Genome-wide association study in a diverse grapevine collection provides insights into the genetic basis of berry size and cluster architecture traits

Geovani Luciano de Oliveira[1,2], Felipe Roberto Francisco[3], Yohans Alves de Moura[1,2], Guilherme Francio Niederauer[1,2], Roberto Fritsche-Neto[4], Anete Pereira de Souza[1,2,5], Mara Fernandes Moura Furlan[3*]

[1] Molecular Biology and Genetic Engineering Center (CBMEG), Universidade Estadual de Campinas (UNICAMP), Campinas, Brazil

[2] Center for Plant Molecular Breeding (CeM²P), Universidade Estadual de Campinas (UNICAMP), Campinas, SP, Brazil

[3] Advanced Fruit Research Center, Instituto Agronômico (IAC), Jundiaí, SP, Brazil

[4] Horticulture Sciences, College of Agriculture and Life Sciences, North Carolina State University, Raleigh, NC, 27695-7609, Campus Box 7609, USA

[5] Department of Plant Biology, Biology Institute, Universidade Estadual de Campinas (UNICAMP), Campinas, Brazil

* Corresponding author

E-mail: mara.moura@sp.gov.br (MFMF)

Abstract

Berry and cluster size are pivotal determinants of grapevine productivity and consumer preferences and remain major targets in grapevine breeding. However, given their complexity as quantitative traits under polygenic control, a deeper understanding of their genetic determinants is needed. The gene pool of the Brazilian grapevine has made a significant contribution to enhancing grapevine performance in tropical and subtropical regions. In this study, we conducted a genome-wide association study (GWAS) using a diverse panel of 288 *Vitis* spp. accessions from the Instituto Agronômico Germplasm Bank, Brazil. This panel was phenotyped for six cluster architecture traits over 12 years and genotyped using the Vitis18kSNP array. Using two different algorithms, the GWAS identified 56 significant SNPs distributed across 17 chromosomes, validating previously identified quantitative trait loci (QTLs) and revealing novel associations. Four closely spaced markers on Chr1 suggest the presence of a QTL influencing five traits simultaneously. A strong association signal, with phenotypic variance explained (PVE) values of approximately 29–35%, indicated a major QTL for berry length (BL) and width (BWi) on Chr14. Additionally, major-effect SNP loci were identified for cluster weight (CW) on Chr1, cluster length (CL) on Chr7 and 14, cluster width (CWi) on Chr6 and 18, and berry weight (BW) on Chr4, with PVE values ranging from 18–27%. Furthermore, 80 genes associated with berry traits and 52 genes associated with cluster traits were identified as putative candidate genes in the genomic regions associated with significant SNPs. These candidate genes are involved in the regulation of growth and development, hormone regulation, protein synthesis, stress response, and other physiological processes essential for cell health and functionality. Our results provide valuable insights into the genetic determinants of grape berry size and cluster architecture, offering critical data to support future functional studies and enhance the efficiency of related breeding programs.

Keywords: quantitative trait loci, subtropical viticulture, berry size, GWAS, genetic resources, grapevine breeding

Introduction

Grapevine (*Vitis* spp.) is a key fruit crop worldwide, with berries consumed fresh or used in the production of wine, juice, jellies, and other food products [1,2]. Each grape product demands specific berry traits, such as color, composition, taste, and size, to meet diverse market and industry requirements [3]. Over centuries of domestication and breeding, grapes have accumulated significant morphological and physiological diversity, resulting in thousands of varieties tailored for specific purposes. The variations in these traits hold considerable potential as valuable resources for grape breeding [4,5].

Grapevine, as a woody perennial fruit crop, presents challenges associated with its substantial size and prolonged juvenile phase, which make plant cultivation efforts and evaluation in breeding programs time-consuming and costly [6]. This issue is particularly pertinent because the assessment of fruit characteristics can only commence once the plant reaches maturity, typically in the third to fifth year of its life cycle [7]. To address this, early genotype selection technologies, including the use of genomic tools for marker-assisted selection (MAS), have been suggested to accelerate and improve the plant breeding process. These advances enable breeders to identify and select individuals with desired genotypes at an early stage, facilitating more efficient and targeted breeding programs. This promises a significant reduction in the breeding cycle by up to a decade, along with notable cost savings ranging from 16–34% [8,9]. Identifying phenotype-genotype associations is crucial for plant breeding. However, several fundamental issues, especially regarding complex traits, still need to be resolved before MAS can fully realize its potential in practical breeding programs [10]. Fruit yield and quality are the most commonly measured characteristics in grapevines but are poorly understood.

These traits are quantitatively extremely complex due to the number of segregating loci that control all the components involved in productivity, as well as the influence of nongenetic factors [11–13].

Berry size is a crucial agronomic trait that strongly influences both quality and yield in grapes. This feature is essential for table grapes, as it is critical for consumer preferences and acceptance and directly influences market value [14,15]. Berry size also plays a key role in determining juice composition and affects the overall quality of wine [16–19]. Cluster size is another essential factor influencing fruit quality and production. Together with berry size, cluster size is vital for defining cluster architecture and significantly affects disease pressure, sun exposure, and uniform ripening [20,21]. Given the importance of these factors for breeding, numerous studies have utilized quantitative trait locus (QTL) mapping in biparental progenies to elucidate the genetic architecture underlying key berry size and cluster architecture traits in grapes and identify QTLs on various chromosomes [15,22–30]. Notably, most major QTLs related to berry size have been found on Chr18, closely linked to the well-characterized SDI (*seed development inhibitor*) locus [31–34].

Linkage mapping has significantly advanced our understanding of important berry and cluster-related traits in grapes. However, the results have shown variability across different genotypes and environments, underscoring the complexity of the genetic control of these traits [35,36]. Moreover, many of the biparental progenies analyzed in these studies were derived exclusively from *V. vinifera* genotypes, which may have produced inconclusive results and an underestimation of the genetic complexity of the crop. This approach also has inherent limitations, including restricted allelic diversity in parental lines, low recombination frequency in offspring, underestimation of polygenic contributions for predictive modeling, and reduced applicability across broader genetic backgrounds [37].

To overcome these constraints, genome-wide association studies (GWASs) offer an alternative approach in which genetic architecture and gene interactions influencing traits are investigated without the need for a specifically designed mapping population [38–41]. When the

genotypes in a germplasm collection exhibit high genetic diversity, GWASs can exploit historical and evolutionary recombination events that have accumulated during domestication and breeding to associate markers with phenotypes of interest. Provided that the marker density is sufficient, the recombination history captured in these panels allows linkage disequilibrium (LD) to be explored at a finer scale and improves the precision of QTL detection [13,42,43]. In recent years, several GWAS investigations have focused on traits related to berry and cluster traits in grapes [37,39,40,44–46]. However, the genetic foundations of many important traits, such as berry width and length, as well as cluster weight, width, and length, remain underexplored. The extensive diversity within cultivated grapevines requires further investigation. Intra- and interspecific crosses conducted throughout breeding cycles to introduce novelty, resistance, adaptability, and hybrid vigor have resulted in hybrids with heterogeneous genetic backgrounds [38,47]. Analyzing a more diverse germplasm in additional association studies may uncover new sources of variation for previously characterized traits [48]. Moreover, integrating genomic technologies with traditional breeding methods could significantly increase the ability to address current and future environmental challenges. Under these circumstances, the conservation of germplasms across plant species and their utilization, leveraging complementary approaches such as genetic and genomic resources, are vital for achieving rapid advances in productivity [49].

One of the main objectives of grapevine breeding programs in Brazil is to develop cultivars that combine adaptations to tropical and subtropical regions with fruit quality, productivity, and resistance to biotic stresses [50]. To support this goal, the Instituto Agronômico (IAC) has, over several years, established a diverse germplasm collection that includes key *V. vinifera* cultivars, complex interspecific hybrids from both national and international sources, along with non-*vinifera Vitis* species [51], and the phenotypic performance of these accessions has been characterized under Brazil's subtropical conditions. Leveraging the broad genetic base of this diverse panel of grapevine resources, we carried out a GWAS analysis on a set of 288 preselected genotypes to investigate six

yield-associated traits to identify novel genomic regions associated with these traits and confirm previously detected QTLs, as well as locate candidate genes within the identified genomic regions. The results of this study provide valuable resources for grapevine breeding by offering insights into the molecular mechanisms potentially involved in cluster architecture, establishing a foundation for subsequent validation experiments, the functional characterization of candidate genes, and the implementation of targeted gene editing strategies.

Materials and methods

Plant material and DNA extraction

The tested population consisted of 288 accessions of *Vitis* spp. from the IAC Grapevine Germplasm Bank, located in Jundiaí, São Paulo (SP), Brazil (23°06’53”S, 46°57’38”W, elevation 745 m). In accordance with the Köppen classification [52], the climate is categorized as Cfa, characterized by subtropical conditions featuring dry winters and hot summers, and the soil in the area is classified as Cambisol Dystrophic Red [53].

The accessions were previously characterized for their trueness-to-type and molecular diversity [51,54], and only distinct genotypes were selected. The material consisted of *Vitis vinifera* cultivars, interspecific hybrids, and non-*vinifera Vitis* species. Among the interspecific hybrids, the most represented groups were those derived from the IAC breeding program, the genotypes developed by the breeding program conducted by Albert Seibel in France at the end of the 19th century, commonly referred to as the Seibel series, and hybrids with a predominant contribution from *Vitis labrusca* (see S1 Table for detailed data on the accessions). The germplasm is established on a single experimental farm, with the accessions distributed across three physically adjacent fields (A,

B, and C; S1 Fig). Each accession was planted in only one of the fields and was represented by a single plot of clonally propagated vines planted side by side. In fields A and B, each plot contained three vines, whereas in field C, each plot contained four vines. All plants were grafted onto 'IAC 766 Campinas' rootstock and trained in a vertical shoot positioning (VSP) system, with spacing of 2.5 m between rows and 1 m between vines. Production pruning was carried out annually in August, leaving one or two buds per branch. After pruning, 2.5% hydrogen cyanamide was applied to induce and stimulate a more uniform budburst. Additional culturing and phytosanitary management practices were also performed according to the standard practices for local growers in São Paulo (SP), Brazil. For sampling purposes, young leaves were collected from a single plant from each accession.

Total genomic DNA was extracted from the leaf tissue with a DNeasy® Plant Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. DNA quality was assessed via both agarose gel electrophoresis and a NanoDrop 8000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, United States). The DNA concentration was determined with a Qubit 3.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, United States).

Phenotypic data

Phenotypic characterization of clusters and berries was conducted over twelve production cycles from 2010–2021. Harvests were carried out between December and February of each cycle, and six traits related to the berries and clusters were evaluated. For each accession, six clusters were collected per plot, and the cluster weight (CW, in grams), cluster length (CL, in centimeters), and cluster width (CWi, in centimeters) were recorded. The same six clusters were also used to obtain berry-related traits, resulting in six replicated berry measurements per accession. Measurements were

taken from ten berries sampled from the upper, middle, and lower portions of each cluster in a 3:4:3 ratio. The sum across the ten berries was used to obtain the berry weight (BW, in grams), berry length (BL, in centimeters), and berry width (BWi, in centimeters). Weight was measured using an analytical balance, and length and width were determined using a metric scale.

Outliers from the phenotypic data were removed using the "outlierTest" function in the *car* package in R [55], employing the Bonferroni outlier test with a statistical threshold of $p < 0.05$. The best linear unbiased estimators (BLUEs) of grapevine genotype effects were estimated for each trait to obtain adjusted genotypic values suitable for downstream analyses. Genotypic BLUE values were obtained through spatial adjustment based on rows and columns, using the 'mmec' function from the R package *sommer* [56]. The mixed model employed was as follows:

$$Y_{ij} = X_g \beta_g + Z_{gxe}(G\ x\ E) + Z_r R + Z_c C + e$$

where Y_{ij} represents the phenotype of the *i*-th genotype in the *j*-th field environment (A, B or C). Genotype effects (β_g) were modeled as fixed at this stage to obtain unbiased adjusted means (BLUEs), with the corresponding design matrix X_g. The term Z_{gxe} represents the random effect of the genotype × environment interaction, while Z_r and Z_c correspond to random row and column effects, respectively. Rows and columns were defined according to the physical position of each accession within its respective field, allowing these effects to capture the spatial heterogeneity within each field. The residual error term *e* is assumed to follow a normal distribution, $e \sim N(0, \sigma^2_e)$.

Variance components were derived from the output of the *sommer* package, and broad-sense heritability (H^2) was estimated using a modified version of the mixed model in which genotype effects were treated as random. Heritability was calculated as $H^2 = VG/(VG + VE)$, where VG represents the estimated genetic variance and VE the environmental variance. The adjusted genotypic values (BLUEs) were used as phenotypic inputs for GWAS and correlation analyses. Pearson correlations between traits were estimated from BLUEs, and the correlation matrix was plotted in R using *ggplot2* and the *ggpairs* function in Ggally. In this study, BLUEs were used as phenotypic

inputs for GWAS instead of best linear unbiased predictors (BLUPs), because population structure and relatedness were explicitly modeled at the association stage through principal components and a kinship matrix, thereby avoiding double shrinkage of genotypic effects.

SNP genotyping and linkage disequilibrium

The 288 accessions were genotyped using the commercial Vitis18kSNP array (Illumina Inc., San Diego, CA, United States) at Neogen (Pindamonhangaba, Brazil), following the Infinium HD Assay Ultra Protocol (Illumina Inc., San Diego, CA, USA). The raw SNP data were visualized and clustered in GenomeStudio 2.0 software (Illumina Inc., San Diego, CA, USA) with a GenCall threshold of 0.15. The SNP scoring tool ASSIsT 1.02 [57] was subsequently utilized, employing default thresholds tailored for germplasm material, to filter the dataset. SNPs classified as “Monomorphic”, “Failed”, or “NullAllele-Failed” were removed. Additional filtering based on minor allele frequency (MAF) was performed using the snpReady software [58], retaining SNPs with MAF > 0.05. The remaining SNPs were used as the high-quality subset for further analysis. Imputation was performed via the k-nearest neighbor imputation (kNNi) algorithm [59].

Linkage disequilibrium (LD) was estimated at the population level using PLINK v1.9 [60] based on pairwise squared allele-frequency correlations (r^2) between SNPs. LD was calculated for all SNP pairs located within a maximum physical distance of 500 kb, using the options --ld-window 99999, --ld-window-kb 500, and --ld-window-r2 0. LD decay was visualized by plotting the mean r^2 values against physical distance using the *ggplot2* R package [61]. LD decay was summarized as the physical distance at which the mean r^2 declined to half of its initial value after accounting for background LD at long physical distances [62].

Genome-wide association study

GWAS was performed for each trait independently via the R statistical genetics package GAPIT version 3 [63]. Population structure was assessed by principal component analysis (PCA) based on the SNP genotype matrix, as implemented in GAPIT3. The first three principal components were used as covariates in the GWAS models to account for population structure effects [64]. Relatedness among accessions was modeled through a genomic relationship (kinship) matrix estimated using the VanRaden method [65], which was incorporated as a random effect, as implemented in GAPIT3 software. Two multi-locus models were used for GWAS, including the fixed and random model circulating probability unification (FarmCPU) [66] and Bayesian information and linkage disequilibrium iteratively nested keyway (BLINK) [67], which have been shown to rank among the best-performing methods across traits with low to high heritability, combining high detection power with good control of type I error and fewer spurious associations than standard GWAS models [68−71]. Significant marker–trait associations were defined as those with Benjamini–Hochberg (B&H) false discovery rate (FDR)-adjusted p-values < 0.05, as calculated by GAPIT3. The online tool MapGene2Chrom [72] was employed to generate a map of significant SNPs. Although the SNP identifiers correspond to the original annotation of the Vitis18kSNP array (https://urgi.versailles.inra.fr/Species/Vitis/GrapeReSeq_Illumina_20K) based on the PN40024 12X.v0 reference genome, the flanking sequences of SNP markers showing significant marker–trait associations were subsequently remapped to the updated grapevine reference genome PN40024.v4 [73] to determine their current chromosomal positions and to support downstream gene annotation analyses. In the case of multiple hits, the best hit (smallest E-value) was selected.

Gene annotation

To identify candidate genes associated with significant SNPs, the grapevine reference genome PN40024.v4 [73] was accessed using the JBrowse [74] feature of Grapedia (https://grapedia.org/genomes). Based on LD patterns of the population, a bin of 30 kb around of each significant SNP, which corresponds a conservative window of 15 kb upstream and downstream was adopted to prioritize genes in close physical proximity to the associated loci, facilitating functional interpretation and reducing background noise in candidate gene identification. We utilized the grape gene reference catalog [75] and the UniProt database [76] to obtain functional annotations for each gene identifier. Further functional annotations for each gene identifier were obtained via Blast2GO (https://www.blast2go.com/) to provide additional context.

Ethics statement

No specific permits were required for plant material sampling or field evaluations conducted in this study because all experimental activities were carried out within institutional experimental fields of the Instituto Agronômico (IAC), Jundiaí, São Paulo, Brazil. These field sites are owned, managed, and maintained by the Instituto Agronômico (IAC), which authorized access for the purposes of this research. This study did not involve endangered or protected species.

Results

Phenotype analysis

Phenotypic data on the cluster and berry traits studied were collected over twelve consecutive fruiting seasons. Substantial phenotypic variation was observed among the evaluated grapevine genotypes (S2 Fig). The trait values were adjusted according to the mixed model from which the BLUEs were extracted for further analysis (S2 Table). The variance components estimated for each trait and broad-sense heritability (H^2) are presented in S3 Table. Broad-sense heritability (H^2) was consistently high across all berry traits, with closely aligned values. The highest heritability was observed for berry weight (BW, 0.95), followed by berry length (BL, 0.92) and berry width (BWi, 0.90). With respect to cluster traits, the H^2 values varied, with high heritability for cluster weight (CW, 0.77), relatively high heritability for cluster length (CL, 0.66), and moderate heritability for cluster width (CWi, 0.50).

