






Article

Inference of the genetic basis of fruit texture in highbush blueberries using genome-wide association analyses

Luis Felipe V. Ferrão ¹, Camila Azevedo^{1,2}, Juliana Benevenuto ¹, Molla Fentie Mengist¹, Claire Luby³, Marti Pottorff⁴, Gonzalo I.P. Casorzo¹, Ted Mackey⁴, Mary Ann Lila³, Lara Giongo⁵, Nahla Bassil ⁴, Penelope Perkins-Veazie³, Massimo Iorizzo ^{3,*} and Patricio R. Munoz ^{1,*}

¹Blueberry Breeding and Genomics Lab, Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, USA

²Statistic Department, Federal University of Vicosa, Vicosa, Brazil

³Plants for Human Health Institute, North Carolina State University, Kannapolis, NC USA

⁴USDA-ARS, Horticulture Crops Research Unit, Corvallis, OR 97333, USA

⁵Fondazione Edmund Mach - Research and Innovation Centre Italy

*Corresponding authors. E-mails: miorizz@ncsu.edu, p.munoz@ufl.edu

Abstract

The global production and consumption of blueberry (*Vaccinium* spp.), a specialty crop known for its abundant bioactive and antioxidant compounds, has more than doubled over the last decade. To hold this momentum, plant breeders have begun to use quantitative genetics and molecular breeding to guide their decisions and select new cultivars that are improved for fruit quality. In this study, we leveraged our inferences on the genetic basis of fruit texture and chemical components by surveying large breeding populations from northern highbush blueberries (NHBs) and southern highbush blueberries (SHBs), the two dominant cultivated blueberries. After evaluating 1065 NHB genotypes planted at the Oregon State University, and 992 SHB genotypes maintained at the University of Florida for 17 texture-related traits, evaluated over multiple years, our contributions consist of the following: (i) we drew attention to differences between NHB and SHB materials and showed that both blueberry types can be differentiated using texture traits; (ii) we computed genetic parameters and shed light on the genetic architecture of important texture attributes, indicating that most traits had a complex nature with low to moderate heritability; (iii) using molecular breeding, we emphasized that prediction could be performed across populations; and finally (iv) the genomic association analyses pinpointed some genomic regions harboring potential candidate genes for texture that could be used for further validation studies. Altogether, the methods and approaches used here can guide future breeding efforts focused on maximizing texture improvements in blueberries.

Introduction

Blueberry (*Vaccinium corymbosum* and hybrids) is a crop with expanding production and consumption worldwide. Considered a rich source of bioactive and antioxidant compounds, blueberry's status as a "superfood" has been a major factor driving the increase of production and consumption [1]. In the USA, for example, consumption of blueberries has experienced a more than a 300% increase in per capita consumption since 2005 and is now at an all-time high of over 0.9 kg per person. Globally, blueberry production has more than doubled over the last decade.

Several species are commercially important in the *Vaccinium* genus. The *V. corymbosum* L. is the dominant cultivated species and can be further separated into northern highbush blueberries (NHBs) and southern highbush blueberries (SHBs), depending on their chilling requirements and winter hardiness. NHB is native to the Eastern and Mid-Western regions of the USA and requires more than 800 hours of chilling at 0–7°C to break bud dormancy. Commercially, NHB cultivars are grown across the Northern regions of the USA, as well as in Canada, Chile, New Zealand, Northern Europe, and China. Most of the current cultivars of NHB were developed in New Jersey, North Carolina, and Michigan; however, as production has been increasing in

the Pacific Northwest, there are now several public and private breeding programs there. Starting in 1950, the breeding program based at the University of Florida led pioneering hybridizations between NHB and species native and adapted to the southeastern USA (e.g. *Vaccinium darrowii*) aiming for better adaptation to higher temperatures and disease pressure [2]. This introgression led to a reduction in the chilling hours required to break flower bud and ultimately established a new blueberry varietal group, the southern highbush (SHB). To date, new plantings of SHB in low chilling geographic regions has been the main driver of the global expansion of blueberry production. SHB plants establish more quickly than NHB, thus producing fruit earlier. Thus, precociousness together with the early-ripening cultivars enable growers to reach an early market window, when prices are higher and competition is lower, making the crop more profitable for growers in tropical and subtropical regions of Africa, America, and Asia [3].

To maintain the marketability of both NHB and SHB cultivars, the blueberry industry seeks to meet the current demand for improved fruit quality [4]. Good texture is considered a key element, affecting both consumer appeal and postharvest decay. Despite the importance of this trait, the genetic improvements for

Received: 25 April 2024; Accepted: 4 August 2024; Published: 21 August 2024; Corrected and Typeset: 1 October 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nanjing Agricultural University. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. Texture, parameters evaluated using the texture analyzer (TA) instruments in blueberries

Method	Texture parameter	Abbreviation	Description	Unit
Penetration with 2 mm flat probe	Force at 1 mm	F1mm	Force at 1 mm depth	N
	Maximum force	FM	Force at point of skin puncture	N
	Distance to maximum force	DFM	Distance between point of contact and skin puncture	mm
	Slope maximum force	SFM	Slope between force at 1 mm and maximum force	N mm ⁻¹
	Area maximum force	AFM	Area underlying the curve between point of contact and maximum force	N mm
	Linear distance to maximum force	LDFM	Length of curve between point of contact and maximum force	-
	Mean internal force	MIF	Mean force between first peak after minimum force* and 80% strain	N
	Area internal force	AIF	Area underlying the curve between first peak after minimum force and 80% strain	N mm
	Number of internal peaks	NIP	Number of peaks between first peak after minimum force and 80% strain (threshold: 1 g)	-
	Force linear distance	FLD	Length of curve between point of contact and 80% strain	-
	Area force linear distance	AFLD	Area between point of contact and 80% strain	N mm
	Burst strain	BrSt	Strain at maximum force	%
	Young's modulus 10%	YM10PD_Cond	Slope between force at minimum time and force at 10% strain**	MPa/%
	Young's modulus 20% burst strain	YM20_BrSt	Slope between force at minimum time and force at 20% of BrSt	MPa/%
	Young's modulus 80% burst strain	YM80_BrSt	Slope between force at minimum time and force at 80% of BrSt	MPa/%
	Young's modulus 100% burst strain	YM100_BrSt	Slope between force at minimum time and force at 100% of BrSt	MPa/%
	Young's modulus 1–2%	YM1to2	Slope between force at 1% and 2% strain	MPa/%

difficulty of phenotyping texture, we demonstrated the potential of genomic prediction to be implemented at the scope of NHB and SHB breeding programs. To our knowledge, this is the first and largest study assessing the genetic architecture of different mechanical texture parameters in blueberry.

