (season, region, site) made a very small contribution to the phenotypic variability of some of these reproductive traits. To our knowledge, this is the first time that the broad-sense heritability values of the traits analyzed in this work are reported. Nevertheless, our H^2 values are within the range of those found for other fruit crops, such as peach and almond trees [58,59]. Another reason that hinders the genetic analysis of fruit set and related traits is the lack of an accurate and fast method for their objective quantification. To accurately determine these traits, flowers and berries should be counted in the same clusters, what requires a non-invasive method for this arduous task. Here, we used a recently reported image-based system capable to count flower caps after their digital scanning to ease this process [13]. As observed in Table 1, we obtained anomalous fruit set values above 100 % in some varieties, probably due to the trapping of some flower caps in the inner part of the clusters after cluster closure, an especially common feature in cultivars with highly compact clusters [60]. Nevertheless, differences between manual and automatic counts were low compared to other automatic systems (see discussion in Ibáñez et al. [13]), and it proved to provide a satisfactory flower number counting data for a proper calculation of derived parameters (such as fruit set rate or coulure and millerandage indices).

To explore the genetic basis of fruit set and related traits we used a collection of wine, table and multi-purpose grapevine cultivars with a great diversity for yield-related variables, including flower number, berry number and fruit set rate [13]. The correlation network obtained indicated a complex interaction between the six traits studied (Fig. 1b). As observed by Ibáñez et al. [13], we found that fruit set correlates negatively with the initial number of flowers in the inflorescence and with the indices used to measure the two abnormal reproductive phenomena considered in this work (coulure and millerandage). Additionally, the final number of berries in the cluster correlates positively with flower number and fruit set rate, and negatively with coulure and millerandage indices, in all cases with low correlation coefficients. These results support the compensation effect between the initial number of flowers and the final number of berries indicated by Dry et al. [11] and Ibáñez et al. [13], with cultivars having an initial high number of flowers per inflorescence showing low fruit set rates and vice versa. Here, correlation analyses indicate that the shedding of ovaries or very young berries is the most critical factor determining the fruit set rate at a species level. Little is known about the physiological basis of flower shedding, although it is suggested to be under a complex hormonal control, in which auxin transport and ethylene-related gene expression seem to play a leading role [61]. On the other hand, we found that the number of seeds per berry only correlates significantly with millerandage (Fig. 1a). Millerandage occurs when a high number of live green ovaries and/or seedless berries are present in the cluster [11]. Live green ovaries result from the lack of ovule fertilization [14], whereas seedless berries in seeded cultivars occur by parthenocarpy or by stenospermocarpy [15]. In parthenocarpy, small berries without seed traces develop in the absence of fertilization, whereas in stenospermocarpy, seed development aborts after fertilization and embryo formation, generating berries with seminal rudiments or seed traces [40,41]. Fertilization-derived problems also reduce the final number of seeds per berry [62], explaining the correlation observed here between seed number and millerandage.

We identified several significantly associated SNPs in the gene sequence of diverse MADS-box transcription factors (Table 2 and Supplementary File 8). MADS-box genes are known to play a wide range of functions in angiosperms, including flowering induction, flower organ identity and the control of flowering time [63]. The grapevine genome contains 90 MADS-box genes, which can be divided into the MADS type I and MADS type II monophyletic lineages [44]. MIKC-type genes (named after their characteristic domain structure that includes a MADS, an Intervening, a Keratin and a C-terminal domains) belong to the MADS-box type II genes, and they include a series of genes with well-documented functions in flower organogenesis in diverse species [64]. Multiple works indicate a conserved function of the MIKC-type genes in grapevine, as reported for *AGAMOUS-LIKE* (*AGL*), *APETALA* (*AP*), *SHATTERPROOF* (*SHP*), and *SHORT VEGETATIVE PHASE* (*SVP*) genes [65–69]. Our results indicated the presence of 51 SNPs in *VviAGL21* and 4 in *VviAGL12* associated with fruit set and coulure (Table 2). In Arabidopsis, *AtAGL21* regulates auxin accumulation in lateral roots [70], and *AtAGL12* is an auxin-responsive factor involved in root development and in flowering transition [71]. In grapevine, auxin transport has been related to the major or minor ability of berries to persist in the cluster during grapevine abscission after fruit set [61]. Besides, we found a series of SNPs in the *VviAGL6* gene sequence significantly associated with coulure and berry number (Table 2). As observed in diverse angiosperms (Arabidopsis, maize, petunia, rice), *AGL6* genes are involved in floral meristem regulation, floral organogenesis, ovule and seed development, and male and female gametophytes development [72].