Overall, all the correlation coefficients obtained were positive and significant (S3 Fig). BW was strongly and equally correlated with BL and BWi ($r > 0.95$). The correlations between cluster traits ranged from moderate to high, with the strongest correlation observed between CW and CWi ($r = 0.80$) and the weakest correlation observed between CL and CWi ($r = 0.61$). Lower correlations were detected between traits for the different organs (berry and cluster), with the highest correlation occurring between CWi and BL ($r = 0.56$) and the lowest occurring between BW and CL ($r = 0.17$).

Genotyping

The grapevine diversity panel was genotyped using the Vitis18KSNP array, and a total of 18,071 SNPs were scored in GenomeStudio. No poor-quality samples (call rate < 0.8) were detected, and all the samples were retained. After quality filtering, 1,437 SNPs (8%) were classified as monomorphic, 1,520 SNPs (8.4%) as failed, 2,105 SNPs (11.6%) as shifted homozygotes, and 745 SNPs (4.1%) as failed null alleles. All these groups were subsequently excluded from the downstream analysis. Additionally, the snpReady software removed 708 (4%) SNPs with a minor

allele frequency (MAF) of less than 0.05 and 441 (2.4%) SNPs with a call rate of less than 0.85. Ultimately, a subset of 11,115 high-quality SNPs was obtained, with an imputation rate of 1.67% (S4 Table).

This filtered SNP dataset was used for PCA to assess the clustering of genetic variation in the grapevine accessions (Fig 1A). PC1 explained 12.38% of the variation in the genotypic data, whereas PC2 and PC3 explained 4.86% and 4.47% of the variation, respectively (Fig 1B). Overall, the PCA revealed a low degree of population stratification in the panel (Fig 1A). Nevertheless, a structure related to the species background of the accessions was evident. Most *V. vinifera* cultivars formed a dense cluster occupying the central–left region of the plot. In contrast, accessions with a greater contribution from wild *Vitis* species showed greater genetic differentiation from the *V. vinifera* group. Interspecific hybrids, including the Seibel series and IAC hybrids, occupied an intermediate position between *V. vinifera* and wild *Vitis* accessions, whereas wild species tended to be more dispersed toward the opposite side of the plot relative to *V. vinifera*. Among the interspecific hybrids, a clearer structure was observed between those belonging to the Seibel series, which clustered mainly in the upper part of the plot, and those related to *V. labrusca*, which were preferentially located in the lower part. Thus, although the overall stratification was weak, the PCA captured a gradient from predominantly *V. vinifera* genotypes to interspecific hybrids and wild *Vitis* species.

Fig 1. **Principal component analysis (PCA) of grapevine genotypes.** (A) Scatter plot with the genomic data represented in two principal components (PC1 and PC2). Each point represents a genotype from the diversity panel. (B) Genetic variation explained by the first ten PCs.

LD decayed rapidly with increasing physical distance between markers (S4 Fig). The initial mean r^2 was approximately 0.16 and declined to half of this value at ~37 kb after accounting for background LD.

Genome-wide association study

GWAS was conducted using the BLINK and FarmCPU statistical models to explore associations between SNP markers and fruit traits in 288 grapevine accessions. The QQ plots generated from the GWAS models showed a pronounced deviation from the expected p-value distribution only in the upper tail, indicating that population structure and familial relatedness were effectively controlled by the selected covariates while still allowing the detection of true associations (Fig 2). The Manhattan plots summarize the genome-wide association results obtained from the two statistical models for all traits evaluated (Fig 3). GWAS using BLINK and FarmCPU identified 41 and 32 SNP loci, respectively, with significant marker–trait associations. When considering SNP loci defined by their genomic position, 17 loci were detected by both methods, resulting in a total of 56 unique SNP loci with significant associations across all traits (S5 Table, Fig 4). Among these SNP loci, six were associated with two different traits, and two loci were associated with three traits simultaneously (S6 Table, Fig 4). The updated positions of the SNPs associated with the traits, mapped to the reference genome of the grapevine PN40024.v4, enabled the identification of the genes harboring these SNPs. Among the identified SNPs, 38 were located within a gene, and 16 were situated within 15 kb of a candidate gene. No gene was found within 90 kb of the SNP chr18_random_1146283_G_T (associated with BW), and the nearest gene to the SNP ae_C_T_chr14_16833185 (associated with BL and BWi) was 51.44 kb away. A total of 132 candidate genes were identified: 112 genes associated with a single trait, 14 genes associated with two traits, and six genes associated with three traits (S7 Table). When only markers with a phenotypic variance explained (PVE) greater than 5% were considered, 29 SNPs met this criterion in at least one of the tested models. Within the 15 kb flanking regions of these markers, 33 candidate

genes were associated with berry-related traits, and 30 were associated with cluster-related traits (Table 1).

Fig 2. Comparison of the QQ plots resulting from the different GWAS statistical models for eight traits in the full grapevine diversity panel. The red diagonal lines indicate the null hypothesis, where the observed and expected p-values would be situated if there were no associations. The shaded area indicates a 95% confidence interval; BW, berry weight; BL, berry length; BWi, berry width; CW, cluster weight; CL, cluster length; CWi, cluster width.

Fig 3. Manhattan plots indicating single-nucleotide polymorphisms (SNPs) associated with the traits targeted in this study, as predicted by the BLINK and FarmCPU models. The vertical lines indicate SNPs commonly detected by both models. (A) BW, berry weight; (B) CW, cluster weight; (C) BL, berry length; (D) CL, cluster length; (E) BWi, berry width; (F) CWi, cluster width.

Fig 4. Distribution of SNPs associated with the traits according to the BLINK and FarmCPU models. The horizontal bars represent the total number of SNPs per trait, and the vertical bars indicate SNPs that are either unique or shared among traits. Colors in the vertical bars distinguish SNPs identified exclusively by one model (BLINK or FarmCPU) and those detected by both models simultaneously. (A) SNPs associated with berry traits: BW, berry weight; BL, berry length; BWi, berry width. (B) SNPs associated with cluster traits: CW, cluster weight; CL, cluster length; CWi, cluster width.

Table 1. Summary of significant markers associated with berry and cluster traits identified through genome-wide association analysis (GWAS), including only those with PVE greater than 5% and the corresponding candidate genes located within a 15 kb region upstream and downstream of each significant marker.

Trait	Associated marker	Chr	Position	p-value	PVE (%)	Model	Gene	Gene Annotation
BL	chr1_5754629_C_T	1	5737661	1.19E-08	6.36	BLINK	Vitvi01g00511	Heat stress transcription factor (HSF) A-8
							Vitvi01g00512	Vacuolar iron transporter 1 (VIT1)
							Vitvi01g00513	Vacuolar iron transporter 1 (VIT1)
	lb_G_A_chr6_1714702[a]	6	1703187	2.44E-07–5.91E-10	0.58–11.02	FarmCPU, BLINK	Vitvi06g00141	Glycosyl hydrolase 1 family (GH1)
							Vitvi06g00143	Glycosyl hydrolase 1 family (GH1)
							Vitvi06g01615	Glycosyl hydrolase 1 family (GH1)
	chr8_21881697_G_T	8	21872602	5.02E-07	7.49	BLINK	Vitvi08g01906	Peptidase M16 family
							Vitvi08g01910	Peptidase M16 family
	chr11_1513738_G_T	11	1513466	1.48E-06	9.25	BLINK	Vitvi11g00146	NB-LRR family
							Vitvi11g00147	Acetyltransferase component of pyruvate dehydrogenase complex (DLAT/E2)
							Vitvi11g00148	Lycopene epsilon cyclase
	ae_C_T_chr14_16833185[a]	14	16527961	1.59E-09–4.33E-08	30.49–35.22	FarmCPU, BLINK	No hit	-
	chr19_23684078_C_T	19	24046863	1.94E-06–2.84E-06	0.19–10.54	FarmCPU, BLINK	Vitvi19g01761	General transcription factor 3C polypeptide
BWi	ae_A_G_chr1_12009877	1	12313223	1.29E-07	6.47	FarmCPU	Vitvi01g01010	TATA box-binding protein-associated factor RNA polymerase I subunit B
	chr2_3892599_A_G	2	3690239	4.09E-08	6.64	BLINK	Vitvi02g00402	Lysophospholipid acyltransferase 1
							Vitvi02g00403	VvMATE50 - MATE efflux family protein
							Vitvi02g00404	Uncharacterized protein
	be_T_G_chr3_12208763	3	11839072	3.23E-06	5.38	BLINK	Vitvi15g01767	Auxin response factor ARF6
							Vitvi15g04149	Protein kinase domain-containing protein
	chr5_15221260_C_T[c]	5	15456289	2.71E-09	5.42	BLINK	Vitvi05g01115	Myb-like protein X

	lb_G_A_chr6_1714702[a]	6	1703187	2.91E-08–1.12E-08	0.11–11.70	FarmCPU, BLINK	Described in BL	-
	chr6_14784507_A_G	6	14712992	3.59E-06–7.73E-11	1.24–12.66	FarmCPU, BLINK	Vitvi06g01132	Pseudouridine-5'-phosphate glycosidase
							Vitvi06g01133	GRAS/Scarecrow-like protein 3 (SCL3)
	chr11_6170462_C_T	11	6176021	1.11E-09	5.55	BLINK	Vitvi11g00570	Prolamin-like domain-containing protein
							Vitvi11g00571	Triphosphate tunnel metalloenzyme
	ae_C_T_chr14_16833185[b]	14	16527961	7.71E-09–2.27E-11	33.04–29.89	FarmCPU, BLINK	No hit	-
BW	chr4_2974988_A_C	4	2969932	2.78E-07	6.61	FarmCPU	Vitvi04g00314	Glycosyltransferases
							Vitvi04g00316	ER membrane protein complex subunit 1
							Vitvi04g00317	U-box domain-containing protein 26
	be_T_C_chr4_6544365	4	6502499	4.01E-11	19.23	BLINK	Vitvi04g00594	Chorismate mutase (CM2)
							Vitvi04g00595	Kinetochore protein NUF2
							Vitvi04g01927	DNA-directed RNA polymerase II subunit B
	chr5_15221260_C_T[c]	5	15456289	2.24E-12	7.50	BLINK	Described in BWi	-
	lb_G_A_chr6_1714702[a]	6	1703187	8.67E-10	14.23	BLINK	Described in BL	-
	chr11_5942540_A_G	11	5948097	1.70E-07	7.77	BLINK	Vitvi11g00552	Ubiquitin carboxyl-terminal hydrolase
							Vitvi11g00554	ARP2/3 complex 21 kDa subunit
	chr19_23949121_G_T	19	24311977	2.75E-10	6.26	BLINK	Vitvi19g01777	Uncharacterized protein
							Vitvi19g01778	NAC domain-containing protein
CL	chr3_16593567_A_G	3	18000659	4.42E-08–1.48E-07	1.58–7.55	FarmCPU, BLINK	Vitvi03g01151	Peptidyl-prolyl cis-trans isomerase (FKBP12)
	chrUn_7297439_A_G	7	1035840	5.30E-07	24.29	BLINK	Vitvi07g02106	Uncharacterized protein
							Vitvi07g02107	Ankyrin repeat-containing protein
							Vitvi07g02108	Uncharacterized protein
	chr8_18600161_C_T	8	18604552	2.75E-06–1.65E-07	0.61–5.86	FarmCPU, BLINK	Vitvi08g01585	Thaumatin-like protein 1
							Vitvi08g01586	Glutamine synthetase
							Vitvi08g02292	Uncharacterized protein
	lb_T_C_chr14_25461770	14	25894747	4.39E-07–1.65E-09	18.34–23.11	FarmCPU, BLINK	Vitvi14g01522	Reductase

							Vitvi14g01523	VvWRKY43 - WRKY domain-containing protein
CWi	be_C_A_chr3_5463929	3	4044392	2.50E-06–3.78E-06	15.76–1.64	FarmCPU, BLINK	Vitvi03g00359	PMT5 - Sugar transporter family
							Vitvi03g00360	PLATZ transcription factor
							Vitvi03g00361	Pentatricopeptide (PPR) repeat-containing protein
	chr4_16706884_C_T	4	16951942	1.10E-07	11.43	BLINK	Vitvi04g01164	Seipin-1
	be_T_C_chr6_7091603	6	7035547	2.04E-08	21.90	BLINK	Vitvi06g00614	Cell division control protein (CDC6)
							Vitvi06g00615	Ribosomal protein P0 (RPP0A) acidic 60S
	chr7_19824302_C_T	7	28297095	1.9E-08	7.53	FarmCPU	Vitvi07g01988	Translation factor GUF1 homolog
							Vitvi07g01989	MACPF domain-containing protein
	chr15_18113172_A_G	15	19889585	3.62E-06–3.86E-08	12.26–8.43	FarmCPU, BLINK	Vitvi15g00994	Zinc finger CCCH domain-containing protein 1
							Vitvi15g00995	Lipase class 3
							Vitvi15g01618	Poly(A) RNA polymerase mitochondrial-like central palm domain-containing protein
	chr18_20491804_A_G	18	23547288	1.09E-06–2.98E-07	6.20–19.56	FarmCPU, BLINK	Vitvi18g01679	TOD1/MUCI70 glycosyltransferase-like domain-containing protein
							Vitvi18g01681	VvCPK16-Calcium-dependent protein kinase
CW	chr1_5933624_A_G	1	5916771	2.05E-06–3.11E-06	0.67–27.14	FarmCPU, BLINK	Vitvi01g00524	Endoplasmic reticulum transmembrane protein
							Vitvi01g00525	U-box domain-containing protein
							Vitvi01g00526	Protein kinase domain-containing protein
	chr9_290178_C_T	9	290018	1.04E-06–5.56E-07	0.09–10.56	FarmCPU, BLINK	Vitvi09g00022	Uncharacterized protein
							Vitvi09g00023	1-phosphatidylinositol-3-phosphate 5-kinase
							Vitvi09g00026	Translation initiation factor eIF2B subunit alpha
							Vitvi09g00027	Translation initiation factor eIF2B subunit alpha
	chr10_17990933_A_C	10	24330783	3.84E-10–1.98E-07	4.87–11.25	FarmCPU, BLINK	Vitvi10g01581	Katanin p80 (WD repeat containing) subunit B 1

BW, berry weight; (B) CW, cluster weight; (C) BL, berry length; (D) CL, cluster length; (E) BWi, berry width; (F) CWi, cluster width. SNPs associated with more than one trait share the same superscript.

SNP associations and candidate genes for berry traits

In the GWAS, a total of 10 SNPs were significantly associated with BL (Fig 4A). These markers are distributed across Chrs 1, 2, 6, 8, 11, 14, 17, 18, and 19 (Fig 5). Three of the markers, those on Chr 6 (lb_G_A_chr6_1714702), Chr 14 (ae_C_T_chr14_16833185), and Chr 19 (chr19_23684078_C_T), were identified by both models. The phenotypic variance of BL explained by the significant SNPs ranged from 0.19–35.2% (S6 Table). The marker ae_C_T_chr14_16833185 on Chr 14, identified by both models, accounted for the highest percentage of the phenotypic variance in BL (PVE = 35.2%). The GWAS identified thirteen SNPs with significant associations with BWi across ten different chromosomes (1, 2, 5, 6, 7, 10, 11, 14, 15, and 16). Three of these SNP markers were identified by both models and explained 11.7% (lb_G_A_chr6_1714702; on Chr 6), 12.6% (chr6_14784507_A_G; on Chr 6), and 33% (ae_C_T_chr14_16833185; on Chr 14) of the phenotypic variance in BWi. In this study, the greatest number of significant marker–trait associations was identified for BW. The models identified a total of 20 SNPs, but none were identified simultaneously by both (Fig 4A). These markers are distributed across fourteen different chromosomes, excluding only chromosomes 7, 9, 10, 12, and 13. The percentage of PVE by these markers ranged from 0.01% (chr18_7027167_A_G; on Chr 14) to 19.2% (be_T_C_chr4_6544365; on Chr 4). Considering that the three berry traits together reflect berry size, QTLs were found for all three on Chrs 1, 2, 6, 11, and 14. Notably, two significant SNPs (chr2_3892599_A_G and lb_G_A_chr6_1714702) were shared among BL, BW, and BWi. Additionally, two SNPs (chr1_5754629_C_T and chr18_7027167_A_G) were common to BL and BW, two SNPs (chr5_15221260_C_T and be_T_G_chr3_12208763) were common to BW and BWi, and one SNP (ae_C_T_chr14_16833185) was shared between BL and BWi (Fig 4A).

Fig 5. Chromosomal localization of significant SNPs associated with berry- and cluster-related traits. The chromosome number is shown at the top of each chromosome, and chromosome sizes are depicted on a vertical scale (Mb).

A set of 80 distinct genes was associated with at least one of the berry size traits (S7 Table). Among these genes, 26 were associated with BL, 50 with BW, and 26 with BWi. Ten of these genes were associated with two traits simultaneously, and interestingly, six genes (Vitvi02g00402, Vitvi02g00403, Vitvi02g00404, Vitvi06g00141, Vitvi06g00143, and Vitvi06g01615) were shared among the three evaluated berry traits. Among the 80 candidate genes, 72 (90%) had at least one Gene Ontology (GO) annotation, resulting in a total of 273 GO terms (S7 Table). The GO terms were classified into three main categories: cellular component (50 GO terms), molecular function (99 GO terms), and biological process (124 GO terms). In general, the noted GOs associated with berry characteristics are linked to the synthesis of organic monomers necessary for cellular structures (carbohydrates, lipids, and pigments). Terms related to growth (cell division and distension), control of oxidative stress, accumulation of substances characteristic of ripe fruit, such as small carbohydrates and aromatic molecules, and pre- and posttranslational regulation are also present (S7 Table). BL specifically identifies significant terms associated with DNA replication and transcription, phospholipid metabolism, carotenoids, protein recycling, and phosphorylation. BW is associated with terms related to secondary wall synthesis, lipids, membrane transporters, aromatic amino acids, transcriptional regulation, and proteolysis. BWi is associated with terms related to responses to gibberellic acid (GA) and jasmonic acid (JA), lipid synthesis, nutrient and ATP transport, fucosylation, and ribosomal synthesis. Annotated terms associated with more than one trait simultaneously were also identified, including those related to transcriptional regulation, auxin response, cellulose, and anthocyanin synthesis. Among the terms commonly associated with all three

berry traits, the annotated terms are related to the synthesis and transport of small carbohydrates, phospholipids, and triglycerides (S7 Table).