Results

Population structure and genetic diversity in cultivated blueberries

To elucidate the genetic basis of texture traits, we investigated two large blueberry populations separated into NHB and SHB types (Fig. 1a). This is the first time in the fruit literature that a such large and representative population was considered for genetic analyses of texture attributes. Therefore, we first focused on estimating some population genetic parameters. The linkage disequilibrium (LD) was computed and represented in the form of frequency by chromosome (Fig. 1b). Overall, in both groups, less than 1% of pairwise markers showed higher LD (>0.8), and a fast LD decay was observed across all the chromosomes. This LD pattern is expected since both populations are highly diverse, containing individuals in early evaluation stages in the breeding pipeline. The cross-pollination form, associated to the high inbreeding depression observed in the species, are also two important sources that leverage the genetic diversity and affect the LD results. We also investigated potential genomic regions that are differentially selected in both populations. After computing genome wide F_{st} between both blueberry groups, several regions across the chromosomes showed signatures of selection (Fig. 1c).

The selective sweep observed at the end of the chromosome 5 had the strongest peak.

To explore the relationship between individuals within and across populations, we first evaluated the genetic diversity using principal component analyses (PCAs). The first two dimensions (PC1 and PC2) explained 56% and 4% of the structural variance, respectively. Two distinct genetic groups were identified, with SHB genotypes positioned in the quadrants I and IV; and the NHB in the quadrants II and III. To complement the within-population analyses, we relied on the probabilistic models implemented in the STRUCTURE software (Fig. 1e). For the SHB population, a sub-structured was observed within this genetic pool, in which two subgroups were identified. A large amount of admixture was observed in the NHB collection with four subgroups identified.

Texture variation in SHB and NHB populations

In this study, we investigated the genetic diversity of 17 texture parameters evaluated in two large breeding populations of NHB and SHB blueberries (Table 1). The NHB population was part of the USDA-ARS and Oregon State University Cooperative breeding program, and the population was divided into a group of 113 individuals, named here as Validation study set (VAL set), which included advanced selections and cultivars, and 952 individuals, named here as Genetic study set (GEN set), consisting of early selections derived from multiple families (Fig. 1a). Similar genetic material was used for the SHB, which were part of the breeding program at the University of Florida. SHB populations included 113 advanced selection and cultivars as part of the VAL set,

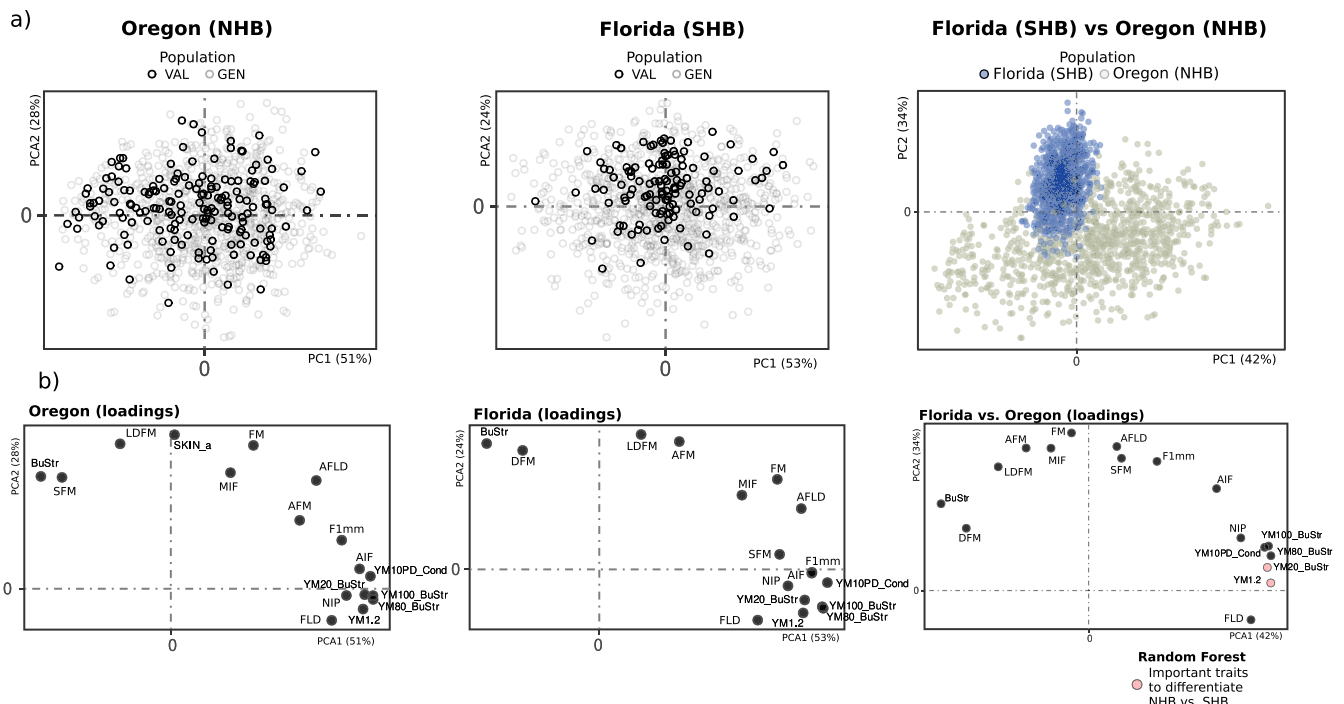


Figure 2. PCA based on 17 texture traits. (a) First two principal components evaluated for populations of Northern Highbush Blueberries (NHB) and Southern Highbush Blueberries (SHB) evaluated in Oregon and Florida, respectively. Each point represents a blueberry genotype, in which the GEN and VAL data sets are represented by different colors. Loading vectors describing how texture traits influence the variation in blueberry data from both locations separately. (b) PCA analyses and loading vectors when both NHB and SHB populations were considered simultaneously. In the loading vectors plot, the highlighted dots represent the most important traits for separating NHB and SHB populations measured via Random Forest model.

and 879 early selections as part of the GEN set (Fig. 1a). Texture attributes were evaluated using a penetration method implemented in a TA.XTPlus Texture Analyzer (Stable Micro Systems, Hamilton, MA, USA) as previously described by [11]. A total of 17 texture-related metrics were retrieved from the TA and used in this study.

In terms of phenotypic variation for texture-related traits, there was considerable diversity for all traits. On average, individuals from the SHB population showed larger values for the difference metrics computed via TA than NHB (Supplemental Material S1), with genotypes in the GEN set, presenting higher phenotypic variation than in VAL set. After combining information from both populations, we assessed the phenotypic diversity of texture attributes using multivariate analyses. In the PCA, we observed a clear distinction between SHB and NHB genotypes (Fig. 2a). Regarding trait importance, the texture attribute related to BuStr and all Young's modulus (YM) trait-related parameters showed high loading scores in the PCA analysis. We relied on unsupervised analyses to identify the key traits to separate both populations. Using Random Forest, we found that four Young's modulus trait-related parameters (YM_BuStr and YM1.2) are the most important parameters to separate SHB and NHB genotypes (Fig. 2b).