Two *VviAP1* MADS-box genes (Vitvi01g00008 and Vitvi14g01341) were found to be associated with the number of flowers per inflorescence and the number of seeds per berry, respectively (Table 2). In Arabidopsis and other flowering plants, *AP1* plays a central role in the transition from the inflorescence meristem into floral meristem, reducing the time needed for flowering [73]. Expression profiles suggest that *VviAP1* conserves a key role in flowering induction and flower morphogenesis in grapevine too [66,68], supporting the association results obtained in this work. Besides, *AtSHP1* and *AtSHP2* genes are known to be required for fruit dehiscence in Arabidopsis, promoting the lignification of valve margin cells that induce silique shatter [74]. Here, we found a series of SNPs in the *VviSHP2* gene sequence associated with coulure and fruit set (Table 2). Thus, our results indicate that the proposed role of this gene in Arabidopsis might be conserved in the grapevine, playing some role in the mechanisms involved in the shedding of unfertilized flowers or very young berries that ultimately affect fruit set rate.

A recent genome-wide analysis of the MADS-box transcription factor family in grapevine indicated a notable expansion of the *SVP* subfamily compared to other woody species, with a total of 10 *VviSVP/VviSVPS* (for *SVP Short*) genes detected across the grapevine genome [44]. Our results revealed significant associations between *VviSVP1* and millerandage, *VviSVP2* and coulure and fruit set, *VviSVP3* and flower number, and *VviSVP4* and coulure (Table 2), suggesting that this subfamily plays a relevant role in the fruit production of this crop. In Arabidopsis, *AtSVP* acts as a repressor of flowering, with an opposite effect to that of *AtAGL24* [75]. Thus, it has been indicated that floral transition takes place when the expression of *AtSVP* reduces and that of *AtAGL24* increases. Similar results in barley suggest that *SVP* genes are part of a conserved mechanism that regulates floral meristem identity across plant species [76]. *VviSVP/VviSVPS* genes show high similarity with the tomato *JOINTLESS* (*SlJ*) gene [44], a MADS-box gene involved in the development of the pedicel abscission zone that controls flower and fruit abscission [77,78]. Results in apple and pear suggest that *J* homologs play similar roles in the abscission of flowers and fruits in these two crops [79,80]. Interestingly, a phylogenetic analysis of the two apple *J* homologs (*MdJa* and *MdJb*) with *SlJ*, *AtSVP*, *AtAGL24* and other *SVP* homologs revealed a high similitude between *AtSVP*, *MdJa*, *MdJb, SlJ,* and the grapevine *VvSVP1* [79], implying a potential similar function for *J* in the grapevine that would support the associations obtained in our work.

The R2R3-MYB transcription factor subfamily regulates numerous developmental processes at the whole plant and cell-specific levels, and

modulate responses to environmental stresses [81]. In grapevine, this subfamily includes 134 genes, many of them involved in flower and seed development or in the synthesis of anthocyanins and flavonoids [82]. Our results revealed significant associations between SNPs in two grapevine R2R3-MYB transcription factors (*VviMYB68* and *VviMYB108b*) and three fruit set-related traits (Table 2). On the one hand, *VviMYB68* was found to associate with coulure and fruit set. An overexpression of *AtMYB68* produces accumulation of lignin in Arabidopsis roots, suggesting that this transcription factor might participate (directly or indirectly) in the regulation of the lignin biosynthesis pathway [81]. In grapevine, the accumulation of lignin in berry pedicels promotes its drop [83], which might explain the results obtained for *VviMYB68* in our work. On the other hand, we detected one SNP in the *VviMYB108b* (also known as *VviMYB78*) gene sequence significantly associated with berry number. Among other phenotypic alterations, Arabidopsis *AtMYB108* mutant lines show low pollen viability and short anther filaments [84], suggesting that *AtMYB108* regulates late stages of stamen development and controls male fertility. In grapevine, a reduced male fertility has been related to poor fruit set, which reduces the number of berries in the cluster and compromises crop yield [7].