Among the several candidate genes associated with berry traits, Vitvi06g00141, Vitvi06g00143, and Vitvi06g01615, which are linked to all three traits, encode members of the glycosyl hydrolase family 1 (GH1). Genes encoding proteins from the multidrug and toxic compound extrusion (MATE) family were also associated with the three traits, with Vitvi02g00403 (VvMATE50) linked to BW, BL, and BWi. In contrast, Vitvi08g00085 (VvMATE22) and Vitvi08g00086 (VvMATE23) were specifically associated with BW. Among the genes jointly associated with BW and BL, Vitvi01g00512 and Vitvi01g00513 encode vacuolar iron transporter 1 (VIT1), whereas Vitvi01g00511 encodes heat stress transcription factor A-8 (HSFA8). Genes associated with both BW and BWi include Vitvi05g01115, encoding MYB-like protein X (MYB-X), and Vitvi15g01767, which encodes auxin response factor 6 (ARF6). Genes exclusively associated with BW include notable candidates such as Vitvi19g01778, encoding an NAC domain-containing protein (NAC TF), and Vitvi11g00554, which encodes a subunit of the actin-related protein 2/3 complex (ARP2/3 complex subunit). With respect to genes linked solely to BL, we identified genes encoding enzymes from the peptidase M16 family (Vitvi08g01906 and Vitvi08g01910); Vitvi11g00146, a member of the nucleotide-binding site leucine-rich repeat (NB-LRR) family; and a gene encoding an acetyltransferase component of the pyruvate dehydrogenase complex (DLAT/E2). Finally, among the genes associated with BWi, Vitvi06g01133 encodes a gibberellin-mediated growth regulator from the GRAS/SCARECROW-LIKE 3 family (GRAS/SCL3), and Vitvi01g00473 encodes jasmonate ZIM-domain protein 6B-like (JAZ/TIFY6B).

SNP associations and candidate genes for cluster traits

Marker–trait association analyses revealed a total of eight SNP markers associated with CL (Fig 4B), explaining 0.1–24.2% of the phenotypic variance (S6 Table). These SNPs are distributed across Chrs 1, 3, 4, 7, 8, 14, 18, and 19 (Fig 5). Notably, four of these markers (chr3_16593567_A_G, chrUn_28901725_A_G, chr8_18600161_C_T, and lb_T_C_chr14_25461770 on Chrs 3, 4, 8, and 14, respectively) were identified simultaneously by both models (Fig 4B). With respect to CW, four SNPs were identified, three of which (chr1_5933624_A_G, chr9_290178_C_T, and chr10_17990933_A_C) were identified simultaneously by both models. These SNPs are distributed across Chrs 1, 5, 9, and 10, explaining 0.6–27.1% of the phenotypic variance. For CWi, the GWAS revealed 11 significantly associated SNPs on Chrs 3, 4, 5, 6, 7, 9, 14, 15, and 18. Four SNPs were consistently identified by both methods, with two of these located in different regions of Chr 15. The PVE values of the significant SNPs ranged from 0.15–21.8% for CWi. Only one significant SNP was associated with more than one cluster trait (Fig 2B). The SNP chr9_290178_C_T on Chr 9 was common to both CW and CWi.

A group of 52 distinct genes was pinpointed as potentially associated with cluster size traits. Among them, 20 are exclusively linked to CL, 6 to CW, and 22 to CWi. Notably, four genes (Vitvi09g00022, Vitvi09g00023, Vitvi09g00026, and Vitvi09g00027) are implicated in both CW and CWi. Furthermore, of these 52 genes, 42 (80.76%) have at least one Gene Ontology (GO) annotation, collectively providing 173 distinct GO terms, with 74 classified as biological processes, 30 as cellular components, and 69 as molecular functions (S7 Table). In general, the annotated terms associated with cluster traits are related to ribosomal RNA processing, protein synthesis, transport, and packaging. For CL, in addition to protein-related processes, terms related to glutamine biosynthesis were identified. In the case of CW, the annotated terms were linked to JA biosynthesis, lipid catabolism, and cytoskeleton depolymerization. CWi was enriched for terms related to DNA and ribosome biosynthesis, lipid catabolism, and calcium ion transport. When CW and CWi are

considered together, the terms associated with both traits indicate a role in translation initiation and calcium-mediated signaling (S7 Table).

Notable among the genes associated with CW and CWi are Vitvi09g00026 and Vitvi09g00027, which encode translation initiation factor 2B subunit alpha (eIF2Bα). In addition, Vitvi01g00572, encoding translation initiation factor 3 subunit 2 (eIF3-S2), was associated with CL. Other interesting genes associated with CL include Vitvi01g00571, belonging to the nitrate transporter 1/peptide transporter family (NRT1/PTR); Vitvi14g01523, encoding a WRKY domain-containing protein 43 (VvWRKY43); Vitvi04g02280, encoding protein phosphatase 2C (PP2C); and Vitvi03g01151, encoding peptidyl-prolyl cis-trans isomerase (FKBP12). With respect to the genes linked to CWi, notable candidates include Vitvi03g00359, which encodes a member of the polyol/monosaccharide transporter (PMT) family; Vitvi06g00614, encoding cell division control protein 6 (CDC6); Vitvi07g01989, encoding membrane attack complex/perforin (MACPF) domain-containing protein; Vitvi15g00994, encoding zinc finger CCCH domain-containing protein 1; and Vitvi15g00915, encoding cytochrome P450 78A40 (CYP78A40) (Table 1 and S7 Table).

Discussion

Association panel design

Genome-wide association mapping is a robust approach for dissecting the genetic architecture of complex traits. The effectiveness of this method relies on the use of association panels with sufficiently large population sizes and diverse genetic backgrounds [42,77,78]. Recent GWASs in grapevine have further demonstrated that, to precisely capture the genome-wide distribution of LD and accurately identify markers associated with complex traits, it is essential to include a wide array

of diverse individuals representing existing phenotypic variability [45,79,80]. In this study, we selected 288 grapevine genotypes from different origins (S1 Table) to encompass a substantial portion of the genetic and phenotypic diversity found in Brazilian germplasms. The IAC grapevine germplasm encompasses approximately 420 accessions, 100 of which are exclusive to this germplasm and originate from the institution's breeding program [54]. In a previous study, this collection was examined with microsatellite (SSR) markers, revealing a broad spectrum based on the species and the intended use of the accessions [51]. To design our grapevine association panel, we considered the results of genetic structure analyses generated from SSR studies. We selected the key genotypes from each genetic cluster, avoiding closely related individuals (i.e., first-degree relatives or mutants).

The 120 accessions from the previously established core collection [51] were prioritized in the selection process to enhance the representation of different alleles of interest across the genetic pools. This selection strategy successfully limited the structuring of the germplasm. In this study, the first three principal components accounted for 21.71% of the genomic population structure, revealing only moderate clustering among the *V. vinifera* accessions. This low level of structure in our population reflects the diverse geographic origins of the grape cultivars and the presence of various intra- and interspecific hybrids, which strengthened the association analyses.

The PCA results were consistent with the patterns previously revealed by SSR markers, showing a clear discrimination based on species background. *V. vinifera* cultivars formed a relatively compact cluster, whereas wild *Vitis* accessions were more distant from this group, reflecting greater genetic differentiation. This pattern agrees with previous findings reporting high genetic divergence between *V. vinifera* and the North American *Vitis* species, which belong to distinct clades [81–83]. *V. labrusca* is part of the North American *Vitis* group, which explains why accessions strongly related to this species were separated from the *V. vinifera* cluster and formed a distinct group. The remaining interspecific hybrids occupied an intermediate position between *V. vinifera* and non-*vinifera Vitis*

accessions, which is consistent with their complex pedigrees. Many of these hybrids originate from crosses involving three or more species; in such cases, both deliberate cross-breeding and natural hybridization across species boundaries generate genotypes with heterogeneous genomic compositions, making it challenging to assign them to a single genetic group because they carry alleles from multiple gene pools [83–85].

Within the interspecific hybrids, an additional level of structure was evident: accessions belonging to the Seibel series tended to cluster together. For the Seibel hybrids, this pattern likely reflects their breeding history, as many were developed within the same breeding program in France and were repeatedly used as parents across different breeding cycles, contributing to the formation of a relatively distinct gene pool. A more detailed description of the genetic structure of this collection is provided in the study by De Oliveira et al. [51]. Overall, the presence of interspecific hybrids with complex pedigrees contributed to a weaker and more diffuse structure across the diversity panel, while *V. vinifera* genotypes, despite differing in geographic origin and use, still belong to the same species and therefore formed a more evident cluster.

LD decayed rapidly with increasing physical distance in the grapevine diversity panel, with the mean r^2 declining to half of its initial value at approximately 37 kb (S4 Fig). This rapid decay pattern is consistent with previous studies in grapevine, which have shown that the extent of LD can vary widely across and within species depending on the genetic composition of the diversity panel [38–40,45,80,86,87]. Considering the high diversity of the panel analyzed in this study, including cultivated varieties, wild species, and interspecific hybrids, we adopted a conservative physical window for candidate gene identification. This strategy aimed to prioritize genes in close physical proximity to significantly associated SNPs, facilitating functional interpretation and reducing background noise by limiting the inclusion of indirectly associated or potentially spurious candidate genes.

QTLs for berry size-related traits

Traits related to berry and cluster size are closely associated with grapevine productivity and quality. Berry size stands out as one of the most economically significant agronomic traits for table grapes, acting as a crucial quality factor that shapes consumer preferences and acceptance [15]. It is also a key determinant of juice composition and final wine quality, particularly in red wines [17]. Furthermore, natural variation in berry size significantly influences the volatile profiles of grape berries [18]. Fruit size is typically assessed by considering three dimensions: weight, length, and width. In the case of grape berries, as observed in this study, these three characteristics are typically reported to have a strong positive correlation [15,45,88,89]. This suggests that berries larger in one dimension, such as length, also tend to be larger in other dimensions, including width and weight [90]. The high broad-sense heritability (H^2) estimates for BW, BL, and BWi obtained in this study indicate that these traits are under strong genetic control. This aligns with previously reported high H^2 values (ranging from 50–97%) for berry size-related traits in other populations [15,24,37,88], suggesting that phenotypic variation is predominantly derived from genetic variance. These findings imply that association analyses to identify favorable alleles linked to molecular markers hold promising potential for MAS.

Several QTLs for berry size-related traits have already been reported, and distinct QTLs have been detected in different genetic backgrounds. An extensive literature review was conducted to investigate potential overlaps between the QTLs identified in this study and previously reported QTLs associated with any berry size-related traits [15,22–32,34,37,40,44–46,91,92]. A major berry weight QTL on Chr18 colocalized with the seedless trait was previously found in several studies near the SSR marker VMC7F2 [22,30–32]. The GWAS findings in this study validate the presence of significantly associated SNPs for BW on Chr18. However, the locations of the two loci identified

here differ from those found through biparental QTL mapping. This discrepancy from previous biparental mapping studies may be due to the use of segregating populations derived from crosses involving seedless cultivars. In GWASs using diversity panels, Guo et al. (2019) [40] and Flutre et al. (2022) [37] reported that approximately 5–6 Mbp of the SNPs identified in the present study were QTLs on Chr18, which also diverged from the BW QTL colocalized with the seedless trait. Some seedless cultivars exhibit reduced berry weight at harvest, necessitating exogenous applications of gibberellic acid and cluster thinning to maximize berry growth potential. This reduction may be due to the absence of growth regulators typically produced by seeds [22,93]. Consequently, the markers colocalized with the QTLs for seed number and seed size indirectly affect berry size [94].

In this study, the SNP markers associated with BW overlapped with previously reported QTLs on chromosomes 1, 2, 3, 4 [37], 5 [44], 6 [46], 8 [22,37,44], 11 [22,26], 14 [37], 17 [22,44], and 19 [37,44] (S6 Table). On Chr1, we identified two potential colocalizations with the QTLs reported by Flutre et al. (2022) [37] for BW. The first is in the region containing the SNP shared by both BW and BL traits (chr1_5754629_C_T) and the SNP identified for BWi (chr1_5295802_C_T), approximately 0.8 Mbp away. The second colocalization is with the SNP for BW (chr1_22609084_C_T), located approximately 2.8 Mbp from one of the QTLs reported by these authors. Our results also confirmed a QTL previously identified on Chr17 by Doligez et al. (2013) [22]. This QTL was reported as a stable and large-effect QTL for berry weight that did not colocalize with QTLs for seed traits. However, this QTL was identified using biparental populations and low-resolution linkage maps. Recently, Thorat et al. (2024) [44] identified a marker associated with BW in the same region. By integrating high-throughput genotyping tools and using a diversity panel, they likely increased the precision of the confidence intervals of the QTLs, aligning more accurately with the QTLs identified in this study.

The SNP located on Chr2 (chr2_3892599_A_G), associated simultaneously with the three berry size traits in this study, was found to be less than 1 Mbp from the marker associated with BWi

in the study by García-Abadillo et al. (2024) [45] and approximately 1.4 Mbp from the region related to BW by Flutre et al. (2022) [37]. The simultaneous association of this region with all three berry size-related traits observed in this study is supported by the strong phenotypic correlations between them ($r > 0.9$). A similar situation was observed on Chr6. The SNP simultaneously associated with all three traits (BW, BL, and BWi) was found to be approximately 3 Mbp from the QTL reported by Marrano et al. (2018) [46] for BW. Few studies have focused on identifying QTLs related to BL and BWi compared with BW, likely because of the high correlation among these traits and the intensive labor required for phenotyping BL and BWi. Additionally, SNPs exclusively associated with BWi had the lowest overlap rate with previously reported QTLs. However, based on our findings, we can infer that many QTLs reported solely for BW may also be associated with other berry size-related traits. In other words, the QTLs identified for one of these traits are likely to influence all dimensions of berry size, probably because the same genetic factors govern these characteristics.

Many regions identified in this study for the three berry traits colocalized with QTLs previously reported for different berry size-related characteristics. On Chr8, three distinct regions were associated with BW. Two of these regions (ae_A_G_chr8_11440804 and chr8_17424579_A_G) had previously been reported for this trait [22,37,44] and were also associated with berry shape and BWi by García-Abadillo et al. (2024) [45], respectively. The third region (chr8_1499673_G_T) exactly matched the QTLs identified by Wang et al. (2022) [15] for BL and BWi. Additionally, on Chr8, the SNP associated with BL (chr8_21881697_G_T) overlapped with the QTLs identified by Flutre et al. (2022) [37] for BW and by García-Abadillo et al. (2024) [45] for BWi. The SNP on Chr16 (chr16_14671632_C_T) associated with BWi likely corresponds to the QTL identified by García-Abadillo et al. (2024) [45] for berry shape. Because this trait was derived by the authors from the ratio of berry height (or BL) to BWi, it can be concluded that these traits are highly correlated and might correspond to the same QTL.

Furthermore, new potential QTLs were identified on chromosomes that are already associated with berry size but in regions that are not consistent with previous studies. Previously unreported associations for berry size were detected in regions of chromosomes 1, 4, 6, 7, 10, 11, 14, 15, 16, 17, 18, and 19. For example, no overlaps were identified for the SNP on Chr14 (ae_C_T_chr14_16833185), which was simultaneously associated with both models for BL and BWi and had a high PVE value (30–35%). The differences in the positions of some QTLs compared with those in previous studies suggest that fruit size-related traits are polygenic and controlled by different causal polymorphisms potentially segregating among various populations [22]. Additionally, factors such as genotyping density, varying phenotyping methods, and the statistical models employed can complicate the precise identification of overlaps with QTLs reported in other studies [37].

QTLs for cluster-related traits

In grape production, subtraits associated with cluster architecture, including cluster and berry size, number of berries per cluster, and cluster compactness, are crucial. These traits not only impact fruit quality and yield but also influence factors such as disease susceptibility, sunlight exposure, and uniform ripening [27,44]. The weight, length, and width of clusters are among the subtraits commonly used to objectively and quantitatively assess cluster compactness and size in research studies. The genetic mechanisms underlying these traits in grapes are complex and influenced by both viticultural practices and environmental factors, resulting in significant year-to-year variation [11,44,95,96]. Consequently, the influence of numerous minor genetic contributions makes it difficult for the literature to reach a consensus on the location of their QTL. Moreover, population or environmental factors could cause these QTLs to be unstable [45]. Nonetheless, certain overlaps were detected. The first concerns the SNP on Chr10 recognized by both models linked to CW, which

coincided with a QTL identified by Richter et al. (2019) [27] connected with the same trait mapped in a biparental population. The region identified on Chr15 associated with CWi corresponded to the QTL found by Flutre et al. (2022) [37] for the same trait. Finally, the SNP linked to CL on Chr14 overlapped with the QTL documented by Thorat et al. (2024) [44] for the same trait, as well as with the QTL discovered by Marrano et al. (2018) [46] related to CW.

In addition, several other instances of QTL colocalization were observed, where the QTLs reported for one cluster differed from the trait associated in this study but were correlated (S6 Table). We found moderate to high correlation levels between cluster length, width, and weight. However, they have been variably reported in other studies, ranging from high to moderate and even low [88,89,97,98]. In addition to being heavily influenced by nongenetic factors, these traits are also affected by other characteristics, such as berry size, the number of berries per cluster, and rachis size, which likely contribute to the variation in their correlation. Nevertheless, these QTL overlaps associated with different cluster traits highlight the interconnected nature of these characteristics and suggest that common genetic factors may influence multiple aspects of cluster architecture.