The relationship between texture-related traits and traditional compression force was examined. For this, we collected 100 berries from multiple genotypes and measured, in the same order, for texture attributes using the TA machine, and compression force (or deflection) using FruitFirm 1000 machine. Among all the TA parameters, SFM showed the highest correlations (0.87) with the compression force (g/mm) by the FruitFirm 1000. The linear equation fitted to transform SFM ($N\ mm^{-1}$) from TA into compression force (g/mm) from FruitFirm 1000 was:

$y_{CF} = 38.7 + 248\beta_1$, where y_{CF} is the predicted compression force value measured in g/mm and β_1 is the SFM value reported via TA.

Texture traits are predictable and show low-to-moderate heritability values

Blueberry breeders have relied on understanding the genetic architecture of complex traits and the use of genomic prediction to guide parental selection in recurrent selection breeding programs. Genome-wide single-nucleotide polymorphisms (SNPs) were obtained using a targeted genotyping by sequencing approach (Flex-Seq) [12] yielding a total of 113,421 markers used to build a genomic relationship matrix of NHB and SHB populations combined.

For genomic selection, we first tested four prediction scenarios (Fig. 3a). Firstly, GBLUP models were fitted separately, within NHB and SHB populations (indicated as intra-prediction scenarios)—mimicking the fact the NHB and SHB breeding programs are performed independently. In a more holistic way, models were calibrated in one population and to predict the other in both directions (indicated as inter-prediction scenarios).

The genomic information was also used to estimate heritability values and set up theoretical limits for genetic improvements in both populations. Low to moderate genomic heritability values were computed across traits, highlighting the quantitative nature of texture-related traits in blueberries. In general, NHB populations showed higher heritability values than the SHB population (Fig. 3b). Interestingly, some traits presented heterogeneous values across populations, indicating that phenotypic variation in NHB and SHB populations might be under distinct genetic control or that environment has a larger effect on one population than the other. For example, NIP had higher narrow-sense heritability values in SHB when compared to NHB.

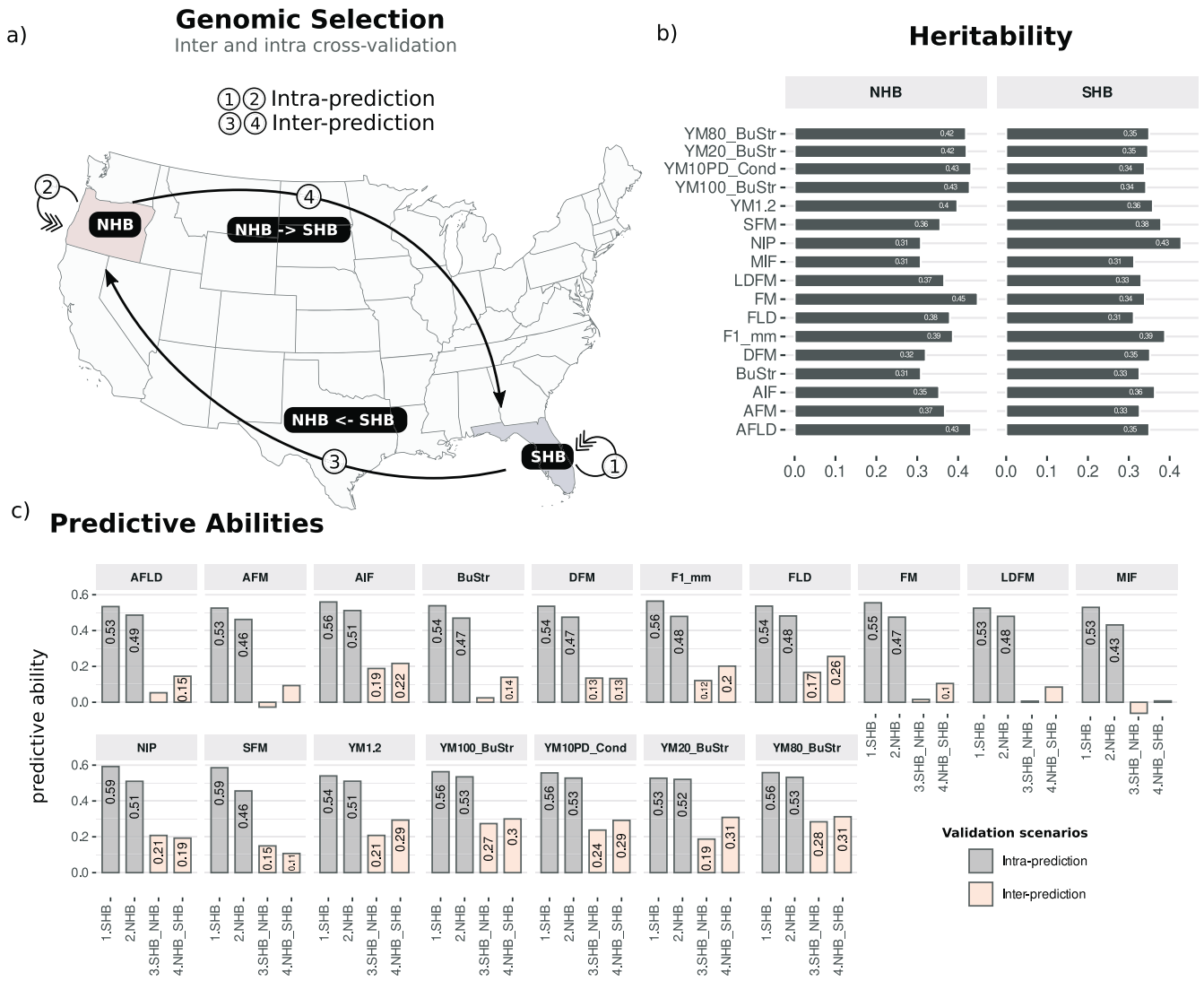


Figure 3. Genomic prediction of texture-related traits in blueberry. (a) Scenarios of cross-validation tested in Northern (NHB) and Southern (SHB) highbush blueberries from Oregon and Florida States, respectively; (b) Genomic heritability for 17 texture traits evaluated in NHB and SHB; (c) Predictive ability measured as the Person's correlation between predicted and observed phenotypic values for inter and intra-population validation scenarios.

Finally, we computed predictive abilities to test the relevance of genomic selection for blueberry texture. When comparing intra-population predictions, SHB showed better prediction abilities for all traits (Fig. 3c). We observed prediction abilities ranging from 0.53 (AFLD, AFM, LDFM, MIF, and YM20_BuStr) to 0.59 (NIP and SFM), with an average value of 0.54. For NHB, values ranged from 0.43 (MIF) to 0.53 (YM100_BuStr, YM10PD_Cond, and YM80_BuStr) with an average value of 0.48. The largest difference between predictive ability was observed for SFM, with an improvement of 0.13 observed in the SHB, when compared to NHB.