Together with *LEAFY* (*LFY*), *TERMINAL FLOWER 1* (*TFL1*) controls Arabidopsis inflorescence architecture [85]. Misexpression of the main *TFL1* homolog in grapevine (*VviTFL1A*) delays the time of anthesis and modifies the branching architecture of the inflorescence, as observed in a somatic variant of the grapevine cultivar Carignan [86]. Further work identified a series of *VviTFL1A* polymorphisms that associate with the phenotypic diversity observed for flowering time, berry weight and cluster width in a core collection of 140 grapevine cultivars [87], suggesting a role of *VviTFL1A* on phenology and cluster traits, and thus a conservation of the role reported in Arabidopsis. Similarly, a previous work linked some polymorphisms in the *VviNAC26* gene sequence (the closest homolog to Arabidopsis *AtNAP* or *AtNAC029* [88]) with berry dimensions and cluster weight variability [34]. Here, we detected one *VviTFL1A* (Vitvi17g00229) SNP associated with millerandage, and two SNPs in the *VviNAC26* gene sequence (Vitvi01g01038) associated with the number of berries in the cluster (Table 2), suggesting additional effects for these two genes in grapevine fruit production. Lastly, we found some SNPs in a gene that codes for a protein of unknown function (Vitvi18g02631) that associates with coulure, millerandage, and fruit set (Table 2). The lack of information on this gene hinders to hypothesize about its biological function in grapevine, but expression analyses in cv. Corvina indicates that it is differentially expressed in flowers at full-flowering, more specifically in flower stamens [89]. This information suggests that Vitvi18g02631 might be involved in male fertility, which would support the association results observed in this work.

Our results support the polygenic nature of fruit set-related traits in grapevine, being under the control of numerous QTLs on different chromosomes. This challenging and complex genetic basis would require pyramiding multiple beneficial QTL alleles in a single grapevine genotype to improve grape production through additive complementary mechanisms, as previously shown for major crops like rice [90] and wheat [91]. Nevertheless, this approach would benefit from more detailed information regarding the interaction between the causing QTLs prior to their pyramiding in a single genotype, and it will always depend on the aims of the breeding program. Alternatively, in the cases where the associated SNPs were confirmed as the cause of the phenotypic variation, the information provided by this work could be used for modifying these specific genetic loci through cutting-edge genome editing technologies, such as the CRISPR/Cas9 system, which has shown promising results in grapevine [92,93]. As an example, low-yielding cultivars with a reduced number of initial flowers could be benefited from the incorporation of alleles that increase fruit set to ensure an adequate yield. On the contrary, highly-productive cultivars with a high number of flowers per inflorescence could be improved through the incorporation of alleles associated with reduced fruit set rates, to reduce the final number of berries in the cluster that, in turn, will reduce disease pressure due to a less compact cluster architecture. In any case, our results are very useful to detect some genes with potential roles on the biological mechanisms determining fruit set and fruit set-related traits. Further functional analyses will be required to confirm their putative biological function in grapevine.

5. Conclusion

A major challenge in viticulture, given the predicted impacts of climate change on grape formation and growth, is to optimize vineyard production whilst maintaining grape quality. Understanding the genetic and molecular mechanisms responsible for the determination of grapevine yield components can be relevant to develop long-term strategies in that direction. The results reported in this work aid to understand the complex genetic basis of fruit set and other fruit set-related traits, providing a series of candidate genes and alleles likely involved in their genetic architecture. Among the genes tested in this work, we found a relevant role of a series of MADS-box transcription factors, which in Arabidopsis and other crops were found to be involved in flowering and fruit set processes. Among them, some members of the *VviAGL*, *VviAP*, *VviSHP* and *VviSVP* gene subfamilies stood out as firm candidates to understand their DNA sequence diversity and test their biological effect in fruit set through specific functional analyses. In addition, the detection of alleles that increase or reduce fruit set performance opens the door to improve grapevine cultivars with better grape production to meet human needs and to cope with future environment conditions.

Author contributions

Conceptualization, LHZ, JI and JT; Methodology, formal analysis and data interpretation, LHZ, JI and JT; Phenotypic trials, LHZ, EB and SI; NGS data management, RTP, JG and PC-B; Manuscript drafting, LHZ; Manuscript review and editing, LHZ, JT and JI; Manuscript critical discussion, LHZ, RTP, JG, PC-B, JMM-Z, JI and JT; Funding acquisition, JI, JG and JMM-Z.

Funding sources

This work was supported by the Spanish Ministerio de Economía y Competitividad (MINECO) [project AGL2014−59171-R] and by the Agencia Estatal de Investigación (AEI) [project BIO2017-86375-R], both co-funded by FEDER (UE). JT is funded by a Juan de la Cierva-Incorporación grant [IJC2018-035036-I]. LHZ was partially funded by a STSM from the COST Action CA17111 INTEGRAPE, supported by COST (European Cooperation in Science and Technology).

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

Authors acknowledge T. Flutre for his assistance with the calculation of BLUP values, M. Angulo, S. Hernáiz and M.I. Montemayor for their technical assistance and J.L. Pérez-Sotés and the Servicio de Investigación Agraria y Sanidad Vegetal (Gobierno de La Rioja) for plant material maintenance.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.plantsci.2021.110875.