Additionally, as shown on Chr1, Chr7, Chr8, and Chr14, some regions in this study were associated with both berry and cluster traits simultaneously (Fig 5). Despite being significant, we identified a weak to moderate correlation between cluster and berry traits (S3 Fig). However, cluster traits are indirectly linked to berry traits. For example, as cluster size increases, the likelihood of larger berries may also increase, or larger and heavier berries will contribute more to the total cluster weight [97]. Of course, numerous other factors influence the relationships between these traits, such as the number of berries per cluster, berry density, environmental factors, and gene□gene interactions, which can also moderate the expression of these traits. These factors can result in lower phenotypic correlations while still reflecting a shared genetic basis. It remains challenging to suggest a pleiotropic effect of a locus on several phenotypic traits. However, the proximity of the QTLs for berry- and cluster-related subtraits in these regions offers an opportunity for MAS. This situation

could be leveraged using a small range of molecular markers from these QTL regions to select for larger or smaller berry volumes, accompanied by larger or medium-sized clusters, depending on the intended use for table grapes or wine. By doing so, it may be possible to capture several coinherited traits, potentially leading to a more efficient selection process that aligns with the desired characteristics for each application [27].

Efforts have been made to study the genetic determinants of berry and cluster size through QTL mapping using segregating populations [15,22,27,29–31,34], yielding valuable insights into the genetic inheritance of these complex traits. However, these analyses used populations with relatively limited genetic variability and often failed to capture all recombination events. Using the GWAS approach, we identified many associated markers for each of the studied berry and cluster traits, with some markers explaining substantial phenotypic variation. Despite new associations being identified in this study, overlap with previously reported QTLs was observed (S6 Table). Most of these overlaps were with recent studies that also conducted GWAS on diversity panels [37,44,45]. This reproducibility stems from the technique's ability to leverage historical recombination events to identify haplotype blocks associated with phenotypes of interest across the diversity of crops, identifying stable QTLs that are not population specific.

Candidate genes for berry size traits

Fruit development is typically categorized into three distinct stages: fruit set, fruit growth, and fruit maturation [99]. Fleshy fruiting species undergo phases of cell division and expansion following fruit set, which allows the determination of fruit size before maturation and ripening. Plant hormones regulate various aspects of growth and development, including fruit size, by responding to both environmental and endogenous signals [100]. At the cellular and molecular levels, phytohormones

can influence gene expression and transcription, as well as cell division and growth [101]. In this study, we identified several genes related to hormones that are associated with berry size. Most of these genes were linked to the plant hormones auxin, gibberellic acid (GA), and JA. Studies have shown that auxin and GA are early indicators of fruit setting and growth [102–106]. Genes related to these hormones have been previously linked to berry size and shape traits [22,26,44,87,91]. Auxins play crucial roles in regulating nearly all aspects of plant growth and development by influencing cell division, cell growth, and cell differentiation [107,108]. Auxin response factors (ARFs) are key transcription factors that mediate auxin signaling in plants by binding to auxin-response elements in the promoter regions of early auxin-response genes, thereby regulating auxin-induced gene responses [109,110]. Class A ARFs can also respond to environmental signals through auxin-independent pathways, regulating plant growth and development [111]. QTLs associated with important traits have been increasingly linked to auxin, suggesting that auxin engineering holds tremendous potential for manipulating crop architecture in plant breeding [112].

GA is known to be essential for regulating berry expansion. Research has shown that exogenous GA can promote the division and expansion of grape pulp cells, increase sugar accumulation, and increase fruit size [100,113–115]. The gene Vitvi06g01133, associated with BWi, has been identified in grapevine as a GRAS gene belonging to the SCL3 subfamily and is predominantly expressed in the stem, seed, and berry flesh [116]. Proteins from the SCL3 subfamily are key components in the GA signaling pathway and cooperate to mediate cell elongation during root development in *Arabidopsis thaliana* [117]. The GRAS gene family is a plant-specific group of transcription factors that play crucial roles in various metabolic pathways, including plant growth, development, fruit ripening, and stress response [118,119]. The jasmonate–ZIM domain protein acts as an inhibitory factor in the JA pathway, which is involved in regulating plant development and the response of plants to biotic and abiotic stresses [120–123].

Additionally, several genes involved in responses to abiotic and biotic stresses were associated with berry size traits in this study. This associations likely occur because abiotic and biotic stresses can significantly affect plant growth, yield, and quality by disrupting physiological processes, weakening defense mechanisms, and reducing overall productivity [124]. Heat stress is among the major threats in viticulture [125], significantly affecting berry ripening, quality, and yield when air temperatures exceed 35°C [126]. As a result, high temperatures directly impact berry size [127]. The gene Vitvi01g00511, associated with both BW and BL, encodes an HSF, the most important transcription factor family involved in the response of plants to heat [128]. To reduce the severity of destructive effects, plants use a range of defensive strategies. NB-LRR family genes encode intracellular resistance (R) proteins that directly or indirectly recognize and activate defense responses, including the hypersensitive response and systemic acquired resistance [129,130]. Several transcription factors play crucial roles in regulating the internal defense system of plants. The MYB and NAC families of transcriptional regulators play essential roles in regulating various vital biological processes during plant growth, development, and stress responses [131,132]. Specifically, genes encoding NAC domain-containing proteins have been associated with the early development of grape flowers and berries [133], with several studies suggesting their involvement in determining berry size variation [15,87,92,134,135].

In plants, members of the GH1 family have been identified as key players in various biological processes related to growth and development [136,137]. Notably, substantial evidence indicates that GH1 genes are also involved in the activation of defense compounds in response to stress and the release of active plant phytohormones [138–140]. Studies have demonstrated that in grapevines, β-glucosidase enzymes, a subclass of the GH1 family, participate in responses to waterlogging stress [141], low-temperature stress [142], and abscisic acid (ABA) biosynthesis [143].

Several transport-related genes were associated with the three berry traits across different chromosomes and play crucial roles in plant defense, growth, and development. The transport and

conversion of photosynthetic products are closely tied to sugar accumulation in fruit, with total soluble sugars increasing during growth and peaking at maturity or ripening [144,145]. The genes Vitvi02g00403 (VvMATE50), Vitvi08g00085 (VvMATE22), and Vitvi08g00086 (VvMATE23) belong to the MATE family, which is among the largest families of transporter genes in plants and is involved in various physiological processes throughout plant growth and development [146]. In grapevine, the gene VvMATE50 belongs to a group of genes involved in the transport or signaling of phytohormones such as ABA, auxin, and salicylic acid (SA) and plays a significant role in grape growth and defense responses. The genes VvMATE22 and VvMATE23 are part of a MATE group and function in the transport of flavonoids and alkaloids [147]. According to Watanabe et al. (2022) [147], VvMATE22 is highly expressed in seed and berry tissues during early developmental stages, with expression levels decreasing as development progresses. These findings suggest that VvMATE22 plays a crucial role in regulating the transport of substances that impact berry growth and development. Among the transport-related genes identified in this study, two encode VIT1, which plays a crucial role in maintaining iron homeostasis and regulating various cellular functions essential for fruit development [148,149].

In plant cells, actin filaments indirectly influence cell shape by regulating the transport properties of organelles and cargo molecules, which in turn alter the mechanical properties of the cell wall [150]. The Arp2/3 complex serves as a critical regulator of actin filament dynamics across a wide range of eukaryotic cells [151]. This complex nucleates new actin filaments and stimulates the formation of branched actin networks, which are essential for various cellular processes, including cell shape modulation, cell wall assembly, intracellular trafficking, and cell expansion [152–156]. In the context of grape berry development, it is plausible that the Arp2/3 complex regulates the extent of cell expansion in the pericarp, directly controlling the volumetric increase that defines berry size. Moreover, its participation in vesicle trafficking may affect the deposition of cell wall components,

influencing fruit hardness and elasticity, which are indirectly associated with berry development potential.

Additional genes identified through GWAS that may directly or indirectly affect the determination of final fruit size include members of the peptidase M16 family, which are involved in protein processing and play critical roles in maintaining cellular viability by ensuring protein and energy homeostasis [157–159]. In addition, a gene encoding DLAT/E2 was identified, which may contribute to the regulation of cell expansion. An adequate supply of acetyl-CoA provided by this complex is essential for the biosynthesis of membrane lipids, which are integral constituents of the cell wall, and for energy production (ATP), processes that are critical during fruit growth and development [160,161].

Candidate genes for cluster-related traits

Cluster compactness and size significantly influence the commercial value of both wine and table grapes. These traits are highly complex, with various berry and cluster attributes contributing to their variation across different varieties, each influenced by their genetic architecture [13,162]. Among these attributes, berry dimensions are key factors that directly affect cluster compactness and size [163]. Notably, significant SNPs on Chr1 and 14 associated with CL are near genes encoding proteins from the NRT1/PTR and WRKY domain-containing families. These families were identified by García-Abadillo et al. (2024) [45] as potential candidates for controlling both berry length and width. NRT1/PTR genes, originally identified as nitrate or peptide transporters in plants, are involved in transporting a wide range of substrates, such as amino acids, nitrate, auxin, JA, abscisic acid (ABA), GA, and glucosinolates [164], which play essential roles in the molecular mechanisms influencing fruit size determination [100]. Grape WRKY transcription factors play crucial regulatory

roles in both fruit development and overall grapevine growth. They are also vital for mediating plant responses to various biotic and abiotic stresses [165]. Specifically, the gene encoding VvWRKY43 identified in this study is potentially involved in regulating the stilbene biosynthetic pathway [166]. Stilbenes are important compounds that protect plants from pests, pathogens, and abiotic stresses [167].

Plants have evolved a wide range of defense strategies to cope with biotic and abiotic stresses; however, these responses are often energetically costly and divert resources away from growth, development, and reproduction, including during fruit development [168]. In this study, distinct genes associated with CL and CWi were identified, many of which are involved in stress-related pathways. Members of the MACPF family play crucial roles in maintaining the balance between plant growth and immunity [169], contributing to both biotic and abiotic stress responses, likely through interactions with JA and SA signaling pathways [170–172]. Additionally, genes encoding PP2C have been shown to participate in various stress responses by modulating growth factors, regulating phytohormones, and controlling metabolic processes in several important plant species [173]. In grapevine, PP2Cs are actively involved in the ABA signal transduction pathway [174] and have been proposed as candidate genes affecting BL [44]. Zinc finger CCCH-type proteins also play significant roles in plant adaptation to various abiotic stresses, including oxidative stress, salinity, drought, flooding, and cold temperatures, across different species [175]. In grapevine, these proteins are potentially involved in not only stress responses but also growth and developmental processes [176]. Furthermore, they are involved in various developmental processes in plants, including seed development [177], flowering and senescence [178], plant architecture [179], and secondary wall synthesis [180].

Genes related to translation and protein synthesis were identified in genomic regions near SNPs associated with all three cluster traits. In particular, genes encoding eIF3-S2 and eIF2Bα were detected. These factors play essential roles in the initiation phase of protein synthesis, a critical step

for assembling the ribosomal machinery required to translate mRNA into functional proteins. Their activity ensures the efficiency and fidelity of protein biosynthesis, which is fundamental for supporting cellular processes such as proliferation, differentiation, and adaptation to environmental cues [181,182]. The efficiency of translation, which is regulated by the eIF2Bα and eIF-3 complexes, directly impacts the production of proteins necessary for cell division and expansion [183,184], which are critical processes for fruit development. Other noteworthy genes involved in cell cycle regulation and associated with cluster traits include FKBP12 and CDC6. Plant FKBPs perform diverse functions in essential biological processes related to growth and development [185]. Specifically, FKBP12 has been reported to participate in cell cycle regulation and embryo development, as well as in controlling the direction of pollen tube growth in *Arabidopsis thaliana* and *Picea wilsonii* [186–188]. Moreover, CDC6 plays a central role in the initiation of DNA replication, ensuring accurate and coordinated genome duplication during the cell cycle [189]. Modulation of the expression or activity of these genes may directly impact the rate of cell division in developing tissues, thereby influencing fruit growth and ultimately impacting grape cluster size.

Cluster traits are strongly associated with flower development and fertilization, as cluster growth depends on both cell proliferation and expansion during the development of floral and peduncle structures [190,191]. Reproductive development is an energy-demanding process, from floral initiation to fruit maturation [192]. The transport of soluble sugars across membranes is a tightly regulated process mediated by sugar transporters belonging to multiple gene families [144,193,194]. In grapevine, PMTs have been detected in different vegetative organs [195]. Specifically, the VvPMT5 gene is expressed at high levels in mature leaves, petioles, and tendrils, with its expression peaking during the fruit set stage [195]. The observed expression pattern suggests a potential role for the PMT5 gene in regulating the development and structural definition of the cluster axis. Genes from the PLATZ transcription factor family have also been implicated in flower development in grapevine [196]. They are also known to participate in various aspects of plant

growth and development across different species, including stress responses, cell proliferation, and seed development [197–199]. Notably, in rice, the PLATZ gene *GL6* positively regulates cell division by increasing the number of cells in the spikelet hull, ultimately leading to larger grains [200].

Extensive research has demonstrated that several CYP78A genes are also involved in the regulation of seed development and organ size [201]. Arabidopsis CYP78A9 plays a role in controlling floral organ size and has redundant functions with those of CYP78A6 in integument development [202,203]. Similarly, TaCYP78A3 influences wheat seed size by affecting the extent of integument cell proliferation [204]. In maize, CYP78A1 stimulates leaf growth by extending the duration of cell division [205]. In sweet cherry, PaCYP78A6 acts redundantly with PaCYP78A9 to influence fruit size during growth and development [206]. These findings support the potential involvement of the candidate gene CYP78A40 in determining the final size of the grape cluster.

Other candidate genes identified in this study may influence various grape cluster traits at different levels. The regions identified in GWASs typically contain blocks of credible SNPs, and a major challenge is to pinpoint the causal gene in that locus of interest [207]. Our association study revealed numerous QTLs with multiple genes for various grape yield traits. Further investigation into the expression patterns of these candidate genes at specific stages of fruit development, along with an analysis of their regulatory networks, will provide deeper insights into how these genes contribute to the observed variations in the studied berry and cluster traits.

Conclusion

This study demonstrated the feasibility of conducting a GWAS in a perennial fruit crop such as grapevine for complex traits related to yield and quality, which are crucial for breeding. Using a grapevine diversity panel with accessions phenotyped over 12 seasons under subtropical conditions,

we identified numerous SNP markers that are significantly associated with berry and cluster-related traits. Notably, 67.8% of these markers were located within annotated genes, with some explaining a substantial portion of the observed variation. Additionally, we confirmed several previously known genomic regions and identified new regions linked to the traits studied. This study also connected different traits to QTLs reported in earlier research, strongly supporting the hypothesis that these regions have pleiotropic effects. Several predicted genes have emerged as strong candidates for influencing berry size and cluster architecture, as they are involved in critical functions such as growth regulation, hormone control, protein synthesis, stress response, and other essential physiological processes. We anticipate that the identified marker–trait associations and candidate genes will significantly advance our understanding of the molecular basis underlying cluster architecture and berry size, paving the way for further validation, functional investigation, and gene editing approaches. Ultimately, these findings provide valuable tools for grapevine breeders, contributing to the development of elite cultivars and benefiting the grapevine research community.

Acknowledgments

The authors are grateful to the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), the Coordenac□ão de Aperfeic□oamento de Pessoal de Nível Superior (CAPES), and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for their support of the project and the researchers.

References

1. OIV. State of the world vine and wine sector in 2022. 2023 [cited 22 May 2023]. Available:

https://www.oiv.int/sites/default/files/documents/2023_SWVWS_report_EN.pdf

2. Nanni A, Parisi M, Colonna M. Wine by-products as raw materials for the production of biopolymers and of natural reinforcing fillers: A critical review. Polymers (Basel). 2021;13: 381. doi:10.3390/polym13030381

3. Alston JM, Sambucci O. Grapes in the world economy. In: Cantu D, Walker MA, editors. The Grape Genome. Cham: Springer International Publishing; 2019. pp. 1–24. doi:10.1007/978-3-030-18601-2

4. Terral JF, Tabard E, Bouby L, Ivorra S, Pastor T, Figueiral I, et al. 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. 2010;105: 443–455. doi:10.1093/aob/mcp298

5. Ferreira V, Pinto-Carnide O, Arroyo-García R, Castro I. Berry color variation in grapevine as a source of diversity. Plant Physiol Biochem. 2018;132: 696–707. doi:10.1016/j.plaphy.2018.08.021

6. Gray DJ, Li ZT, Dhekney SA. Precision breeding of grapevine (*Vitis vinifera* L.) for improved traits. Plant Sci. 2014;228: 3–10. doi:10.1016/j.plantsci.2014.03.023

7. This P, Lacombe T, Thomas MR. Historical origins and genetic diversity of wine grapes. Trends Genet. 2006;22: 511–519. doi:10.1016/j.tig.2006.07.008

8. Anand A, Subramanian M, Kar D. Breeding techniques to dispense higher genetic gains. Front Plant Sci. 2023;13. doi:10.3389/fpls.2022.1076094

9. Edge-Garza DA, Luby JJ, Peace C. Decision support for cost-efficient and logistically feasible marker-assisted seedling selection in fruit breeding. Mol Breed. 2015;35: 223. doi:10.1007/s11032-015-0409-z

10. Delrot S, Grimplet J, Carbonell-Bejerano P, Schwandner A, Bert P-F, Bavaresco L, et al. Genetic and genomic approaches for adaptation of grapevine to climate change. Genomic

designing of climate-smart fruit crops. Cham: Springer International Publishing; 2020. pp. 157–270. doi:10.1007/978-3-319-97946-5_7

11. Ohana-Levi N, Cohen Y, Munitz S, Michaelovsky R, Ferman Mintz D, Hagag N, et al. The response of yield, number of clusters, and cluster weight to meteorological factors and irrigation practices in grapevines: A multi-experiment study. Sci Hortic (Amsterdam). 2024;326: 112761. doi:10.1016/j.scienta.2023.112761

12. Wenter A, Zanotelli D, Montagnani L, Tagliavini M, Andreotti C. Effect of different timings and intensities of water stress on yield and berry composition of grapevine (cv. Sauvignon blanc) in a mountain environment. Sci Hortic (Amsterdam). 2018;236: 137–145. doi:10.1016/j.scienta.2018.03.037

13. Tello J, Ibáñez J. Review: Status and prospects of association mapping in grapevine. Plant Sci. 2023;327: 0–2. doi:10.1016/j.plantsci.2022.111539