Overall, we hypothesized that NHB and SHB breeding programs could be integrated by sharing predictive models across populations. To test it, we compared inter- and intra-population analyses and observed that all prediction abilities were drastically reduced with models calibrated and tested across populations (Fig. 3b). More specifically, for some traits, inter-population will not work regardless of which population the models are calibrated. For example, low predictions were observed for LDFM, AFM, and MIF. For other traits, we could see a directional effect, with models working only when trained in NHB to predict SHB, including

AFLD, BuStr, and FM. For all other traits, we could see reasonable results for inter-population predictions, with some traits showing predictive accuracies higher than 0.25 (e.g. YM100_BuStr and YM80_BuStr).

Texture attributes are consistent across fruits and shows high repeatability values

Another important question is when multiple fruits are evaluated from the same plant, how many samples are necessary to represent a genotype? To answer this question, we estimated the repeatability values based on 10 berries collected from the same genotype, including a permanent environment effect in the mixed model. Like heritability, the repeatability coefficient is also related to the variability in the population, where lower values indicate that a larger number of technical replications (i.e., more berries per plants) must be considered for an efficient selection [13]. All repeatability values computed in both populations were moderate-to-high, indicating that the option for 10 fruits measured during the texture analyses were sufficient to describe the phenotypic variance of a given genotype (Supplemental

GWAS texture

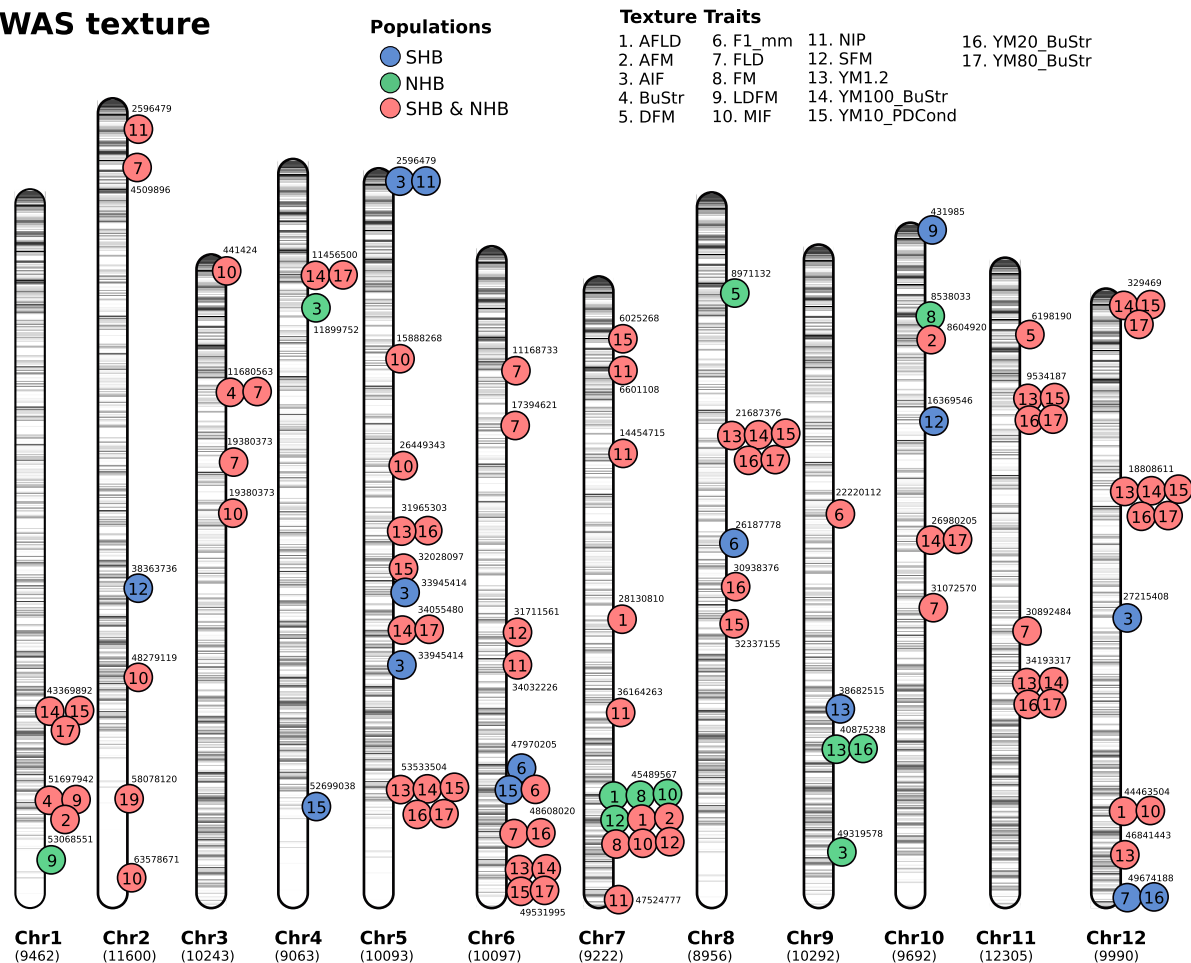


Figure 4. SNP-trait associations detected via GWAS (GWAS). In total, 17 texture-related traits were tested in two blueberry breeding populations (SHB and NHB). GWAS was carried out in the two populations independently and then combining both data sets. Numbers below each chromosome indicate the total number of SNPs. The circles represent the position of a significant marker with the lowest p-value within a window of 5 Mb in the genome. The colors and numbers inside the circles represent the population and traits, respectively, for which the association was detected as indicated in the legend. The numbers outside the circles indicate the position, in base pairs, where the marker with the lowest P value was identified.

Fig. S1). Repeatability sets the theoretical upper limits for heritability. In most cases, the heritability values were substantially lower than repeatability estimates, indicating that fruits are relatively uniform within plants, but the environmental variation is an important source of variation in blueberries.

Genomic regions were found associated with most texture traits

We performed GWAS analyses for twenty texture-related traits using the Q+K linear mixed and the additive model. Analyses were carried out for each population independently, and then combined SHB and NHB populations (Supplemental Fig. S2, S3, and S4). An illustration displaying the genomic location associated with each trait is summarized in the Fig. 4. The position of the significant markers with the lowest associated P value in the genomic region is displayed. In total, we reported 11 genomic regions for NHB, 15 for SHB, and 85 when both SHB and NHB populations were considered together in the GWAS.