References

[1] OIV, Statistical Report on World Vitiviniculture, International Organisation of Vine and Wine, 2019, 2019.
[2] B.I. Reisch, C.L. Owens, P.S. Cousins, Grape, in: M.L. Badenes, D.H. Byrne (Eds.), Fruit Breeding, Handbook of Plant Breeding, Springer, New York, 2012, pp. 225–262.
[3] P. This, T. Lacombe, M.R. Thomas, Historical origins and genetic diversity of wine grapes, Trends Genet. 22 (2006) 511–519.
[4] M.J. Carmona, J. Chaib, J.M. Martínez-Zapater, M.R. Thomas, A molecular genetic perspective of reproductive development in grapevine, J. Exp. Bot. 59 (2008) 2579–2596.
[5] P.E. McGovern, D.L. Glusker, L.J. Exner, M.M. Volgt, Neolithic resinated wine, Nature 381 (1996), 480-480.
[6] J.F. Terral, E. Tabard, L. Bouby, S. Ivorra, T. Pastor, I. Figueiral, S. Picq, J. B. Chevance, C. Jung, L. Fabre, C. Tardy, M. Compan, R. Bacilieri, T. Lacombe, P. This, Evolution and history of grapevine (*Vitis vinifera*) under domestication: new morphometric perspectives to understand seed domestication syndrome and reveal origins of ancient European cultivars, Ann. Bot. 105 (2010) 443–455.
[7] J. Tello, M.I. Montemayor, A. Forneck, J. Ibáñez, A new image-based tool for the high throughput phenotyping of pollen viability: evaluation of inter- and intra-cultivar diversity in grapevine, Plant Methods 9 (14) (2018) 3.
[8] Z. Migicovsky, J. Sawler, K.M. Gardner, M.K. Aradhya, B.H. Prins, H. R. Schwaninger, C.D. Bustamante, E.S. Buckler, G.-Y. Zhong, P.J. Brown, S. Myles, Patterns of genomic and phenomic diversity in wine and table grapes, Hortic. Res. 4 (2017), 17035.
[9] C. Srinivasan, M.G. Mullins, Phisiology of flowering in the grapevine - a review, Am. J. Enol. Vitic. 32 (1981) 47–63.
[10] P.R. Petrie, P.R. Clingeleffer, Effects of temperature and light (before and after budburst) on inflorescence morphology and flower number of Chardonnay grapevines (*Vitis vinifera* L.), Aust. J. Grape Wine Res. 11 (2005) 59–65.
[11] P.R. Dry, M.L. Longbottom, S. McLoughlin, T.E. Johnson, C. Collins, Classification of reproductive performance of ten winegrape varieties, Aust. J. Grape Wine Res. 16 (2010) 47–55.
[12] R. Bessis, J.C. Fournioux, Zone d'abscission et coulure de la vigne, Vitis 31 (1992) 9–21.
[13] J. Ibáñez, E. Baroja, J. Grimplet, S. Ibáñez, Cultivated grapevine displays a great diversity for reproductive performance variables, Crop Breed. Genet. Genomics 2 (2020) e200003.
[14] A.P. Friend, M.C.T. Trought, Delayed winter spur-pruning in New Zealand can alter yield components of Merlot grapevines, Aust. J. Grape Wine Res. 13 (2007) 157–164.
[15] M.L. Longbottom, P.R. Dry, M. Sedgley, Effects of sodium molybdate foliar sprays on molybdenum concentration in the vegetative and reproductive structures and on yield components of *Vitis vinifera* cv. Merlot, Aust. J. Grape Wine Res. 16 (2010) 477–490.
[16] C. Collins, P.R. Dry, Response of fruitset and other yield components to shoot topping and 2-chlorethyltrimethyl-ammonium chloride application, Aust. J. Grape Wine Res. 15 (2009) 256–267.
[17] P. May, Flowering and Fruitset in Grapevines, Lythrum Press, Adelaide, South Australia, 2004.
[18] L. Lu, J. Liang, X. Chang, H. Yang, T. Li, J. Hu, Enhanced vacuolar invertase activity and capability for carbohydrate import in GA-treated inflorescence correlate with increased fruit set in grapevine, Tree Genet. Genomes 13 (2017) 21.
[19] T. Baby, C. Collins, S.D. Tyerman, M. Gilliham, Salinity negatively affects pollen tube growth and fruit set in grapevines and cannot be ameliorated by silicon, Am. J. Enol. Vitic. 67 (2016) 218–228.
[20] E. Duchene, C. Schneider, J.P. Gaudillere, Effects of nitrogen nutrition timing on fruit set of grapevine, cv, Grenache, Vitis 40 (2001) 45–46.