14. Wang Z, Zhou J, Xu X, Perl A, Chen S, Ma H. Adoption of table grape cultivars: An attribute preference study on Chinese grape growers. Sci Hortic (Amsterdam). 2017;216: 66–75. doi:10.1016/j.scienta.2017.01.001

15. Wang H, Yan A, Wang X, Zhang G, Liu Z, Xu H, et al. Identification of QTLs and candidate genes controlling berry size in table grape by integrating QTL and transcriptomic analysis. Sci Hortic (Amsterdam). 2022;305: 111403. doi:10.1016/j.scienta.2022.111403

16. Chen W-K, He F, Wang Y-X, Liu X, Duan C-Q, Wang J. Influences of berry size on fruit composition and wine quality of *Vitis vinifera* L. cv. ‘Cabernet Sauvignon’ grapes. South African J Enol Vitic. 2018;39. doi:10.21548/39-1-2439

17. Martin D, Lindsay M, Kilmartin P, Dias Araujo L, Rutan T, Yvon M, et al. Grape berry size is a key factor in determining New Zealand Pinot noir wine composition. OENO One. 2022;56: 389–411. doi:10.20870/oeno-one.2022.56.2.5436

18. Xie S, Tang Y, Wang P, Song C, Duan B, Zhang Z, et al. Influence of natural variation in

berry size on the volatile profiles of *Vitis vinifera* L. cv. Merlot and Cabernet Gernischt grapes. Amancio S, editor. PLoS One. 2018;13: e0201374. doi:10.1371/journal.pone.0201374

19. Walker RR, Blackmore DH, Clingeleffer PR, Kerridge GH, Rühl EH, Nicholas PR. Shiraz berry size in relation to seed number and implications for juice and wine composition. Aust J Grape Wine Res. 2005;11: 2–8. doi:10.1111/j.1755-0238.2005.tb00273.x

20. Tello J, Ibáñez J. What do we know about grapevine bunch compactness? A state-of-the-art review. Aust J Grape Wine Res. 2018;24: 6–23. doi:10.1111/ajgw.12310

21. Torres-Lomas E, Lado-Bega J, Garcia-Zamora G, Diaz-Garcia L. Segment anything for comprehensive analysis of grapevine cluster architecture and berry properties. Plant Phenomics. 2024;6. doi:10.34133/plantphenomics.0202

22. Doligez A, Bertrand Y, Farnos M, Grolier M, Romieu C, Esnault F, et al. New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). BMC Plant Biol. 2013;13: 217. doi:10.1186/1471-2229-13-217

23. Viana AP, Riaz S, Walker MA. Genetic dissection of agronomic traits within a segregating population of breeding table grapes. Genet Mol Res. 2013;12: 951–964. doi:10.4238/2013.April.2.11

24. Houel C, Chatbanyong R, Doligez A, Rienth M, Foria S, Luchaire N, et al. Identification of stable QTLs for vegetative and reproductive traits in the microvine (*Vitis vinifera* L.) using the 18 K Infinium chip. BMC Plant Biol. 2015;15: 205. doi:10.1186/s12870-015-0588-0

25. Zhao YH, Guo YS, Lin H, Liu ZD, Ma HF, Guo XW, et al. Quantitative trait locus analysis of grape weight and soluble solid content. Genet Mol Res. 2015;14: 9872–9881. doi:10.4238/2015.August.19.21

26. Ban Y, Mitani N, Sato A, Kono A, Hayashi T. Genetic dissection of quantitative trait loci for berry traits in interspecific hybrid grape (*Vitis labruscana* × *Vitis vinifera*). Euphytica. 2016;211: 295–310. doi:10.1007/s10681-016-1737-8

27. Richter R, Gabriel D, Rist F, Töpfer R, Zyprian E. Identification of co-located QTLs and genomic regions affecting grapevine cluster architecture. Theor Appl Genet. 2019;132: 1159–1177. doi:10.1007/s00122-018-3269-1

28. Underhill A, Hirsch C, Clark M. Image-based phenotyping identifies quantitative trait loci for cluster compactness in grape. J Am Soc Hortic Sci. 2020;145: 363–373. doi:10.21273/JASHS04932-20

29. Fanizza G, Lamaj F, Costantini L, Chaabane R, Grando MS. QTL analysis for fruit yield components in table grapes (*Vitis vinifera*). Theor Appl Genet. 2005;111: 658–664. doi:10.1007/s00122-005-2016-6

30. Correa J, Ravest G, Laborie D, Mamani M, Torres E, Muñoz C, et al. Quantitative trait loci for the response to gibberellic acid of berry size and seed mass in table grape (*Vitis vinifera* L.). Aust J Grape Wine Res. 2015;21: 496–507. doi:10.1111/ajgw.12141

31. Cabezas JA, Cervera MT, Ruiz-García L, Carreño J, Martínez-Zapater JM. A genetic analysis of seed and berry weight in grapevine. Gustafson JP, editor. Genome. 2006;49: 1572–1585. doi:10.1139/g06-122

32. Mejía N, Gebauer M, Muñoz L, Hewstone N, Muñoz C, Hinrichsen P. Identification of QTLs for seedlessness, berry size, and ripening date in a seedless x seedless table grape progeny. Am J Enol Vitic. 2007;58: 499–507. doi:10.5344/ajev.2007.58.4.499

33. Costantini L, Battilana J, Lamaj F, Fanizza G, Grando M. Berry and phenology-related traits in grapevine (*Vitis vinifera* L.): From quantitative trait loci to underlying genes. BMC Plant Biol. 2008;8: 38. doi:10.1186/1471-2229-8-38

34. Doligez A, Bouquet A, Danglot Y, Lahogue F, Riaz S, Meredith C, et al. Genetic mapping of grapevine (*Vitis vinifera* L.) applied to the detection of QTLs for seedlessness and berry weight. Theor Appl Genet. 2002;105: 780–795. doi:10.1007/s00122-002-0951-z

35. Vezzulli S, Doligez A, Bellin D. Molecular mapping of grapevine genes. 1st ed. In: Cantu D,

Walker MA, editors. The Grape Genome. 1st ed. Cham: Springer Nature; 2019. pp. 103–136. doi:10.1007/978-3-030-18601-2

36. Moura-furlan MF, de Oliveira GL, Rodrigues CS, Paioli-Pires EJ. The role of plant breeding in grapevine production. Advances in Botanical Research. Elsevier; 2024. pp. 1–40. doi:10.1016/bs.abr.2024.01.001

37. Flutre T, Le Cunff L, Fodor A, Launay A, Romieu C, Berger G, et al. A genome-wide association and prediction study in grapevine deciphers the genetic architecture of multiple traits and identifies genes under many new QTLs. G3 Genes, Genomes, Genet. 2022;12. doi:10.1093/g3journal/jkac103

38. Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK, et al. Genetic structure and domestication history of the grape. Proc Natl Acad Sci. 2011;108: 3530–3535. doi:10.1073/pnas.1009363108

39. Laucou V, Launay A, Bacilieri R, Lacombe T, Andre MT De, Hausmann L, et al. Extended diversity analysis of cultivated grapevine *Vitis vinifera* with 10K genome-wide SNPs. PLoS One. 2018;13: 1–27. doi:https://doi.org/10.1371/journal.pone.0192540

40. Guo DL, Zhao HL, Li Q, Zhang GH, Jiang JF, Liu CH, et al. Genome-wide association study of berry-related traits in grape [*Vitis vinifera* L.] based on genotyping-by-sequencing markers. Hortic Res. 2019;6. doi:10.1038/s41438-018-0089-z

41. Zhang H, Fan X, Zhang Y, Jiang J, Liu C. Identification of favorable SNP alleles and candidate genes for seedlessness in *Vitis vinifera* L. using genome-wide association mapping. Euphytica. 2017;213: 1–13. doi:10.1007/s10681-017-1919-z

42. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, et al. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell. 2009;21: 2194–2202. doi:10.1105/tpc.109.068437

43. Zadokar A, Sharma P, Sharma R. Comprehensive insights on association mapping in perennial

fruit crops breeding – Its implications, current status and future perspectives. Plant Sci. 2025;350: 112281. doi:10.1016/j.plantsci.2024.112281

44. Thorat KD, Upadhyay A, Samarth RR, Machchhindra SR, Jagtap MA, Kushwaha K, et al. Genome-wide association analysis to identify genomic regions and predict candidate genes for bunch traits in grapes (*Vitis vinifera* L.). Sci Hortic (Amsterdam). 2024;328. doi:10.1016/j.scienta.2024.112882

45. García-Abadillo J, Barba P, Carvalho T, Sosa-Zuniga V, Lozano R, Carvalho HF, et al. Dissecting the complex genetic basis of pre and post-harvest traits in *Vitis vinifera* L . using genome-wide association studies . Hortic Res. 2024. doi:10.1093/hr/uhad283

46. Marrano A, Micheletti D, Lorenzi S, Neale D, Grando MS. Genomic signatures of different adaptations to environmental stimuli between wild and cultivated *Vitis vinifera* L. Hortic Res. 2018;5: 34. doi:10.1038/s41438-018-0041-2

47. Foria S, Magris G, Jurman I, Schwope R, De Candido M, De Luca E, et al. Extent of wild-to-crop interspecific introgression in grapevine (*Vitis vinifera*) as a consequence of resistance breeding and implications for the crop species definition. Hortic Res. 2022;9. doi:10.1093/hr/uhab010

48. Amyotte B, Bowen AJ, Banks T, Rajcan I, Somers DJ. Mapping the sensory perception of apple using descriptive sensory evaluation in a genome wide association study. PLoS One. 2017;12. doi:10.1371/journal.pone.0171710

49. Morgil H, Can Gercek Y, Tulum I. Single nucleotide polymorphisms (SNPs) in plant genetics and breeding. The recent topics in genetic polymorphisms. IntechOpen; 2020. doi:10.5772/intechopen.91886

50. Maia JDG, Camargo UA, Tonietto J, Zanus MC, Quecini V, Ferreira ME, et al. Grapevine breeding programs in Brazil. In: Reynolds A, editor. Grapevine breeding programs for the wine industry. Elsevier; 2015. pp. 247–271. doi:10.1016/B978-1-78242-075-0.00011-9

51. de Oliveira GL, de Souza AP, de Oliveira FA, Zucchi MI, de Souza LM, Moura MF. Genetic structure and molecular diversity of Brazilian grapevine germplasm: Management and use in breeding programs. Goel S, editor. PLoS One. 2020;15: e0240665. doi:10.1371/journal.pone.0240665

52. Alvares CA, Stape JL, Sentelhas PC, de Moraes Gonçalves JL, Sparovek G. Köppen's climate classification map for Brazil. Meteorol Zeitschrift. 2013;22: 711–728. doi:10.1127/0941-2948/2013/0507

53. Santos HG, Jacomine PKT, Anjos LHC, Oliveira VA, Lumbreras JF, Coelho MR, et al. Sistema brasileiro de classificação de solos. 5th ed. Brasília, DF, Brazil: Embrapa; 2018.

54. de Oliveira GL, Niederauer GF, de Oliveira FA, Rodrigues CS, Hernandes JL, de Souza AP, et al. Genetic diversity, population structure and parentage analysis of Brazilian grapevine hybrids after half a century of genetic breeding. Sci Hortic (Amsterdam). 2023;311. doi:10.1016/j.scienta.2023.111825

55. Fox J, Weisberg S. An R companion to applied regression. Third. Thousand Oaks, CA: Sage; 2019.

56. Covarrubias-Pazaran G. Genome-assisted prediction of quantitative traits using the R package sommer. PLoS One. 2016;11: e0156744. doi:10.1371/journal.pone.0156744

57. Di Guardo M, Micheletti D, Bianco L, Koehorst-van Putten HJJ, Longhi S, Costa F, et al. ASSIsT: an automatic SNP scoring tool for in- and outbreeding species. Bioinformatics. 2015;31: 3873–3874. doi:10.1093/bioinformatics/btv446

58. Granato I, Fritsche-Neto R. snpReady: Preparing genotypic datasets in order to run genomic analysis. 2018 [cited 15 May 2023]. Available: https://cran.r-project.org/web/packages/snpReady/index.html

59. Hastie T, Tibshirani R, Narasimhan B. impute: Imputation for microarray data. 2020 [cited 10 Feb 2023] pp. 1–5. Available: https://git.bioconductor.org/packages/impute

60. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81: 559–575. doi:10.1086/519795

61. Villanueva RAM, Chen ZJ. ggplot2: Elegant graphics for data analysis (2nd ed.). meas interdiscip res perspect. 2019;17: 160–167. doi:10.1080/15366367.2019.1565254

62. Vos PG, Paulo MJ, Voorrips RE, Visser RGF, van Eck HJ, van Eeuwijk FA. Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. Theor Appl Genet. 2017;130: 123–135. doi:10.1007/s00122-016-2798-8

63. Wang J, Zhang Z. GAPIT Version 3: Boosting Power and Accuracy for Genomic Association and Prediction. Genomics, Proteomics Bioinforma. 2021;19: 629–640. doi:10.1016/j.gpb.2021.08.005

64. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006;38: 904–909. doi:10.1038/ng1847

65. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91: 4414–4423. doi:10.3168/jds.2007-0980

66. Liu X, Huang M, Fan B, Buckler ES, Zhang Z. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. Listgarten J, editor. PLOS Genet. 2016;12: e1005767. doi:10.1371/journal.pgen.1005767

67. Huang M, Liu X, Zhou Y, Summers RM, Zhang Z. BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. Gigascience. 2019;8. doi:10.1093/gigascience/giy154

68. Kaler AS, Gillman JD, Beissinger T, Purcell LC. Comparing different statistical models and multiple testing corrections for association mapping in soybean and maize. Front Plant Sci. 2020;10. doi:10.3389/fpls.2019.01794

69. Cebeci Z, Bayraktar M, Gökçe G. Comparison of the statistical methods for genome-wide association studies on simulated quantitative traits of domesticated goats (*Capra hircus* L.). Small Rumin Res. 2023;227: 107053. doi:10.1016/j.smallrumres.2023.107053

70. Sandhu KS, Burke AB, Merrick LF, Pumphrey MO, Carter AH. Comparing performances of different statistical models and multiple threshold methods in a nested association mapping population of wheat. Front Plant Sci. 2024;15. doi:10.3389/fpls.2024.1460353

71. Prateep-Na-Thalang N, Tongyoo P, Phumichai C, Duangjit J. Comparing different statistical models for association mapping and genomic prediction of fruit quality traits in tomato. Sci Hortic (Amsterdam). 2024;327: 112838. doi:10.1016/j.scienta.2023.112838

72. Chao J, Li Z, Sun Y, Aluko OO, Wu X, Wang Q, et al. MG2C: a user-friendly online tool for drawing genetic maps. Mol Hortic. 2021;1: 16. doi:10.1186/s43897-021-00020-x

73. Velt A, Frommer B, Blanc S, Holtgräwe D, Duchêne É, Dumas V, et al. An improved reference of the grapevine genome reasserts the origin of the PN40024 highly homozygous genotype. Whiteman N, editor. G3 Genes|Genomes|Genetics. 2023. doi:10.1093/g3journal/jkad067

74. Skinner ME, Uzilov A V., Stein LD, Mungall CJ, Holmes IH. JBrowse: A next-generation genome browser. Genome Res. 2009;19: 1630–1638. doi:10.1101/gr.094607.109

75. Navarro-Payá D, Santiago A, Orduña L, Zhang C, Amato A, D'Inca E, et al. The grape gene reference catalogue as a standard resource for gene selection and genetic improvement. Front Plant Sci. 2022;12. doi:10.3389/fpls.2021.803977

76. UniProtConsortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47: D506–D515. doi:10.1093/nar/gky1049

77. Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods. 2013;9: 29. doi:10.1186/1746-4811-9-29

78. Hall D, Tegstrom C, Ingvarsson PK. Using association mapping to dissect the genetic basis of

complex traits in plants. Brief Funct Genomics. 2010;9: 157–165. doi:10.1093/bfgp/elp048

79. Hu L, Xu T, Cai Y, Qin Y, Zheng Q, Chen T, et al. Identifying candidate genes for grape (*Vitis vinifera* L.) fruit firmness through genome-wide association studies. J Agric Food Chem. 2025;73: 8413–8425. doi:10.1021/acs.jafc.5c00085

80. Zhang C, Yang Y, Zhang S, Yadav V, Zhong H, Zhang F, et al. Mining candidate genes for grape seed traits based on a genome-wide association study. Hortic Plant J. 2025;11: 1847–1864. doi:10.1016/j.hpj.2024.02.015

81. Wan Y, Schwaninger HR, Baldo AM, Labate JA, Zhong G-Y, Simon CJ. A phylogenetic analysis of the grape genus (*Vitis* L.) reveals broad reticulation and concurrent diversification during neogene and quaternary climate change. BMC Evol Biol. 2013;13: 141. doi:10.1186/1471-2148-13-141

82. Klein LL, Miller AJ, Ciotir C, Hyma K, Uribe-Convers S, Londo J. High-throughput sequencing data clarify evolutionary relationships among North American *Vitis* species and improve identification in USDA *Vitis* germplasm collections. Am J Bot. 2018;105: 215–226. doi:10.1002/ajb2.1033

83. Zecca G, Labra M, Grassi F. Untangling the evolution of american wild grapes: Admixed species and how to find them. Front Plant Sci. 2020;10: 1–17. doi:10.3389/fpls.2019.01814

84. Riaz S, Tenscher AC, Heinitz CC, Huerta-Acosta KG, Andrew Walker M. Genetic analysis reveals an east-west divide within North American *Vitis* species that mirrors their resistance to Pierce’s disease. PLoS One. 2020;15: 7–10. doi:10.1371/journal.pone.0243445

85. Cipriani G, Spadotto A, Jurman I, Gaspero G Di, Crespan M, Meneghetti S, et al. The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. Theor Appl Genet. 2010;121: 1569–1585. doi:10.1007/s00122-010-1411-9