Few associations were reported when both populations were evaluated either independently or when both populations were combined for the traits AFM, BuStr, DFM, FM, and LDFM. For all young modulus (YM) texture traits, many significant (>10) associations were reported. The Young modulus parameters

(YM1.2, YM10_PDCCond, YM20_BuStr, YM80_BuStr, YM100_BuStr) for example showed a high polygenic architecture with hits distributed across 9 different chromosomes. Also, QTL for these parameters clustered around the same regions in chromosome 5 (position 53 533 504), chromosome 6 (position 49 531 955), chromosome 8 (position 21 687 376), chromosome 11 (position 9 534 987 and 34 193 317), and chromosome 12 (position 329 469 and 18 808 611) which suggest similar genetic control. This is consistent with the hypothesis that given that these parameters are representing the same mechanical textural component. Chromosomes 5 and 7 showed the highest number of associations. At the end of the chromosome 7 (position 45 489 567), for example, the same genomic region pinpointed significant associations for 10 different traits. Only for the trait YM1.2, we could map a consistent region across both populations separately, which was located at chromosome 9 with two close genomic regions (position 38 682 515 and 40 875 238).

Candidate genes were found at significant GWAS regions.

Once significant genomic regions associated with texture-related traits were detected through GWAS analyses, we searched for candidate genes for the most relevant traits, named YM20_BuStr,

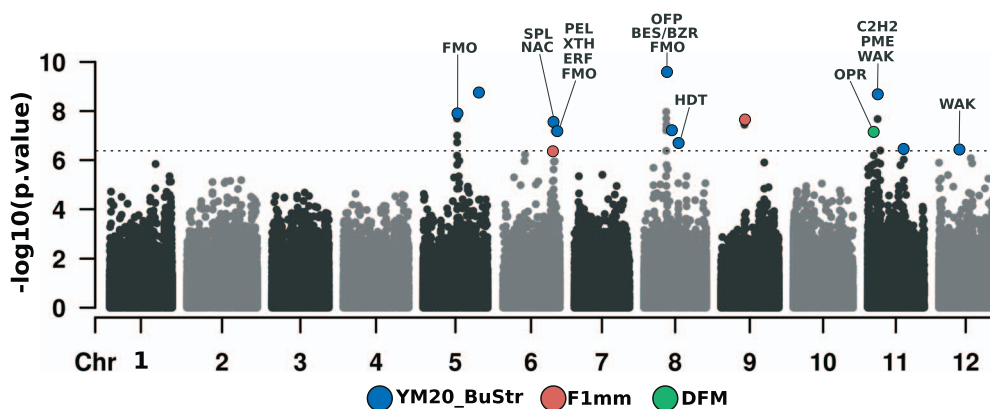


Figure 5. Overlapping Manhattan plots for GWAS hits detected for three texture-related traits (YM20_BuStr, F1mm, and DFM). The dots represent SNPs. The x-axis shows the SNP position across the 12-haploid set of chromosomes of the “W8520_P0” reference genome. The y-axis shows the statistical significance of SNP association with the traits, with dashed line representing threshold at Bonferroni-type correction. Abbreviations of plausible candidate gene names detected in significant GWAS regions are presented as follows: FMO = flavin-containing monooxygenase; NAC = NAC domain-containing protein; SPL = Squamosa promoter-binding-like protein; ERF = Ethylene-responsive transcription factor; XTH = Xyloglucan endotransglucosylase/hydrolase; PEL = Pectate lyase; BES1/BZR1 = BES1 BZR1 homolog protein; OFP = Ovate family transcription repressor; HDT = histone deacetylase; OPR = 12-oxophytodienoate reductase; WAK = wall-associated receptor kinase; PME = Pectinesterase/pectinesterase inhibitor; C2H2 = C2H2 zinc finger protein.

F1mm, and DFM (Supplementary Table S2). Given the large number of attributes generated by the TA analyses, we focused our discussion on traits that were identified when both SHB and NHB populations were analyzed together and that have a biological meaning. For example, YM20_BuSt, from the multivariate analyses was revealed as an important trait to differentiate SHB and NHB materials. At the sensory level, F1mm has a good correlation with firmness measured in the consumer panel. Finally, DFM seems an important component for post-harvest.

Genes directly or indirectly related to cell wall metabolism and fruit firmness variation in other species were identified flanking significant SNPs for YM20_BuStr and DFM traits (Fig. 5). No plausible genes could be pointed for F1mm with the functional description of the genes retrieved so far. The potential roles of these individual genes in texture modification are further elaborated in the discussion section.

Discussion

In recent decades, numerous breeding schemes have been proposed in plant breeding. In general, two main goals are addressed in outcrossing species, such as blueberry: (i) identify the best performing materials for commercial cultivar release and (ii) identify genotypes that can be used as parents in future crosses [9]. To make it more efficient, estimating genetic parameters in diverse populations provides the means to determine the genetic control of complex traits and thus guide future breeding decisions. Although extensive analyses in this area have been reported for SHB [14], and for NHB [15], to our knowledge, no scientific study has addressed populations from both types of blueberry at the same time and considered more detailed texture attributes.

Fruit quality comprises a complex series of traits that contribute to the attractiveness and suitability of fruits for consumption. Texture at maturity is arguably one of the most important attributes for blueberry consumers and retailers, since it contributes to consumer liking, resistance to bruising during harvesting and transportation, and shelf-life. So far, most of the breeding studies applied to blueberry are focused on mechanical attributes, while the relevance of acoustic and other mouthfeel properties remains elusive at this point. Studies assessing mechanical texture indicated that in blueberry multiple components are con-

tributing to texture [16]. To date, no study has been conducted in blueberry to evaluate genetic parameters associated to texture attributes. To fill this gap, here we used TA to evaluate a large diversity of texture attributes in SHB and NHB populations. Our results indicate that both firmness (also known as Chord stiffness), usually measured with the compression method with FirmTech equipment, and SFM, measured here with the TA, represent the slope of the chord drawn between any specific point on the force–deformation curve. Slight deviations from these measurements obtained with the two machines are expected since the points set by the operators are specific to each experiment and can affect the slope values [6]. Mean SFM value in SHB was higher than NHB. This result is consistent with the blueberry literature, indicating that SHB are overall firmer than NHB genotypes [17]. However to note, that SFM is not one of the parameters that best explains the difference among genotypes in the NHB and SHB populations (Fig. 1) and has not been identified as a key trait to explain sensorial texture or postharvest texture changes [16]. This highlights the importance of using methods that are more aligned with the texture measurements by the TA.

At the breeding level, molecular breeding tools could leverage early selections and increase genetic gains. Two main forms of marker-assisted breeding have been successfully applied in fruit crops: (i) genomic selection (GS), that relies on all markers considered simultaneously to predict the genetic merit of an individual for more complex traits; (ii) marker-assisted selection (MAS), that focuses on the use of previously identified targeted molecular markers for more oligogenic traits. Nonetheless, both methods can be integrated in the breeding pipeline, with MAS used in early-stage selection, for example, to screen a large number of seedlings for simple traits, followed by a GS approach to predict the genetic merit for complex traits for a reduced number of individuals out of the MAS selection.