[21] R.P. Smithyman, G.S. Howell, D.P. Miller, The use of competition for carbohydrates among vegetative and reproductive sinks to reduce fruit set and Botrytis bunch rot in Seyval blanc grapevines, Am. J. Enol. Vitic. 49 (1998) 163–170.
[22] D. Acimovic, L. Tozzini, A. Green, P. Sivilotti, P. Sabbatini, Identification of a defoliation severity threshold for changing fruitset, bunch morphology and fruit composition in Pinot Noir, Aust. J. Grape Wine Res. 22 (2016) 399–408.
[23] K. Padmalatha, H. Weksler, A. Mugzach, A.K. Acheampong, C. Zheng, T. Halaly-Basha, E. Or, ABA application during flowering and fruit set reduces berry number and improves cluster uniformity, Am. J. Enol. Vitic. 68 (2017) 275–282.
[24] H.W. Caspari, A. Lang, P. Alspach, Effects of girdling and leaf removal on fruit set and vegetative growth in grape, Am. J. Enol. Vitic. 49 (1998) 359–366.
[25] A.P. Viana, S. Riaz, M.A. Walker, Genetic dissection of agronomic traits within a segregating population of breeding table grapes, Genet. Mol. Res. 12 (2013) 951–964.
[26] G. Fanizza, F. Lamaj, L. Costantini, R. Chaabane, M.S. Grando, QTL analysis for fruit yield components in table grapes (*Vitis vinifera*), Theor. Appl. Genet. 111 (2005) 658–664.
[27] J. Correa, M. Mamani, C. Muñoz-Espinoza, D. Laborie, C. Muñoz, M. Pinto, P. Hinrichsen, Heritability and identification of QTLs and underlying candidate genes associated with the architecture of the grapevine cluster (*Vitis vinifera* L.), Theor. Appl. Genet. 127 (2014) 1143–1162.
[28] R. Richter, D. Gabriel, F. Rist, R. Topfer, E. Zyprian, Identification of co-located QTLs and genomic regions affecting grapevine cluster architecture, Theor. Appl. Genet. 132 (2019) 1159–1177.
[29] J. Tello, R. Torres-Pérez, J. Grimplet, J. Ibáñez, Association analysis of grapevine bunch traits using a comprehensive approach, Theor. Appl. Genet. 129 (2016) 227–242.
[30] J. Tello, R. Aguirrezábal, S. Hernaiz, B. Larreina, M.I. Montemayor, E. Vaquero, J. Ibáñez, Multicultivar and multivariate study of the natural variation for grapevine bunch compactness, Aust. J. Grape Wine Res. 21 (2015) 277–289.
[31] B.G. Coombe, Adoption of a system for identifying grapevine growth stages, Aust. J. Grape Wine Res. 1 (1995) 104–110.
[32] D. Bates, M. Maechler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4, J. Stat. Softw. 67 (2015) 1–48.
[33] J. Tello, R. Torres-Pérez, T. Flutre, J. Grimplet, J. Ibáñez, *VviUCC1* nucleotide diversity, linkage disequilibrium and association with rachis architecture traits in grapevine, Genes 11 (2020) 598.
[34] J. Tello, R. Torres-Pérez, J. Grimplet, P. Carbonell-Bejerano, J.M. Martínez-Zapater, J. Ibáñez, Polymorphisms and minihaplotypes in the *VvNAC26* gene associate with berry size variation in grapevine, BMC Plant Biol. 15 (2015) 253.
[35] G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, Mol. Ecol. 14 (2005) 2611–2620.
[36] D. Earl, B.M. vonHoldt, STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method, Conserv. Genet. Resour. 4 (2012) 359–361.
[37] Y. Vigouroux, J.C. Glaubitz, Y. Matsuoka, M.-M. Goodman, J. Sánchez, J. Doebley, Population structure and genetic diversity of New World maize races assessed by DNA microsatellites, Am. J. Bot. 95 (2008) 1240–1253.
[38] J. Wang, An estimator for pairwise relatedness using molecular markers, Genetics 160 (2002) 1203–1215.
[39] J. Pew, P.H. Muir, J. Wang, T.R. Frasier, Related: an R package for analysing pairwise relatedness from codominant molecular markers, Mol. Ecol. Resour. 15 (2015) 557–561.