86. Myles S, Chia J-M, Hurwitz B, Simon C, Zhong GY, Buckler E, et al. Rapid genomic

characterization of the genus *Vitis*. PLoS One. 2010;5: e8219. doi:10.1371/journal.pone.0008219

87. Zhang C, Cui L, Fang J. Genome-wide association study of the candidate genes for grape berry shape-related traits. BMC Plant Biol. 2022;22: 42. doi:10.1186/s12870-022-03434-x

88. Rodrigues CS, Moura MF, de Oliveira GL, da Silva MJR, Tecchio MA. Best linear unbiased prediction in combination with path analysis in processing grapes. Sci Agric. 2024;81. doi:10.1590/1678-992X-2022-0218

89. Migicovsky Z, Sawler J, Gardner KM, Aradhya MK, Prins BH, Schwaninger HR, et al. Patterns of genomic and phenomic diversity in wine and table grapes. Hortic Res. 2017;4: 17035. doi:10.1038/hortres.2017.35

90. Abiri K, Rezaei M, Tahanian H, Heidari P, Khadivi A. Morphological and pomological variability of a grape (*Vitis vinifera* L.) germplasm collection. Sci Hortic (Amsterdam). 2020;266: 109285. doi:10.1016/j.scienta.2020.109285

91. Wu Y, Wang Y, Fan X, Zhang Y, Jiang J, Sun L, et al. QTL mapping for berry shape based on a high-density genetic map constructed by whole-genome resequencing in grape. Hortic Plant J. 2023;9: 729–742. doi:10.1016/j.hpj.2022.11.005

92. Muñoz-Espinoza C, Di Genova A, Sánchez A, Correa J, Espinoza A, Meneses C, et al. Identification of SNPs and InDels associated with berry size in table grapes integrating genetic and transcriptomic approaches. BMC Plant Biol. 2020;20: 1–21. doi:10.1186/s12870-020-02564-4

93. Mejía N, Soto B, Guerrero M, Casanueva X, Houel C, de los Ángeles Miccono M, et al. Molecular, genetic and transcriptional evidence for a role of VvAGL11 in stenospermocarpic seedlessness in grapevine. BMC Plant Biol. 2011;11: 57. doi:10.1186/1471-2229-11-57

94. Chialva C, Eichler E, Grissi C, Muñoz C, Gomez-Talquenca S, Martínez-Zapater JM, et al. Expression of grapevine AINTEGUMENTA-like genes is associated with variation in ovary

and berry size. Plant Mol Biol. 2016;91: 67–80. doi:10.1007/s11103-016-0443-1

95. Al-Saif AM, Fahmy MA, Baghdady GA, El-Razik AMA, Kabsha EA, Farouk MH, et al. The impact of bud load on berry quality, yield, and cluster compactness in H4 strain grapevines. Agronomy. 2023;13: 2431. doi:10.3390/agronomy13092431

96. Pedro Júnior MJ, Hernandes JL, Bardin-Camparotto L, Blain GC. Plant parameters and must composition of 'Syrah' grapevine cultivated under sequential summer and winter growing seasons. Bragantia. 2017;76: 345–351. doi:10.1590/1678-4499.146

97. Vafaee Y, Ghaderi N, Khadivi A. Morphological variation and marker-fruit trait associations in a collection of grape (*Vitis vinifera* L.). Sci Hortic (Amsterdam). 2017;225: 771–782. doi:10.1016/j.scienta.2017.08.007

98. Khadivi-Khub A, Salimpour A, Rasouli M. Analysis of grape germplasm from Iran based on fruit characteristics. Rev Bras Bot. 2014;37: 105–113. doi:10.1007/s40415-014-0054-5

99. Azzi L, Deluche C, Gévaudant F, Frangne N, Delmas F, Hernould M, et al. Fruit growth-related genes in tomato. J Exp Bot. 2015;66: 1075–1086. doi:10.1093/jxb/eru527

100. Zhao X, Muhammad N, Zhao Z, Yin K, Liu Z, Wang L, et al. Molecular regulation of fruit size in horticultural plants: A review. Sci Hortic (Amsterdam). 2021;288. doi:10.1016/j.scienta.2021.110353

101. Hussain Q, Shi J, Scheben A, Zhan J, Wang X, Liu G, et al. Genetic and signalling pathways of dry fruit size: targets for genome editing based crop improvement. Plant Biotechnol J. 2020;18: 1124–1140. doi:10.1111/pbi.13318

102. Zhao J, Li Y, Ding L, Yan S, Liu M, Jiang L, et al. Phloem transcriptome signatures underpin the physiological differentiation of the pedicel, stalk and fruit of cucumber (*Cucumis sativus* L.). Plant Cell Physiol. 2016;57: 19–34. doi:10.1093/pcp/pcv168

103. Pattison RJ, Csukasi F, Catalá C. Mechanisms regulating auxin action during fruit development. Physiol Plant. 2014;151: 62–72. doi:10.1111/ppl.12142

104. Liao X, Li M, Liu B, Yan M, Yu X, Zi H, et al. Interlinked regulatory loops of ABA catabolism and biosynthesis coordinate fruit growth and ripening in woodland strawberry. Proc Natl Acad Sci. 2018;115. doi:10.1073/pnas.1812575115

105. de Jong M, Mariani C, Vriezen WH. The role of auxin and gibberellin in tomato fruit set. J Exp Bot. 2009;60: 1523–1532. doi:10.1093/jxb/erp094

106. Godoy F, Kühn N, Muñoz M, Marchandon G, Gouthu S, Deluc L, et al. The role of auxin during early berry development in grapevine as revealed by transcript profiling from pollination to fruit set. Hortic Res. 2021;8: 1–15. doi:10.1038/s41438-021-00568-1

107. Benková E, Michniewicz M, Sauer M, Teichmann T, Seifertová D, Jürgens G, et al. Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell. 2003;115: 591–602. doi:10.1016/S0092-8674(03)00924-3

108. Scarpella E, Marcos D, Friml J, Berleth T. Control of leaf vascular patterning by polar auxin transport. Genes Dev. 2006;20: 1015–1027. doi:10.1101/gad.1402406

109. Korasick DA, Westfall CS, Lee SG, Nanao MH, Dumas R, Hagen G, et al. Molecular basis for Auxin Response Factor protein interaction and the control of auxin response repression. Proc Natl Acad Sci. 2014;111: 5427–5432. doi:10.1073/pnas.1400074111

110. Li S-B, Xie Z-Z, Hu C-G, Zhang J-Z. A review of auxin response factors (ARFs) in plants. Front Plant Sci. 2016;7. doi:10.3389/fpls.2016.00047

111. Yan M, Yan Y, Wang P, Wang Y, Piao X, Di P, et al. Genome-wide identification and expression analysis of auxin response factor (ARF) gene family in panax ginseng indicates its possible roles in root development. Plants. 2023;12: 3943. doi:10.3390/plants12233943

112. Li L, Chen X. Auxin regulation on crop: from mechanisms to opportunities in soybean breeding. Mol Breed. 2023;43: 16. doi:10.1007/s11032-023-01361-9

113. Ravest G, Mamani M, Giacomelli L, Moser C, Pastenes C, Hinrichsen P. Bioactive gibberellins show differential abundance at key phenological stages for berry growth in table

grapes. Am J Enol Vitic. 2017;68: 478–484. doi:10.5344/ajev.2017.16036

114. Suehiro Y, Mochida K, Tsuma M, Yasuda Y, Itamura H, Esumi T. Effects of gibberellic acid/cytokinin treatments on berry development and maturation in the yellow-green skinned ‘Shine Muscat’ grape. Hortic J. 2019;88: 202–213. doi:10.2503/hortj.UTD-046

115. Casanova L, Casanova R, Moret A, Agustí M. The application of gibberellic acid increases berry size of “Emperatriz” seedless grape. Spanish J Agric Res. 2009;7: 919–927. doi:10.5424/sjar/2009074-1105

116. Grimplet J, Agudelo-Romero P, Teixeira RT, Martinez-Zapater JM, Fortes AM. Structural and functional analysis of the GRAS gene family in grapevine indicates a role of GRAS proteins in the control of development and stress responses. Front Plant Sci. 2016;7. doi:10.3389/fpls.2016.00353

117. Heo J-O, Chang KS, Kim IA, Lee M-H, Lee SA, Song S-K, et al. Funneling of gibberellin signaling by the GRAS transcription regulator SCARECROW-LIKE 3 in the Arabidopsis root. Proc Natl Acad Sci. 2011;108: 2166–2171. doi:10.1073/pnas.1012215108

118. Wang T-T, Yu T-F, Fu J-D, Su H-G, Chen J, Zhou Y-B, et al. Genome-wide analysis of the GRAS gene family and functional identification of GmGRAS37 in drought and salt tolerance. Front Plant Sci. 2020;11. doi:10.3389/fpls.2020.604690

119. Neves C, Ribeiro B, Amaro R, Expósito J, Grimplet J, Fortes AM. Network of GRAS transcription factors in plant development, fruit ripening and stress responses. Hortic Res. 2023;10. doi:10.1093/hr/uhad220

120. Song H, Fu X, Li J, Niu T, Shen J, Wang X, et al. Phylogenetic analysis and expression profiles of jasmonate ZIM-domain gene family provide insight into abiotic stress resistance in sunflower. Front Plant Sci. 2022;13. doi:10.3389/fpls.2022.1010404

121. Taniguchi S, Hosokawa Shinonaga Y, Tamaoki D, Yamada S, Akimitsu K, Gomi K. Jasmonate induction of the monoterpene linalool confers resistance to rice bacterial blight and

its biosynthesis is regulated by JAZ protein in rice. Plant Cell Environ. 2014;37: 451–461. doi:10.1111/pce.12169

122. Singh AP, Pandey BK, Mehra P, Heitz T, Giri J. OsJAZ9 overexpression modulates jasmonic acid biosynthesis and potassium deficiency responses in rice. Plant Mol Biol. 2020;104: 397–410. doi:10.1007/s11103-020-01047-2

123. Peethambaran PK, Glenz R, Höninger S, Shahinul Islam SM, Hummel S, Harter K, et al. Salt-inducible expression of OsJAZ8 improves resilience against salt-stress. BMC Plant Biol. 2018;18: 311. doi:10.1186/s12870-018-1521-0

124. Lal MK, Tiwari RK, Altaf MA, Kumar A, Kumar R. Editorial: Abiotic and biotic stress in horticultural crops: insight into recent advances in the underlying tolerance mechanism. Front Plant Sci. 2023;14. doi:10.3389/fpls.2023.1212982

125. Zha Q, Xi X, He Y, Jiang A. Comprehensive evaluation of heat resistance in 68 *Vitis* Germ Plasm resources. Vitis - J Grapevine Res. 2018;57: 75–81. doi:10.5073/vitis.2018.57.75-81

126. Teslić N, Zinzani G, Parpinello GP, Versari A. Climate change trends, grape production, and potential alcohol concentration in wine from the “Romagna Sangiovese” appellation area (Italy). Theor Appl Climatol. 2018;131: 793–803. doi:10.1007/s00704-016-2005-5

127. Greer DH, Weston C. Heat stress affects flowering, berry growth, sugar accumulation and photosynthesis of Vitis vinifera cv. Semillon grapevines grown in a controlled environment. Funct Plant Biol. 2010;37: 206. doi:10.1071/FP09209

128. Liu G, Chai F, Wang Y, Jiang J, Duan W, Wang Y, et al. Genome-wide identification and classification of HSF family in grape, and their transcriptional analysis under heat acclimation and heat stress. Hortic Plant J. 2018;4: 133–143. doi:10.1016/j.hpj.2018.06.001

129. DeYoung BJ, Innes RW. Plant NBS-LRR proteins in pathogen sensing and host defense. Nat Immunol. 2006;7: 1243–9. doi:10.1038/ni1410

130. van Wersch S, Tian L, Hoy R, Li X. Plant NLRs: The whistleblowers of plant immunity. plant

commun. 2020;1: 100016. doi:10.1016/j.xplc.2019.100016

131. Biswas D, Gain H, Mandal A. MYB transcription factor: A new weapon for biotic stress tolerance in plants. Plant Stress. 2023;10: 100252. doi:10.1016/j.stress.2023.100252

132. Nuruzzaman M, Sharoni AM, Kikuchi S. Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. Front Microbiol. 2013;4. doi:10.3389/fmicb.2013.00248

133. Fernandez L, Ageorges A, Torregrosa L. A putative NAP homolog specifically expressed during grapevine flower and berry development. Vitis. 2006;45: 51–52. doi:10.5073/vitis.2006.45.51-52

134. Muñoz-Espinoza C, Di Genova A, Correa J, Silva R, Maass A, González-Agüero M, et al. Transcriptome profiling of grapevine seedless segregants during berry development reveals candidate genes associated with berry weight. BMC Plant Biol. 2016;16: 1–17. doi:10.1186/s12870-016-0789-1

135. Tello J, Torres-Pérez R, Grimplet J, Carbonell-Bejerano P, Martínez-Zapater JM, Ibáñez J. Polymorphisms and minihaplotypes in the VvNAC26 gene associate with berry size variation in grapevine. BMC Plant Biol. 2015;15: 253. doi:10.1186/s12870-015-0622-2

136. Ketudat Cairns JR, Esen A. β-Glucosidases. Cell Mol Life Sci. 2010;67: 3389–3405. doi:10.1007/s00018-010-0399-2

137. Ketudat Cairns JR, Mahong B, Baiya S, Jeon J-S. β-Glucosidases: Multitasking, moonlighting or simply misunderstood? Plant Sci. 2015;241: 246–259. doi:10.1016/j.plantsci.2015.10.014

138. Kong H, Song J, Ma S, Yang J, Shao Z, Li Q, et al. Genome-wide identification and expression analysis of the glycosyl hydrolase family 1 genes in Medicago sativa revealed their potential roles in response to multiple abiotic stresses. BMC Genomics. 2024;25: 20. doi:10.1186/s12864-023-09918-w

139. Morant AV, Bjarnholt N, Kragh ME, Kjærgaard CH, Jørgensen K, Paquette SM, et al. The β -

glucosidases responsible for bioactivation of hydroxynitrile glucosides in *Lotus japonicus*. Plant Physiol. 2008;147: 1072–1091. doi:10.1104/pp.107.109512

140. Baba SA, Vishwakarma RA, Ashraf N. Functional characterization of CsBGlu12, a β-glucosidase from *Crocus sativus*, provides insights into its role in abiotic stress through accumulation of antioxidant flavonols. J Biol Chem. 2017;292: 4700–4713. doi:10.1074/jbc.M116.762161

141. Wang X, Yan L, Wang B, Qian Y, Wang Z, Wu W. Comparative proteomic analysis of grapevine rootstock in response to waterlogging stress. Front Plant Sci. 2021;12. doi:10.3389/fpls.2021.749184

142. Han X, Li Y-H, Yao M-H, Yao F, Wang Z-L, Wang H, et al. Transcriptomics reveals the effect of short-term freezing on the signal transduction and metabolism of grapevine. Int J Mol Sci. 2023;24: 3884. doi:10.3390/ijms24043884

143. Griesser M, Savoi S, Supapvanich S, Dobrev P, Vankova R, Forneck A. Phytohormone profiles are strongly altered during induction and symptom development of the physiological ripening disorder berry shrivel in grapevine. Plant Mol Biol. 2020;103: 141–157. doi:10.1007/s11103-020-00980-6

144. Li M, Feng F, Cheng L. Expression patterns of genes involved in sugar metabolism and accumulation during apple fruit development. Börnke F, editor. PLoS One. 2012;7: e33055. doi:10.1371/journal.pone.0033055

145. Cao S, Yang Z, Zheng Y. Sugar metabolism in relation to chilling tolerance of loquat fruit. Food Chem. 2013;136: 139–143. doi:10.1016/j.foodchem.2012.07.113

146. Du Z, Su Q, Wu Z, Huang Z, Bao J, Li J, et al. Genome-wide characterization of MATE gene family and expression profiles in response to abiotic stresses in rice (*Oryza sativa*). BMC Ecol Evol. 2021;21: 141. doi:10.1186/s12862-021-01873-y

147. Watanabe M, Otagaki S, Matsumoto S, Shiratake K. Genome-wide analysis of multidrug and

toxic compound extruction transporters in grape. Front Plant Sci. 2022;13. doi:10.3389/fpls.2022.892638

148. Gollhofer J, Timofeev R, Lan P, Schmidt W, Buckhout TJ. Vacuolar-iron-transporter1-like proteins mediate iron homeostasis in *Arabidopsis*. Witt SN, editor. PLoS One. 2014;9: e110468. doi:10.1371/journal.pone.0110468

149. Kato T, Kumazaki K, Wada M, Taniguchi R, Nakane T, Yamashita K, et al. Crystal structure of plant vacuolar iron transporter VIT1. Nat Plants. 2019;5: 308–315. doi:10.1038/s41477-019-0367-2

150. Yanagisawa M, Zhang C, Szymanski DB. ARP2/3-dependent growth in the plant kingdom: SCARs for life. Front Plant Sci. 2013;4. doi:10.3389/fpls.2013.00166

151. Harries PA, Pan A, Quatrano RS. Actin-related protein2/3 complex component ARPC1 is required for proper cell morphogenesis and polarized cell growth in *Physcomitrella patens*. Plant Cell. 2005;17: 2327–2339. doi:10.1105/tpc.105.033266

152. Mathur J, Mathur N, Kernebeck B, Hülskamp M. Mutations in Actin-Related Proteins 2 and 3 affect cell shape development in *Arabidopsis*. Plant Cell. 2003;15: 1632–1645. doi:10.1105/tpc.011676

153. Pratap Sahi V, Cifrová P, García-González J, Kotannal Baby I, Mouillé G, Gineau E, et al. *Arabidopsis thaliana* plants lacking the ARP2/3 complex show defects in cell wall assembly and auxin distribution. Ann Bot. 2017. doi:10.1093/aob/mcx178

154. Jiang K, Sorefan K, Deeks MJ, Bevan MW, Hussey PJ, Hetherington AM. The ARP2/3 complex mediates guard cell actin reorganization and stomatal movement in *Arabidopsis*. Plant Cell. 2012;24: 2031–2040. doi:10.1105/tpc.112.096263