Prior to the predictive and association analyses, we took advantage of this large and diverse blueberry collection, including two representative genetic pools, to investigate some population genetic parameters. First, we observed a rapid LD decay and a large frequency of pairwise markers in low LD, in both populations. This result is fairly expected given the cross-pollination nature of the species and it justifies the use of high-density marker for all the downstream genetic analyses. Also

relevant, we shed some light on sweeps across the genome, which opens an important window of opportunities to explore signature of selection. In particular, at the end of the chromosome 5, we noticed large differences in the allele frequencies across both populations (large Fst), that could be better explored in additional studies.

In terms of genetic architecture, in this study, we identified texture traits to have a complex genetic architecture, with low to moderate genomic heritability values and marker associations explaining only a small fraction of the genetic variation through GWAS analyses. Therefore, a GS breeding approach was explored for texture traits. Predictive abilities were larger when models were calibrated and tested within NHB and SHB populations separately. In contrast, and as expected, poor accuracies were obtained for inter-population predictions, hindering the use of the same predictive model across breeding programs for these two types of blueberries.

To define potential targets for MAS, on the other hand, we relied on GWAS analyses and gene mining to narrow down potential candidates related to three important traits (YM20_BuStr, F1mm and DFM). In a total of 17 texture traits measured via TA, we have focused our discussions on traits that have shown previous evidence of being associated with sensory and post-harvest characteristics [8, 11, 16]. For example, when exploring the shelf-life predictability in blueberry [16] showed that DFM is the best predictor for overall change in texture. In the same studies, the authors reported that F1mm trait showed a good ability to separate firm and soft genotypes, suggesting that F1mm has a close connection with firmness sensory perception. The use of Y20_BuSt, on the other hand, was pointed out in our PCA results as one of the most important attributes to differentiate SHB and NHB blueberries, and for this reason, was further investigated.

The genomic regions surrounding significant associations for two of the most relevant traits (YM20_BuStr and DFM) showed promising candidate genes that could be potential targets for future functional validation. At the molecular level, textural changes during ripening and postharvest of fleshy fruits have been strongly attributed to cell wall modifications [18]. For YM20_BuStr, we found candidate genes encoding enzymes that could directly modify plant cell wall and have been shown to contribute to fruit softening in several species, such as: a “xyloglucan endotransglucosylase/hydrolase (XTH)” that plays a role in disassembling xyloglucans and loosening the cell wall [19, 20]; a “pectate lyase (PEL)”, which degrades de-esterified pectin in the primary wall and middle lamella [21, 22]; and a “pectinesterase/pectinesterase inhibitor (PME)”, which acts in the modification of cell walls via demethylesterification of pectins [23].

In addition to cell wall modifying enzymes, genes encoding transcription factors that could regulate cell wall targets were also identified as candidate genes. For YM20_BuStr, we found transcription factors from different families whose members have been reported to be involved in fruit softening, such as “NAC domain-containing protein (NAC)” [24] a “squamosa promoter-binding-like protein (SPL)” [25], an “ethylene-responsive transcription factor-like protein (ERF)” [18], a “BES1 BZR1 homolog protein” [26], and a “C2H2 zinc finger protein” [27].

Genes that are part of broader regulatory and hormonal mechanisms of fruit ripening were also detected. For DFM, we found a gene encoding a “12-oxophytodiene reductase 3 (OPR3)”, which is involved in the biosynthesis of jasmonic acid and oxylipins. Jasmonate is a plant hormone involved in fruit ripening, and its external application has been shown to affect texture in different

fruit species by regulating cell wall enzyme activity [28, 29]. This gene was also found to be upregulated in a melon line with higher flesh firmness and whole fruit hardness, but lower juiciness, when compared to its parental genotype [30]. For YM20_BuStr, genes encoding a “flavin-containing monooxygenase (FMO)” were found in different peaks. A member of this family, named YUCCA, catalyzes a rate-limiting step in the tryptophan-dependent auxin biosynthesis [31]. In peaches, an insertion of a transposon-like sequence in the 5'-flanking region of the YUCCA gene was found to cause the stony hard phenotype, likely due to a reduction of the auxin levels at the late-ripening stage, resulting in low ethylene production and inhibition of fruit softening [32].

Among other categories of candidate genes highlighted, there was a “histone deacetylase 15 (HDT)” that could be recruited to epigenetically repress key ripening genes [33]. In tomato, a histone deacetylase (SlHDT3) was found to regulate fruit softening, altering the expression of ethylene biosynthesis genes and cell wall metabolism [34]. We also pinpointed genes encoding “wall-associated receptor kinase (WAK)” that participate in cell integrity and are receptors for pectin and oligogalacturonic acids binding in the cell wall. In grapes, WAKs genes (VvWARK2 and VvWARK8) were found in a stable QTL for mesocarp firmness pericarp puncture hardness [35] and were highly expressed in a firm-fleshed compared to a soft-fleshed cultivar [36]. A similar pattern of expression of WAK genes was found when comparing normal and mealy peach fruits in cold storage [37].

For future directions, we first highlight the representativeness of this research. Genetic parameters of 20 texture traits in blueberry were estimated, using two representative breeding populations of NHB and SHB. Multiple candidate genes were pinpointed, and genomic selection reveals a valid alternative to predict texture. Despite important results, a measure of caution should be taken in the interpretation of the genetic parameters. Data collection was performed in a single year and was potentially upwardly biased by Genotype-by-Environment interactions. However, an experiment of this magnitude cannot be easily repeated. Another important aspect is the real impact of the TA values on consumer preferences. Some preliminary studies in fruits have shown a correlation between some TA metrics with crunchiness and other textural aspects [5]. However, the main texture attributes measured via TA analyses associated to consumer acceptability are still elusive and should be further explored in future sensory analyses.

Altogether, our contributions in this paper are threefold in that (i) we showed that texture traits can differentiate genotypes from NHB and SHB groups; (ii) we computed genetic parameters and suggested that most of the texture attributes have a complex genetic architecture; (iii) we draw attention to the potential of genomic selection to accelerate and increase genetic gains for these traits; and (iv) we provided GWAS hits and candidate genes to be further validated.

Material and methods

Plant material

In this study, we used two breeding populations from NHB and SHB blueberry types. These breeding programs are conducted independently at USDA-ARS and Oregon State University (Corvallis, OR) and at University of Florida (Gainesville, FL), respectively. Both programs use the basic crossing design following a phenotypic recurrent selection model where parents are selected to maximize the genetic gains and minimize potential inbreeding [14].

The NHB population used in this study includes a total of 1065 genotypes. The population was further divided into genetic (GEN) and validation (VAL) study sets. The GEN study set comprises 952 individuals in the first stage of evaluation in the breeding program, derived from 85 controlled crosses made from 51 different parents. The VAL study set includes 113 individuals of advanced selections and cultivars from the NHB breeding program. Both the VAL and GEN study sets were planted in 2016 and 2017 in a high-density nursery at the Oregon State University Lewis Brown Farm in Corvallis, OR. For the NHB population, phenotypic information was collected in 2020, 2021, and 2022.