[40] C. Royo, R. Torres-Perez, N. Mauri, N. Diestro, J.A. Cabezas, C. Marchal, T. Lacombe, J. Ibáñez, M. Tornel, J. Carreno, J.M. Martínez-Zapater, P. Carbonell-Bejerano, The major origin of seedless grapes is associated with a missense mutation in the MADS-box gene *VviAGL11*, Plant Physiol. 177 (2018) 1234–1253.
[41] C. Royo, P. Carbonell-Bejerano, R. Torres-Pérez, A. Nebish, O. Martinez, M. Rey, R. Aroutiounian, J. Ibáñez, J.M. Martínez-Zapater, Developmental, transcriptome, and genetic alterations associated with parthenocarpy in the grapevine seedless somatic variant Corinto bianco, J. Exp. Bot. 67 (2016) 259–273.
[42] J.M. Martínez-Zapater, P. Carbonell-Bejerano, C. Royo, R. Torres-Pérez, N. Diestro, J. Grimplet, N. Mauri, J. Ibáñez, Genetic variation for grapevine reproductive development, Acta Hort. 1248 (2019) 319–326.
[43] J. Diaz-Riquelme, D. Lijavetzky, J.M. Martínez-Zapater, M.J. Carmona, Genome-wide analysis of MIKCC-Type MADS Box genes in grapevine, Plant Physiol. 149 (2009) 354–369.
[44] J. Grimplet, J.M. Martínez-Zapater, M.J. Carmona, Structural and functional annotation of the MADS-box transcription factor family in grapevine, BMC Genomics 17 (2016) 80.
[45] A. Canaguier, J. Grimplet, G. Di Gaspero, S. Scalabrin, E. Duchene, N. Choisne, N. Mohellibi, C. Guichard, S. Rombauts, I. Le Clainche, A. Berard, A. Chauvear, R. Bounon, C. Rustenholz, M. Morgante, M.-C. Le Paslier, D. Brunel, A. Adam-Blondon, A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3), Genom. Data 14 (2017) 56–62.
[46] B. Langmead, S.L. Salzberg, Fast gapped-read alingment with Bowtie 2, Nat. Methods 9 (2012) 357–359.
[47] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, a.G.P.D.P. Subgroup, The sequence alignment/map format and SAMtools, Bioinformatics 25 (2009) 2078–2079.
[48] K.W. Broman, H. Wu, S. Sen, G.A. Chuchill, R/qtl: QTL mapping in experimental crosses, Bioinformatics 19 (2003) 889–890.
[49] P.J. Bradbury, Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, E.S. Buckler, TASSEL: software for association mapping of complex traits in diverse samples, Bioinformatics 23 (2007) 2633–2635.
[50] P. Duggal, E.M. Gillanders, T.N. Holmes, J.E. Bailey-Wilson, Establishing an adjusted p-value threshold to control the family-wide type I error in genome wide association studies, BMC Genomics 9 (2008) 516.
[51] J.R. Conway, A. Lex, N. Gehlenborg, UpSetR: an R package for the visualization of intersecting sets and their properties, Bioinformatics 33 (2017) 2938–2940.
[52] J. Su, C. Pang, H. Wei, L. Li, B. Liang, C. Wang, M. Song, H. Wang, S. Zhao, X. Jia, G. Mao, L. Huang, D. Geng, C. Wang, S. Fan, S. Yu, Identification of favorable SNP alleles and candidate genes for traits related to early maturity via GWAS in upland cotton, BMC Genomics 17 (2016) 687.
[53] R. Bacilieri, T. Lacombe, L. Le Cunff, M. Di Vecchi-Staraz, V. Laucou, B. Genna, J. P. Péros, P. This, J.-M. Boursiquot, Genetic structure in cultivated grapevine is linked to geography and human selection, BMC Plant Biol. 13 (2013) 25.
[54] F. Emanuelli, S. Lorenzi, L. Grzeskowiak, V. Catalano, M. Stefanini, M. Troggio, S. Myles, J.M. Martínez-Zapater, E. Zyprian, F.M. Moreira, M.S. Grando, Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape, BMC Plant Biol. 13 (2013) 39.
[55] V. Laucou, A. Launay, R. Bacilieri, T. Lacombe, A.F. Adam-Blondon, A. Bérard, A. Chauveau, M.T. de Andrés, L. Hausmann, J. Ibáñez, M.-C. Le Paslier, D. Maghradze, J.M. Martínez-Zapater, E. Maul, M. Ponnaiah, R. Töpfer, J.P. Péros, J.M. Boursiquot, Extended diversity analysis of cultivated grapevine Vitis vinifera with 10K genome-wide SNPs, PLoS One 13 (2018) e0192540.
[56] M. Keller, Managing grapevines to optimise fruit development in a challenging environment: a climate change primer for viticulturists, Aust. J. Grape Wine Res. 16 (2010) 56–69.