155. Yanagisawa M, Desyatova AS, Belteton SA, Mallery EL, Turner JA, Szymanski DB. Patterning mechanisms of cytoskeletal and cell wall systems during leaf trichome morphogenesis. Nat Plants. 2015;1: 15014. doi:10.1038/nplants.2015.14

156. García-González J, Kebrlová Š, Semerák M, Lacek J, Kotannal Baby I, Petrášek J, et al. Arp2/3 complex is required for auxin-driven cell expansion through regulation of auxin transporter homeostasis. Front Plant Sci. 2020;11. doi:10.3389/fpls.2020.00486

157. Garrido C, Wollman F-A, Lafontaine I. The evolutionary history of peptidases involved in the processing of organelle-targeting peptides. Genome Biol Evol. 2022;14. doi:10.1093/gbe/evac101

158. Trösch R, Jarvis P. The stromal processing peptidase of chloroplasts is essential in *Arabidopsis*, with knockout mutations causing embryo arrest after the 16-cell stage. PLoS One. 2011;6: e23039. doi:10.1371/journal.pone.0023039

159. Kmiec B, Teixeira PF, Glaser E. Shredding the signal: targeting peptide degradation in mitochondria and chloroplasts. Trends Plant Sci. 2014;19: 771–778. doi:10.1016/j.tplants.2014.09.004

160. Tovar□Méndez A, Miernyk JA, Randall DD. Regulation of pyruvate dehydrogenase complex activity in plant cells. Eur J Biochem. 2003;270: 1043–1049. doi:10.1046/j.1432-1033.2003.03469.x

161. Ke J, Behal RH, Back SL, Nikolau BJ, Wurtele ES, Oliver DJ. The role of pyruvate dehydrogenase and acetyl-coenzyme a synthetase in fatty acid synthesis in developing *Arabidopsis* seeds. Plant Physiol. 2000;123: 497–508. doi:10.1104/pp.123.2.497

162. Tello J, Torres-Pérez R, Grimplet J, Ibáñez J. Association analysis of grapevine bunch traits using a comprehensive approach. Theor Appl Genet. 2016;129: 227–242. doi:10.1007/s00122-015-2623-9

163. Tello J, Aguirrezábal R, Hernáiz S, Larreina B, Montemayor MI, Vaquero E, et al. Multicultivar and multivariate study of the natural variation for grapevine bunch compactness. Aust J Grape Wine Res. 2015;21: 277–289. doi:10.1111/ajgw.12121

164. Chao H, He J, Cai Q, Zhao W, Fu H, Hua Y, et al. The expression characteristics of NPF

genes and their response to vernalization and nitrogen deficiency in rapeseed. Int J Mol Sci. 2021;22: 4944. doi:10.3390/ijms22094944

165. Wu W, Fu P, Lu J. Grapevine WRKY transcription factors. Fruit Res. 2022;2: 1–8. doi:10.48130/FruRes-2022-0010

166. Vannozzi A, Wong DCJ, Höll J, Hmmam I, Matus JT, Bogs J, et al. Combinatorial regulation of stilbene synthase genes by WRKY and MYB transcription factors in grapevine (*Vitis vinifera* L.). Plant Cell Physiol. 2018;59: 1043–1059. doi:10.1093/pcp/pcy045

167. Valletta A, Iozia LM, Leonelli F. Impact of Environmental Factors on Stilbene Biosynthesis. Plants. 2021;10: 90. doi:10.3390/plants10010090

168. Karasov TL, Chae E, Herman JJ, Bergelson J. Mechanisms to mitigate the trade-off between growth and defense. Plant Cell. 2017;29: 666–680. doi:10.1105/tpc.16.00931

169. Wang H, Huang T, Ruan Z, Tu R, He Y, Liu Q, et al. A MACPF protein OsCAD1 balances plant growth and immunity through regulating salicylic acid homeostasis in rice. Plant Cell Environ. 2025. doi:10.1111/pce.15366

170. Zhang X, Dai Y-S, Wang Y-X, Su Z-Z, Yu L-J, Zhang Z-F, et al. Overexpression of the *Arabidopsis* MACPF protein AtMACP2 promotes pathogen resistance by activating SA signaling. Int J Mol Sci. 2022;23: 8784. doi:10.3390/ijms23158784

171. Yu L, Liu D, Chen S, Dai Y, Guo W, Zhang X, et al. Evolution and expression of the membrane attack complex and perforin gene family in the Poaceae. Int J Mol Sci. 2020;21: 5736. doi:10.3390/ijms21165736

172. Chen P, Jian H, Wei F, Gu L, Hu T, Lv X, et al. Phylogenetic analysis of the membrane attack complex/perforin domain-containing proteins in *Gossypium* and the role of GhMACPF26 in cotton under cold stress. Front Plant Sci. 2021;12. doi:10.3389/fpls.2021.684227

173. Akter N, Islam MSU, Rahman MS, Zohra FT, Rahman SM, Manirujjaman M, et al. Genome-wide identification and characterization of protein phosphatase 2C (PP2C) gene family in

sunflower (*Helianthus annuus* L.) and their expression profiles in response to multiple abiotic stresses. PLoS One. 2024;19: e0298543. doi:10.1371/journal.pone.0298543

174. Boneh U, Biton I, Schwartz A, Ben-Ari G. Characterization of the ABA signal transduction pathway in *Vitis vinifera*. Plant Sci. 2012;187: 89–96. doi:10.1016/j.plantsci.2012.01.015

175. Han G, Qiao Z, Li Y, Wang C, Wang B. The roles of CCCH Zinc-finger proteins in plant abiotic stress tolerance. Int J Mol Sci. 2021;22: 8327. doi:10.3390/ijms22158327

176. Wang X, Zhong Y, Cheng Z (Max). Evolution and expression analysis of the CCCH Zinc finger gene family in *Vitis vinifera*. Plant Genome. 2014;7. doi:10.3835/plantgenome2014.05.0019

177. Chen Y, Sun A, Wang M, Zhu Z, Ouwerkerk PBF. Functions of the CCCH type zinc finger protein OsGZF1 in regulation of the seed storage protein GluB-1 from rice. Plant Mol Biol. 2014;84: 621–634. doi:10.1007/s11103-013-0158-5

178. Yan Z, Jia J, Yan X, Shi H, Han Y. *Arabidopsis* KHZ1 and KHZ2, two novel non-tandem CCCH zinc-finger and K-homolog domain proteins, have redundant roles in the regulation of flowering and senescence. Plant Mol Biol. 2017;95: 549–565. doi:10.1007/s11103-017-0667-8

179. Wang L, Xu Y, Zhang C, Ma Q, Joo S-H, Kim S-K, et al. OsLIC, a novel CCCH-type zinc finger protein with transcription activation, mediates rice architecture via brassinosteroids signaling. PLoS One. 2008;3: e3521. doi:10.1371/journal.pone.0003521

180. Zhang D, Xu Z, Cao S, Chen K, Li S, Liu X, et al. An uncanonical CCCH-tandem zinc-finger protein represses secondary wall synthesis and controls mechanical strength in rice. Mol Plant. 2018;11: 163–174. doi:10.1016/j.molp.2017.11.004

181. Raabe K, Honys D, Michailidis C. The role of eukaryotic initiation factor 3 in plant translation regulation. Plant Physiol Biochem. 2019;145: 75–83. doi:10.1016/j.plaphy.2019.10.015

182. Marintchev A, Ito T. eIF2B and the integrated stress response: a structural and mechanistic view. Biochemistry. 2020;59: 1299–1308. doi:10.1021/acs.biochem.0c00132

183. Browning KS. Plant translation initiation factors: it is not easy to be green. Biochem Soc Trans. 2004;32: 589–591. doi:10.1042/BST0320589

184. Hinnebusch AG. Molecular mechanism of scanning and start codon selection in eukaryotes. Microbiol Mol Biol Rev. 2011;75: 434–467. doi:10.1128/MMBR.00008-11

185. Dong Q, Mao K, Duan D, Zhao S, Wang Y, Wang Q, et al. Genome-wide analyses of genes encoding FK506-binding proteins reveal their involvement in abiotic stress responses in apple. BMC Genomics. 2018;19: 707. doi:10.1186/s12864-018-5097-8

186. Aghdasi B, Ye K, Resnick A, Huang A, Ha HC, Guo X, et al. FKBP12, the 12-kDa FK506-binding protein, is a physiologic regulator of the cell cycle. Proc Natl Acad Sci. 2001;98: 2425–2430. doi:10.1073/pnas.041614198

187. Vespa L, Vachon G, Berger F, Perazza D, Faure J-D, Herzog M. The immunophilin-interacting protein AtFIP37 from *Arabidopsis* is essential for plant development and is involved in trichome endoreduplication. Plant Physiol. 2004;134: 1283–1292. doi:10.1104/pp.103.028050

188. Yu Y, Li Y, Huang G, Meng Z, Zhang D, Wei J, et al. PwHAP5, a CCAAT-binding transcription factor, interacts with PwFKBP12 and plays a role in pollen tube growth orientation in *Picea wilsonii*. J Exp Bot. 2011;62: 4805–4817. doi:10.1093/jxb/err120

189. Castellano MM, del Pozo JC, Ramirez-Parra E, Brown S, Gutierrez C. Expression and stability of *Arabidopsis* CDC6 are associated with endoreplication. Plant Cell. 2001;13: 2671–2686. doi:10.1105/tpc.010329

190. Shiri Y, Solouki M, Ebrahimie E, Emamjomeh A, Zahiri J. Unraveling the transcriptional complexity of compactness in sistan grape cluster. Plant Sci. 2018;270: 198–208. doi:10.1016/j.plantsci.2018.02.011

191. Díaz-Riquelme J, Martínez-Zapater JM, Carmona MJ. Transcriptional analysis of tendril and inflorescence development in grapevine (*Vitis vinifera* L.). PLoS One. 2014;9: e92339.

doi:10.1371/journal.pone.0092339

192. Lebon G, Wojnarowiez G, Holzapfel B, Fontaine F, Vaillant-Gaveau N, Clement C. Sugars and flowering in the grapevine (*Vitis vinifera* L.). J Exp Bot. 2008;59: 2565–2578. doi:10.1093/jxb/ern135

193. Reuscher S, Akiyama M, Yasuda T, Makino H, Aoki K, Shibata D, et al. The sugar transporter inventory of tomato: genome-wide identification and expression analysis. Plant Cell Physiol. 2014;55: 1123–1141. doi:10.1093/pcp/pcu052

194. Akbar S, Hua X, Zhang Y, Liu G, Wang T, Shi H, et al. Genome-wide analysis of sugar transporter gene family in *Erianthus rufipilus* and *Saccharum officinarum*, expression profiling and identification of transcription factors. Front Plant Sci. 2025;15. doi:10.3389/fpls.2024.1502649

195. Afoufa-Bastien D, Medici A, Jeauffre J, Coutos-Thévenot P, Lemoine R, Atanassova R, et al. The *Vitis vinifera* sugar transporter gene family: phylogenetic overview and macroarray expression profiling. BMC Plant Biol. 2010;10: 245. doi:10.1186/1471-2229-10-245

196. Iocco-Corena P, Chaïb J, Torregrosa L, Mackenzie D, Thomas MR, Smith HM. VviPLATZ1 is a major factor that controls female flower morphology determination in grapevine. Nat Commun. 2021;12: 6995. doi:10.1038/s41467-021-27259-8

197. Zhang H, Liu L, Li Z, Wang S, Huang L, Lin S. PLATZ transcription factors and their emerging roles in plant responses to environmental stresses. Plant Sci. 2025;352: 112400. doi:10.1016/j.plantsci.2025.112400

198. Han X, Rong H, Tian Y, Qu Y, Xu M, Xu L. Genome-wide identification of PLATZ transcription factors in *Ginkgo biloba* L. and their expression characteristics during seed development. Front Plant Sci. 2022;13. doi:10.3389/fpls.2022.946194

199. Baucher M, Guérin C, El Jaziri M, Behr M. Integration of the plant-specific PLATZ transcription factors into gene regulatory networks controlling developmental processes. CRC

Crit Rev Plant Sci. 2024;43: 376–393. doi:10.1080/07352689.2024.2388013

200. Wang A, Hou Q, Si L, Huang X, Luo J, Lu D, et al. The PLATZ transcription factor GL6 affects grain length and number in rice. Plant Physiol. 2019;180: 2077–2090. doi:10.1104/pp.18.01574

201. Li Y, Wei K. Comparative functional genomics analysis of cytochrome P450 gene superfamily in wheat and maize. BMC Plant Biol. 2020;20: 93. doi:10.1186/s12870-020-2288-7

202. Sotelo-Silveira M, Cucinotta M, Chauvin A-L, Chávez Montes RA, Colombo L, Marsch-Martínez N, et al. Cytochrome P450 CYP78A9 is involved in Arabidopsis reproductive development. Plant Physiol. 2013;162: 779–799. doi:10.1104/pp.113.218214

203. Fang W, Wang Z, Cui R, Li J, Li Y. Maternal control of seed size by EOD3/CYP78A6 in *Arabidopsis thaliana*. Plant J. 2012;70: 929–939. doi:10.1111/j.1365-313X.2012.04907.x

204. Ma M, Wang Q, Li Z, Cheng H, Li Z, Liu X, et al. Expression of Ta CYP 78A3 , a gene encoding cytochrome P450 CYP 78A3 protein in wheat (*Triticum aestivum* L.), affects seed size. Plant J. 2015;83: 312–325. doi:10.1111/tpj.12896

205. Sun X, Cahill J, Van Hautegem T, Feys K, Whipple C, Novák O, et al. Altered expression of maize PLASTOCHRON1 enhances biomass and seed yield by extending cell division duration. Nat Commun. 2017;8: 14752. doi:10.1038/ncomms14752

206. Qi X, Liu C, Song L, Li M. Arabidopsis EOD3 homologue PaCYP78A6 affects fruit size and is involved in sweet cherry (*Prunus avium* L.) fruit ripening. Sci Hortic (Amsterdam). 2019;246: 57–67. doi:10.1016/j.scienta.2018.10.041

207. Allayee H, Farber CR, Seldin MM, Williams EG, James DE, Lusis AJ. Systems genetics approaches for understanding complex traits with relevance for human disease. Elife. 2023;12. doi:10.7554/eLife.91004

Supporting information

S1 Fig. Geographic location and schematic layout of the experimental fields of the grapevine germplasm collection used in this study. (A) Location of the experimental site in Brazil, highlighting the state of São Paulo and the municipality of Jundiaí, where the germplasm collection is established. (B) Schematic representation of the spatial arrangement of the three experimental fields (labeled A, B, and C) that compose the germplasm; the scale bar corresponds to 50 m.

S2 Fig. Box plots showing the distribution of raw phenotypic data for each trait and year. BL, berry length; BWi, berry width; BW, berry weight; CL, cluster length; CWi, cluster width; CW, cluster weight.

S3 Fig. Pearson correlation coefficients and frequency distribution histograms of BLUEs for different berry and cluster traits evaluated. Significance: ** $p \leq 0.01$, *** $p \leq 0.001$. Abbreviations: BL, berry length; BWi, berry width; BW, berry weight; CL, cluster length; CWi, cluster width; CW, cluster weight.

S4 Fig. Linkage disequilibrium (LD) decay in the grapevine diversity panel. The physical distance in kilobases (kb) between markers is shown on the X-axis, and mean linkage disequilibrium (r^2) is shown on the Y-axis. Pairwise r^2 values were calculated for SNPs within 500 kb and averaged across distance bins.

S1 Table. Genotype codes and passport data for 288 grapevine accessions used in the GWAS.

S2 Table. Best linear unbiased estimators (BLUEs) for six traits evaluated over 12 seasons. The traits included berry length (BL), berry width (BWi), berry weight (BW), cluster length (CL), cluster width (CWi), and cluster weight (CW).

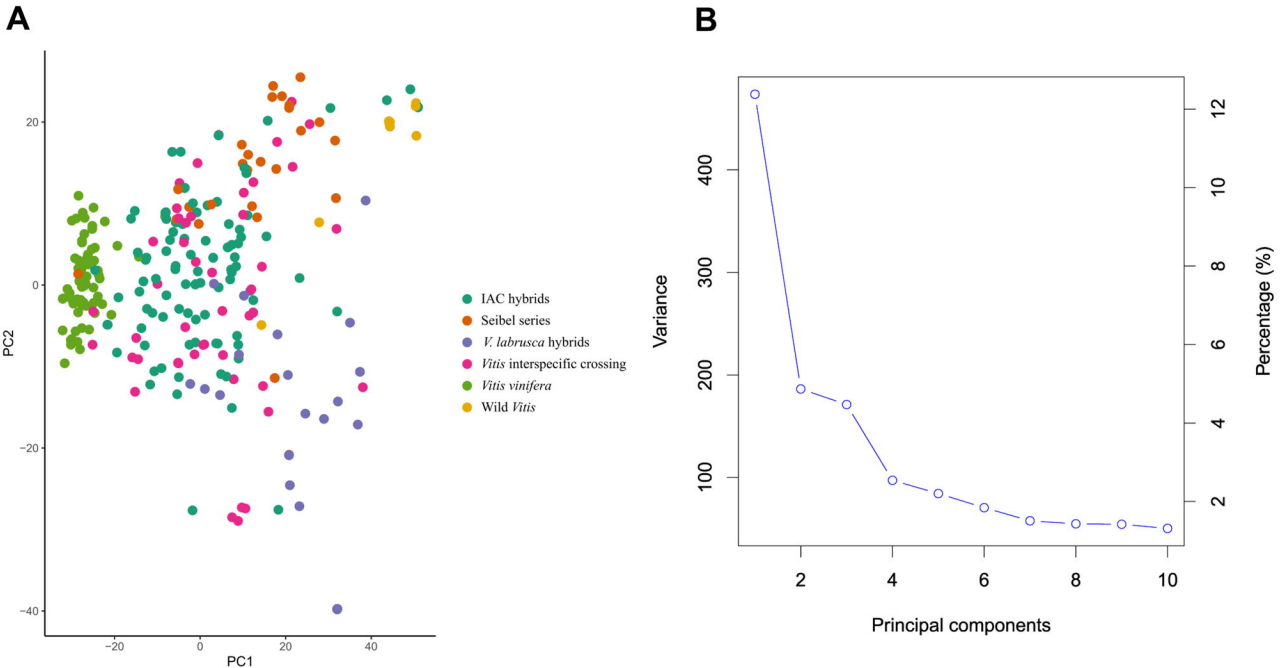
S3 Table. Estimates of variance components and heritability (H^2) for six evaluated traits, obtained from mixed model analysis considering different model terms. ModelTerm indicates the factor or interaction included in the model; VarianceComponent corresponds to the estimated variance; SE is the standard error of the estimate; Zratio is the ratio between the estimate and its standard error, used as an indicator of the relative magnitude of the component; and H^2 represents the broad-sense heritability of the trait.