The SHB population used in this study comprises a total of 992 genotypes. Similarly to NHB, we divided the SHB population into two sets: genetic (GEN) and validation (VAL). The GEN set comprises 879 diverse individuals evaluated in the first stages of the breeding program, derived from 160 controlled crosses made from 205 selected parents in February of 2017 and 2018. The GEN population was established as a high-density nursery at the Plant Science Research and Education Unit (PSREU) in Citra, FL. The VAL set consists of 113 genotypes, including advanced selections and cultivars planted under commercial field practices at Waldo, FL. For the SHB population, phenotypic information was collected across two harvest-years (2021 and 2022).

Texture analyses

For each genotype from both NHB and SHB breeding programs, 10 berries that were fully ripe and free of visible external defects/disease were selected. Texture was evaluated using a penetration method implemented in a TA.XTPlus Texture Analyzer (Stable Micro Systems, Hamilton, MA, USA) as previously described by Giongo et al. [11]. The 10 berries were placed into a plastic condiment cup, held for 24 h at 4°C prior to measurement and warmed to room temperature prior to analysis. The penetration test used a 2 mm flat cylinder probe (TA-52) and test speed of 1 mm s⁻¹ to a final depth of 90% strain onto the equatorial axis of the fruit, perpendicular to the probe. The trigger force was set to 0.5 N with 100 data points collected per second. Using this method for each berry, 17 mechanical parameters were derived from the force deformation curve: distance to maximum force (mm, DFM), slope maximum force (N mm⁻¹, SFM), area maximum force (N mm, AFM), linear distance to maximum force (LDFM), area internal force (N mm, AIF), number of internal peaks (#, NIP), force linear distance (FLD), area force linear distance (N mm, AFLD), peak force at 1 mm (N, F1mm), maximum force (N, FM), mean internal force (N, MIF), and five Apparent Young's modulus (YM) (see Table 1 for detailed description). Data were extracted, and berries with measurements of F1mm < 0.05 (minimum force detection threshold) were removed.

We also compared results from the TA.XTPlus Texture Analyzer with traditional deflection results measured using the FruitFirm 1000 equipment (CVM Inc., Pleasanton, CA, USA). To this end, we designed a straightforward experiment where a pool of 100 random samples was measured in the same order in both machines. For the FruitFirm device, a load cell is used to determine a force-deformation value as a measure of firmness. Compression force (g/mm) from the FruitFirm 1000 was recorded and correlated with the texture attributes described in Table 1.

Genotypic data

Young leaves from each genotype were collected and submitted for DNA extraction, library preparation, and targeted genotyping by sequencing at RAPiD Genomics (LGC Group, Gainesville, FL). A probe panel comprising 22 000 loci were used for targeted

genotyping using the Flex-Seq® Ex-L platform (Flex-Seq Panel Code: FS_1903) as described by [12]. After library preparation, samples were paired-end sequenced using 150 cycles at Illumina NovaSeq platform. Raw reads were cleaned and trimmed, and the remaining reads aligned against the “W85 PO” reference genome [38] using Mosaik v.2.2.3 [39]. SNPs were called for all NHB and SHB genotypes at once with FreeBayes v.1.3.2 using the probe intervals as targets [40]. A total of 173 256 SNPs were identified. Loci were then filtered out by applying the following criteria: only biallelic locus; no monomorphic; maximum missing data of 50%; and minimum and maximum filtering depth of 15 and 150 reads, respectively. A total of 124 070 SNPs were kept after these filtering steps. Read counts were then extracted per allele per individual using vcftools v.0.1.16 [41] and subsequently used to probabilistically assign genotypic classes using the “norm model” in the updog R package v.2.1.0 [42]. From the sequence of this work, we divided our analysis in three data sets contain individuals from the SHB population, NHB population, and grouping both populations. For all the three data sets, we filtered out SNPs with minimum allele frequency less than 1%. The final number of molecular markers retained in the SHB, NHB, and both populations simultaneously were 101 386; 103 187; and 113 421 SNPs, respectively. The genotypic scores were further used in GWAS and GS analyses.

Signature of selection and population structure

Pearson correlation tests (r^2) were performed for pairwise LD estimation per chromosome, as described by de Bem Oliveira [43]. Signature of selections across the chromosomes were investigated by quantifying allelic frequency differences between SHB and NHB populations using the Wright's F_{st} statistics [44]. To assess the genetic diversity, the PCA was performed using the marker-based relationship matrix as input. Tetraploid genomic relationship matrices were computed with the AGHmatrix R-package [45]. Further, we investigated the genetic structure in each population using the probabilistic model implemented in the STRUCTURE software [46], for a range of $K = 2$ to 10. The burn-in period length was set to 20 000 and a total of 200 000 Markov Chain Monte Carlo (MCMC) chains. Because the STRUCTURE algorithm is sensible to the final number of markers included in the analyses, we considered a SNP pruning method to select equidistant markers across all the chromosomes. Thus, a total of 5166 and 4882 SNPs were selected for SHB and NHB, respectively. The best fitting K was identified with STRUCTURE HARVESTER website [47].

Phenotypic analyses and genetic analyses

Descriptive analyses were performed in both breeding populations to check general trends and phenotypic variability. Differences between trait performances of SHB and NHB populations were quantified via analysis of variance (ANOVA) using the $lm()$ function in the R software. Additionally, PCA was carried out to summarize and visualize the variability in both populations, using the FactoMineR R package [48]. To complement the multivariate analyses, we relied on an unsupervised statistical learning method to quantify the importance of texture traits on separating SHB and NHB genotypes. To this end, a Random Forest model was implemented using the default options from the randomForest R package [49] with the number of trees to grow equal to 500 and number of variables randomly sampled equal to $p/3$, where p is the number of variables.

For the phenotypic model, as multiple fruits were collected from the same individual, we estimated the so-called permanent environment effect and computed repeatability values using the

following univariate linear mixed model: $\mathbf{y} = \mathbf{X}\beta + \mathbf{Zu} + \mathbf{Wp} + \mathbf{e}$, where, \mathbf{y} is the vector of phenotypic data, β is the vector of year effect; \mathbf{u} is a vector of random additive effects following a normal distribution with zero mean and (co)variance matrix $A\sigma_a^2$, where \mathbf{A} is the additive relationship matrix and σ_a^2 is the additive genetic variance; \mathbf{p} is the vector of permanent environmental effect and non-additive genetic variance that is independently distributed with means of zero and variance σ_{pe}^2 ; and \mathbf{e} is the vector of residual effect following a normal distribution with zero mean and (co)variance matrix σ_e^2 . \mathbf{X} , \mathbf{Z} , and \mathbf{W} are incidence matrices relating the fixed and random effects to measurements in vector \mathbf{y} . Mixed models were implemented using the ASReml-R 4.0 software [50], while the \mathbf{A} matrix was computed using the AGHmatrix software [45]. Repeatability was computed as $r^2 = \frac{\sigma_a^2 + \sigma_{pe}^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2}$, following a similar strategy reported in [51]. For all the subsequent genomic analyses, we estimated the empirical BLUEs associated to the genotypic effects to be used as the observed phenotypes.