[57] S. Vezzulli, A. Doligez, D. Bellin, Molecular mapping of grapevine genes, in: D. Cantu, M.A. Walker (Eds.), The Grape Genome, Springer International Publishing, 2019, pp. 103–136.
[58] R.J. Chandrababu, R.K. Sharma, Heritability estimates in almond [*Prunus dulcis* (Miller) D.A. Webb], Sci. Hort. 79 (1999) 237–243.
[59] D. Milatovic, D. Nikolic, D. Durovic, Variability, heritability and correlations of some factors affecting productivity in peach, Hort. Sci. (Prague) 37 (2010) 79–87.
[60] D. Molitor, L. Hoffman, M. Beyer, Flower debris removal delays grape bunch rot epidemic, Am. J. Enol. Vitic. 66 (2015) 548–553.
[61] N. Kühn, C. Abello, F. Godoy, S. Delrot, P. Arce-Johnson, Differential behavior within a grapevine cluster: decreased ethylene-related gene expression dependent on auxin transport is correlated with low abscission of first developed berries, PLoS One 9 (2014) e111258.
[62] A. Ebadi, M. Sedgley, P. May, B.G. Coombe, Seed development and abortion in *Vitis vinifera* L., cv. Chardonnay, Int. J. Plant Sci. 157 (1996) 703–712.
[63] M. Ng, M.F. Yanofsky, Function and evolution of the plant MADS-box gene family, Nat. Rev. Genet. 2 (2001) 186–195.
[64] K. Kaufmann, R. Melzer, G. Theissen, MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants, Gene 347 (2005) 183–198.
[65] M.J. Poupin, F. Federici, C. Medina, J.T. Matus, T. Timmerman, P. Arce-Johnson, Isolation of the three grape sub-lineages of B-class MADS-box *TM6*, *PISTILLATA* and *APETALA3* genes which are differentially expressed during flower and fruit development, Gene 404 (2007) 10–24.
[66] M. Calonje, P. Cubas, J.M. Martínez-Zapater, M.J. Carmona, Floral meristem identity genes are expressed during tendril development in grapevine, Plant Physiol. 135 (2004) 1491–1501.
[67] P.K. Boss, E. Sensi, C. Hua, C. Davies, M.R. Thomas, Cloning and characterisation of grapevine (*Vitis vinifera* L.) MADS-box genes expressed during inflorescence and berry development, Plant Sci. 162 (2002) 887–895.
[68] J. Diaz-Riquelme, J.M. Martínez-Zapater, M.J. Carmona, Transcriptional analysis of tendril and inflorescence development in grapevine (*Vitis vinifera* L.), PLoS One 9 (2014) e92339.
[69] F. Palumbo, A. Vannozzi, G. Magon, M. Lucchin, G. Barcaccia, Genomics of flower identity in grapevine (*Vitis vinifera* L.), Front. Plant Sci. 10 (2019) 316.
[70] L.-H. Yu, Z.-Q. Miao, G.-F. Qi, J. Wu, X.-T. Cai, J.-L. Mao, C.-B. Xiang, MADS-box transcription factor AGL21 regulates lateral root development and responds to multiple external and physiological signals, Mol. Plant 7 (2014) 1653–1669.
[71] R.- Tapia-López, B. García-Ponce, J.G. Dubrovsky, A. Garay-Arroyo, R.V. Pérez-Ruíz, S.-H. Kim, F. Acevedo, S. Pelaz, E.R. Alvarez-Buylla, An *AGAMOUS*-related MADS-box gene, *XAL1* (*AGL12*), regulates root meristem cell proliferation and flowering transition in Arabidopsis, Plant Physiol. 146 (2008) 1182–1192.
[72] L. Dreni, D. Zhang, Flower development: the evolutionary history and functions of the *AGL6* subfamily MADS-box genes, J. Exp. Bot. 67 (2016) 1625–1638.
[73] M.A. Mandel, M.F. Yanofsky, A gene triggering flower formation in *Arabidopsis*, Nature 377 (1995) 522–524.
[74] S.J. Lijegren, G.S. Ditta, Y. Eshed, B. Savidge, J.L. Bowman, M.F. Yanofsky, *SHATTERPROOF* MADS-box genes control seed dispersal in Arabidopsis, Nature 404 (2000) 766–770.
[75] V. Gregis, A. Sessa, L. Colombo, M.M. Kater, *AGL24*, *SHORT VEGETATIVE PHASE*, and *APETALA1* redundantly control *AGAMOUS* during early stages of flower development in Arabidopsis, Plant Cell 18 (2006) 1373–1382.
[76] B. Trevaskis, M. Tadege, M.N. Hemming, J. Peacock, E.S. Dennis, C. Sheldon, *Short Vegetative Phase*-Like MADS-Box genes inhibit floral meristem identity in barley, Plant Physiol. 143 (2007) 225–235.