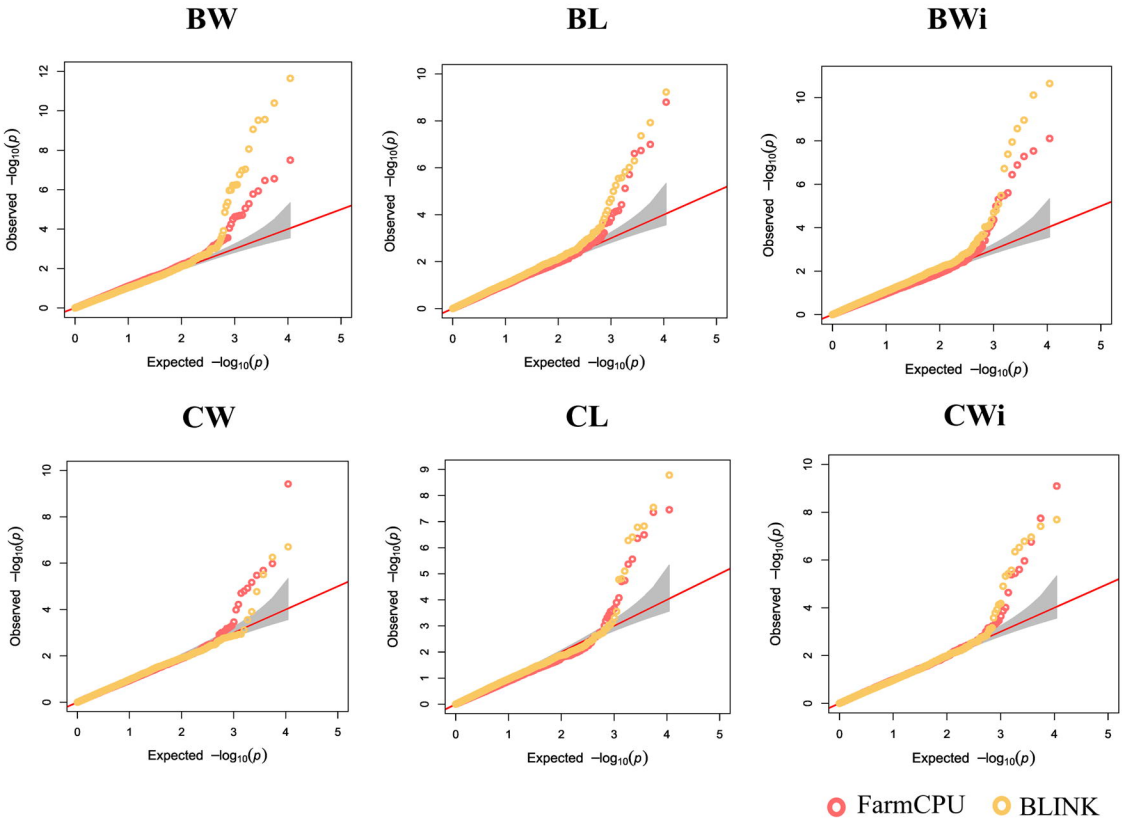
S4 Table. SNP profiles of 288 grapevine accessions from the IAC grapevine germplasm, genotyped at 11,115 high-quality SNPs used for GWAS analysis. "AA": homozygous for the dominant allele; "AB": heterozygous; "BB": homozygous for the recessive allele; "NA": missing data.

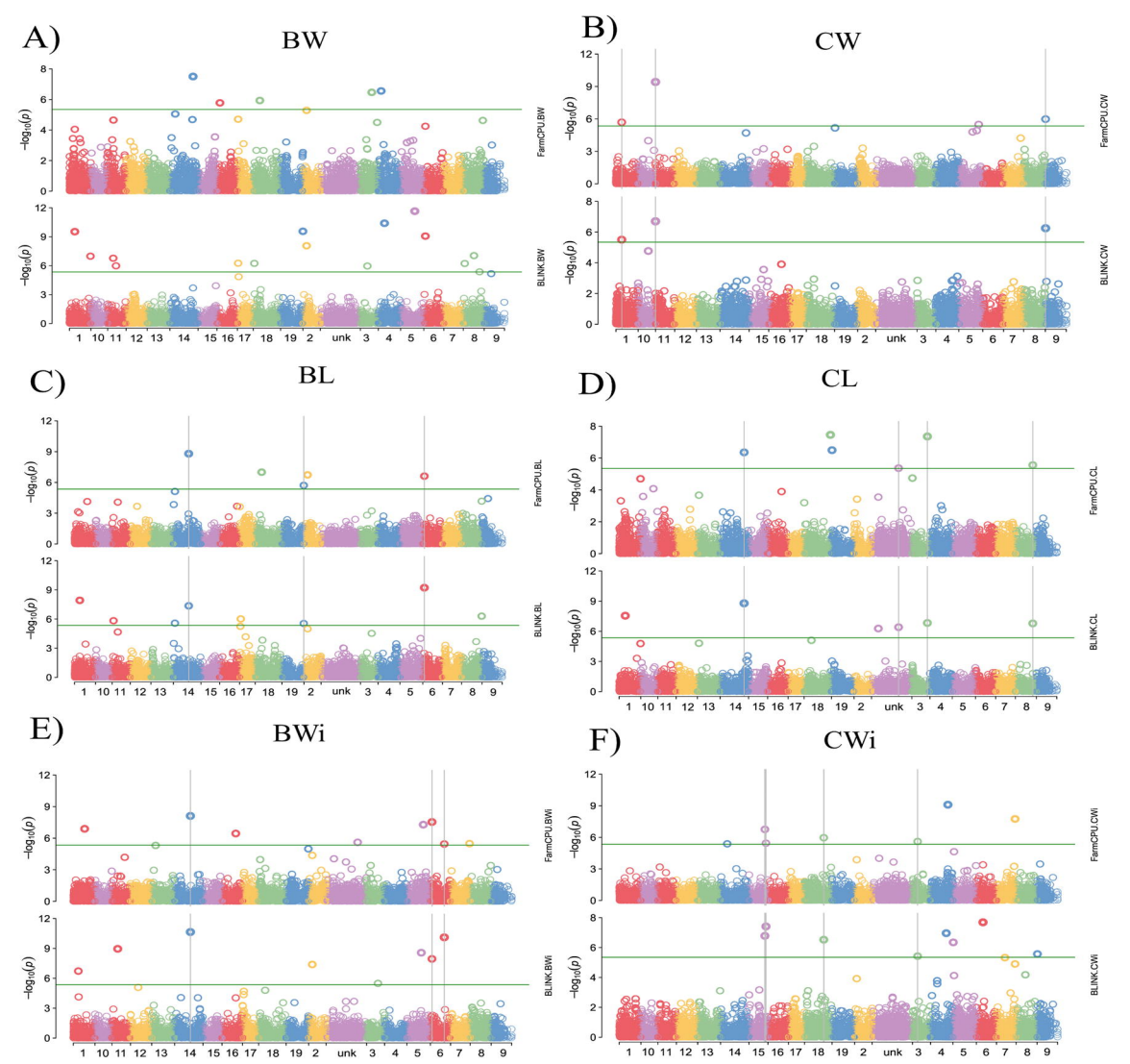
S5 Table. Comparison of significant SNP loci detected by BLINK and FarmCPU. A total of 11,115 SNPs were analyzed in the GWAS. Counts indicate loci detected by each method, those shared, and the total number of unique significant SNP loci across traits.

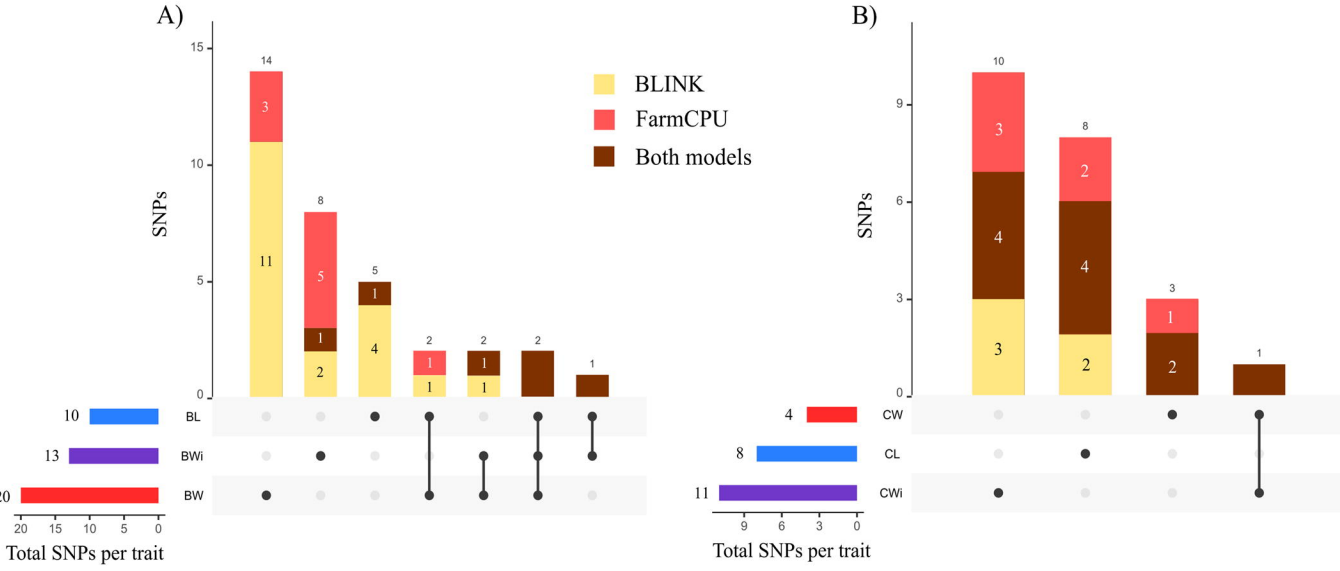
S6 Table. List of significant markers associated with berry and cluster-related traits from a genome-wide association analysis of 288 grapevine accessions using 11,115 SNPs whose MAFs were greater than 0.05.

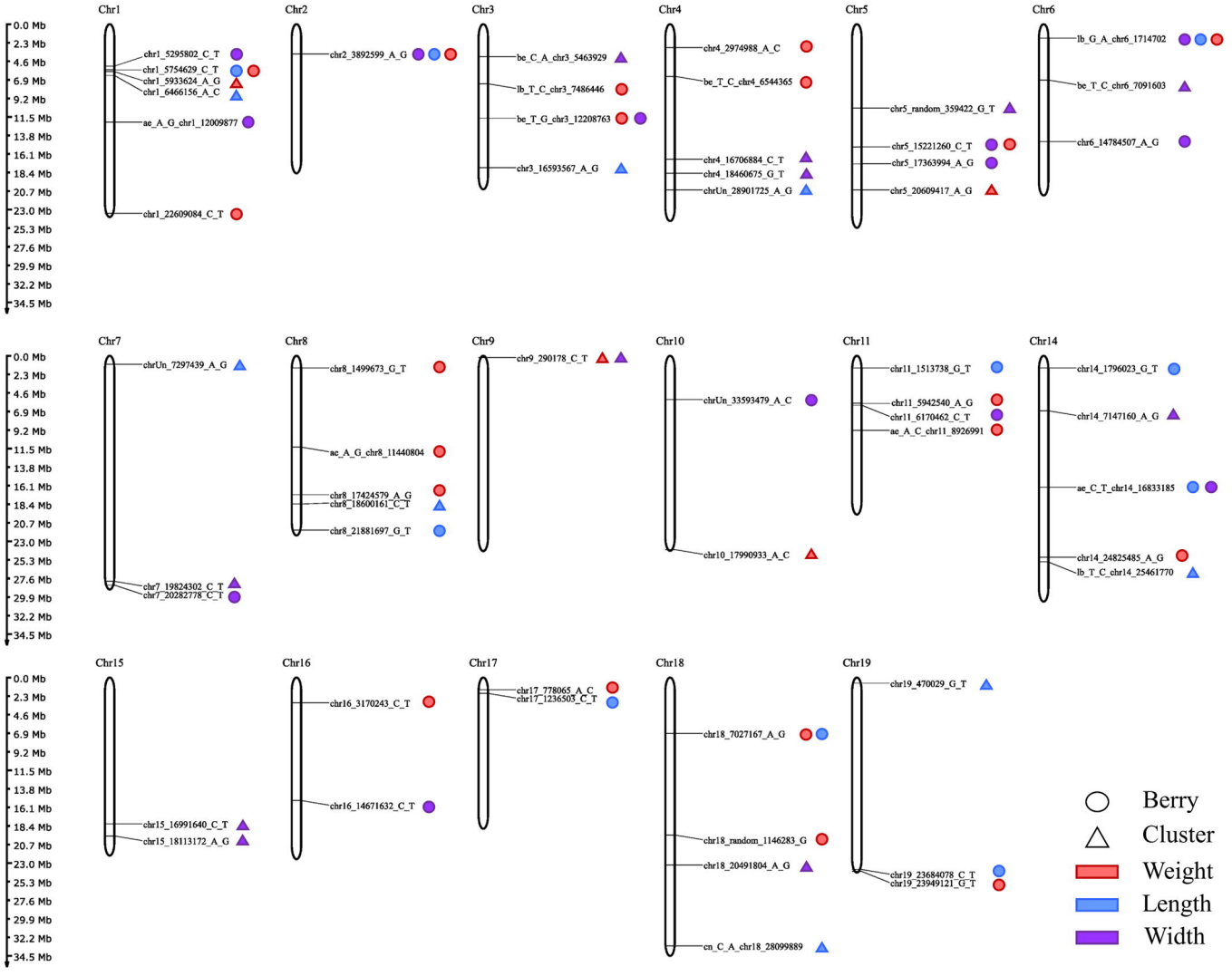
S7 Table. Details of the genes closest to significant marker–trait associations, including their physical positions, annotated functions, and GO terms.


A
PC2
PC1
20
0
−20
−40
−20
0
20
40
IAC hybrids
Seibel series
V. labrusca hybrids
Vitis interspecific crossing
Vitis vinifera
Wild Vitis
B
Variance
Percentage (%)
Principal components
100
200
300
400
2
4
6
8
10
12


BW
BL
BWi
CW
CL
CWi
Observed −log10(p)
Expected −log10(p)
FarmCPU
BLINK


A)
BW
B)
CW
C)
BL
D)
CL
E)
BWi
F)
CWi
FarmCPU.BW
BLINK.BW
FarmCPU.CW
BLINK.CW
FarmCPU.BL
BLINK.BL
FarmCPU.CL
BLINK.CL
FarmCPU.BWi
BLINK.BWi
FarmCPU.CWi
BLINK.CWi
−log10(p)
1 10 11 12 13 14 15 16 17 18 19 2 unk 3 4 5 6 7 8 9


A)
B)
BLINK
FarmCPU
Both models
SNPs
Total SNPs per trait
BL
BWi
BW
CW
CL
CWi


Chr1
chr1_5295802_C_T
chr1_5754629_C_T
chr1_5933624_A_G
chr1_6466156_A_C
ae_A_G_chr1_12009877
chr1_22609084_C_T
Chr2
chr2_3892599_A_G
Chr3
be_C_A_chr3_5463929
lb_T_C_chr3_7486446
be_T_G_chr3_12208763
chr3_16593567_A_G
Chr4
chr4_2974988_A_C
be_T_C_chr4_6544365
chr4_16706884_C_T
chr4_18460675_G_T
chrUn_28901725_A_G
Chr5
chr5_random_359422_G_T
chr5_15221260_C_T
chr5_17363994_A_G
chr5_20609417_A_G
Chr6
lb_G_A_chr6_1714702
be_T_C_chr6_7091603
chr6_14784507_A_G
Chr7
chrUn_7297439_A_G
chr7_19824302_C_T
chr7_20282778_C_T
Chr8
chr8_1499673_G_T
ae_A_G_chr8_11440804
chr8_17424579_A_G
chr8_18600161_C_T
chr8_21881697_G_T
Chr9
chr9_290178_C_T
Chr10
chrUn_33593479_A_C
chr10_17990933_A_C
Chr11
chr11_1513738_G_T
chr11_5942540_A_G
chr11_6170462_C_T
ae_A_C_chr11_8926991
Chr14
chr14_1796023_G_T
chr14_7147160_A_G
ae_C_T_chr14_16833185
chr14_24825485_A_G
lb_T_C_chr14_25461770
Chr15
chr15_16991640_C_T
chr15_18113172_A_G
Chr16
chr16_3170243_C_T
chr16_14671632_C_T
Chr17
chr17_778065_A_C
chr17_1236503_C_T
Chr18
chr18_7027167_A_G
chr18_random_1146283_G
chr18_20491804_A_G
cn_C_A_chr18_28099889
Chr19
chr19_470029_G_T
chr19_23684078_C_T
chr19_23949121_G_T
0.0 Mb
2.3 Mb
4.6 Mb
6.9 Mb
9.2 Mb
11.5 Mb
13.8 Mb
16.1 Mb
18.4 Mb
20.7 Mb
23.0 Mb
25.3 Mb
27.6 Mb
29.9 Mb
32.2 Mb
34.5 Mb
Berry
Cluster
Weight
Length
Width