[77] D. Liu, D. Wang, Z. Qin, D. Zhang, L. Yin, L. Wu, J. Colasanti, A. Li, L. Mao, The SEPALLATA MADS-box protein SLMBP21 forms protein complexes with JOINTLESS and MACROCALYX as a transcription activator for development of the tomato flower abscission zone, Plant J. 7 (2014) 284–296.
[78] L. Mao, D. Begum, H.-W. Chuang, M.A. Budiman, E.J. Szymkowlak, E.E. Irish, R. A. Wing, JOINTLESS is a MADS-box gene controlling tomato flower abscission zone development, Nature 406 (2000) 910–913.
[79] T. Nakano, H. Kato, Y. Shima, Y. Ito, Apple *SVP* family MADS-box proteins and the tomato pedicel abscission zone regulator *JOINTLESS* have similar molecular activities, Plant Cell Physiol. 56 (2015) 1097–1106.
[80] X. Qi, S. Hu, H. Zhou, H. Liu, X. Liu, L. Wang, B. Zhao, X. Huang, S. Zhang, A MADS-box transcription factor of 'Kuerlexiangli' (*Pyrus sinkiangensis Yu*) *PsJOINTLESS* gene functions in floral organ abscission, Gene 642 (2018) 163–171.
[81] C. Feng, E. Andreasson, A. Maslak, H.P. Mock, O. Mattson, J. Mundy, *Arabidopsis MYB68* in development and responses to environmental cues, Plant Sci. 167 (2004) 1099–1107.
[82] D.C.J. Wong, R. Schlechter, A. Vannozzi, J. Höll, I. Hmmam, J. Bogs, G.B. Tornielli, S.D. Castellarin, J.T. Matus, A systems-oriented analysis of the grapevine R2R3-MYB transcription factor family uncovers new insights into the regulation of stilbene accumulation, DNA Res. 23 (2016) 451–466.
[83] M. García-Rojas, M. Meneses, K. Oviedo, C. Carrasco, B. Defilippi, M. González-Agüero, G. León, P. Hinrichsen, Exogenous gibberellic acid application induces the overexpression of key genes for pedicel lignification and an increase in berry drop in table grape, Plant Physiol. Biochem. 126 (2018) 32–38.
[84] A. Mandaokar, J. Browse, MYB108 acts together with MYB24 to regulate jasmonate-mediated stamen maturation in arabidopsis, Plant Physiol. 149 (2008) 851–862.
[85] P. Prusinkiewicz, Y. Erasmus, B. Lane, L.D. Harder, E. Coen, Evolution and development of inflorescence architectures, Science 316 (2007) 1452–1456.
[86] L. Fernandez, L. Torregrosa, V. Segura, A. Bouquet, J.M. Martínez-Zapater, Transposon-induced gene activation as a mechanism generating cluster shape somatic variation in grapevine, Plant J. 61 (2010) 545–557.
[87] L. Fernandez, L. Le Cunff, J. Tello, T. Lacombe, J.M. Boursiquot, A. Fournier-Level, G. Bravo, S. Lalet, L. Torregrosa, P. This, J.M. Martínez-Zapater, Haplotype diversity of *VvTFL1A* gene and association with cluster traits in grapevine (*V.* vinifera), BMC Plant Biol. 14 (2014) 209.
[88] A. Cenci, V. Guignon, N. Roux, M. Rouard, Genomic analysis of NAC transcription factors in banana (*Musa acuminata*) and definition of NAC orthologous groups for monocots and dicots, Plant Mol. Biol. 85 (2014) 63–80.
[89] M. Fasoli, S. Dal Santo, S. Zenoni, G.B. Tornielli, L. Farina, A. Zamboni, A. Porceddu, L. Venturini, M. Bicego, V. Murino, A. Ferrarini, M. Delledonne, M. Pezzotti, The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program, Plant Cell 24 (2012) 3489–3505.
[90] P. Wang, Y. Xing, Z. Li, S. Yu, Improving rice yield and quality by QTL pyramiding, Mol. Breed. 29 (2012) 903–913.
[91] S. Tyagi, R.R. Mir, H. Kaur, P. Chhuneja, B. Ramesh, H.S. Balyan, P.K. Gupta, Marker-assisted pyramiding of eight QTLs/genes for seven different traits in common wheat (*Triticum aestivum* L.), Mol. Breed. 34 (2014) 167–175.
[92] C. Ren, Y. Guo, J. Kong, F. Lecourieux, Z. Dai, S. Li, Z. Liang, Knockout of *VvCCD8* gene in grapevine affects shoot branching, BMC Plant Biol. 20 (2020) 47.
[93] M.-Y. Li, Y.-T. Jiao, Y.-T. Wang, N. Zhang, B.-B. Wang, R.-Q. Liu, X. Yin, Y. Xu, G.-T. Liu, CRISPR/Cas9-mediated *VvPR4b* editing decreases downy mildew resistance in grapevine (*Vitis vinifera* L.), Hortic. Res. 7 (2020) 149.