Genome-wide association analyses

The SNP-trait association analyses were based on a linear mixed model, accounting for population structure (\mathbf{Q}) and the polygenic term (\mathbf{K}) as implemented in the GWASpoly R package. The following model was fitted for each trait: $y = \mathbf{W}\alpha + \mathbf{x}\beta + \mathbf{Zu} + e$, where y is an n -vector of observed phenotypes, pre-corrected for the environment effects for n individuals, \mathbf{W} is an $n \times c$ incidence matrix of covariates (fixed effects); and α is a c -vector of the corresponding coefficients including the intercept; \mathbf{x} is an n -vector of marker genotypes, encoded in the tetraploid version after estimated the allele dosages; β is the fixed effect of the marker; \mathbf{Z} is the incidence matrix that maps the genotypes to observation; \mathbf{u} is an n -vector of the random polygenic effects, with $u \sim MVN(0, G\sigma_u^2)$ where \mathbf{G} is a known $n \times n$ genomic relationship matrix; and e is a random n -vector error term, with $e \sim MVN(0, I\sigma_e^2)$ where \mathbf{I} is an $n \times n$ identity matrix. MVN denotes the n -dimensional multivariate distribution.

To correct for population structure, we used the five first principal components as covariates in the GWAS model and used the leave-one-chromosome-out (LOCO) approach to increase the statistical power by avoiding to double fitting of the candidate marker in the model — phenomenon commonly reported in the GWAS literature as “proximal contamination” [52]. Hypothesis tests were conducted for the additive gene action, and statistical significance defined using the “M.eff” method—a Bonferroni-type correction that accounts for an effective number of markers considering LD information. The relative importance of each marker was tested using a likelihood ratio test under a backward elimination as implemented in the “fit.QTL” function.

Annotation of genetic associations

The genomic position of significant SNPs detected through GWAS analyses for the most relevant traits (YM20_BuStr, F1mm, and DFM) was used to search for candidate genes. Genes within 100 kb on each side of the SNPs were retrieved from the reference genome “W85 P0” and their functional in silico annotation performed using PANNZER2 [53] and eggNOG-mapper servers [54]. For significant SNPs located in an interval less than 1 Mb, the entire region in between was screened for candidate genes. Functional descriptions were further examined from the literature.

Genomic prediction

A genomic best linear unbiased prediction (GBLUP) model was fitted to predict the genetic merits, as follow: $\mathbf{y} = \mathbf{1}\mu + \mathbf{Zu} + \mathbf{e}$,

where \mathbf{y} is an n -vector of phenotypic observations after corrected for the year effect, μ is the population mean, and $\mathbf{1}$ is a vector of same dimension of y being all elements equal to 1, fixed effect; \mathbf{Z} is the design matrix linking the phenotypic records to genotype random effect, with $u \sim MVN(0, G\sigma_u^2)$ where \mathbf{G} is a known $n \times n$ genomic relationship matrix built using the AGHmatrix R package [45]; and $e \sim MVN(0, I\sigma_e^2)$ where \mathbf{I} is an $n \times n$ identity matrix. Variance components estimated from this model were used to compute a narrow-sense “genomic-based heritability”, as: $E[h^2] \approx \frac{\sigma_a^2(1+f)}{\sigma_a^2(1+f) + \sigma_e^2}$, where $(1+f)$ is the mean of the diagonal elements of the \mathbf{G} matrix, that is equivalent to the inbreeding coefficient (f) of the current population. Mixed models were implemented using the ASReml-R 4.0 software [50].

Model performance was evaluated using inter and intra-population prediction models. For intra-population analyses, predictive abilities were computed as the Person's correlation between predicted and observed values within each population (NHB and SHB) and using a 10-fold cross-validation scheme. For inter-prediction analyses, we tested the predictive performance across breeding populations. Namely, models were trained in the SHB population and used to predict individuals in the NHB population and vice versa.

Acknowledgements

This work was supported by SCRI-NIFA Award 2019-51181-30015, VacciniumCAP: Leveraging genetic and genomic resources to enable development of blueberry and cranberry cultivars with improved fruit quality attributes and by the University of Florida royalty fund generated by the licensing of blueberry cultivars.

Data availability

The genomic and phenotypic data are available as supplemental material.

Conflict of interest statement

The authors declare no conflict of interest.

Supplementary data

Supplementary data is available at Horticulture Research online.

References

1. Kalt W, Cassidy A, Howard LR. et al. Recent research on the health benefits of blueberries and their anthocyanins. *Adv Nutr*. 2020;**11**:224–36
2. Hancock JF, Lyrene P, Finn CE. et al. Blueberries and cranberries. In: *Temperate Fruit Crop Breeding*. Springer, 2008,115–50
3. Retamales JB, Hancock JF. Blueberries. Cabi, 2018.
4. Gallardo RK, Zhang Q, Dossett M. et al. Breeding trait priorities of the blueberry industry in the United States and Canada. *HortScience*. 2018;**53**:1021–8
5. Roth M, Muranty H, Di Guardo M. et al. Genomic prediction of fruit texture and training population optimization towards the application of genomic selection in apple. *Hortic Res*. 2020;**7**:7
6. Rivera S, Giongo L, Cappai F. et al. Blueberry firmness—a review of the textural and mechanical properties used in quality evaluations. *Postharvest Biol Technol*. 2022;**192**:112016

7. Giongo L, Poncetta P, Loretto P. et al. Texture profiling of blueberries (*Vaccinium* spp.) during fruit development, ripening and storage. *Postharvest Biol Technol.* 2013;**76**:34–9
8. Oh H, Pottorff M, Giongo L. et al. Dissecting the texture characteristics determining the phenotypic variation and postharvest change of fruit texture in blueberry. In: *Hortscience*. ALEXANDRIA, VA, 2022,S130–0
9. Ferrão LFV, Amadeu RR, Benevenuto J. et al. Genomic selection in an outcrossing autotetraploid fruit crop: lessons from blueberry breeding. *Front. Plant Sci.* 2021;**12**:1075
10. Ferrão LFV, Benevenuto J, Oliveira IB. et al. Insights into the genetic basis of blueberry fruit-related traits using diploid and polyploid models in a GWAS context. *Front Ecol Evol.* 2018;**6**:107
11. Giongo L, Ajelli M, Pottorff M. et al. Comparative multi-parameters approach to dissect texture subcomponents of high-bush blueberry cultivars at harvest and postharvest. *Postharvest Biol Technol.* 2022;**183**:111696
12. Aleza P, Boutsika A, Bassil NV. et al. Development of a targeted genotyping platform for reproducible results within tetraploid and hexaploid blueberry. *Frontiers in Horticulture* 2024;**2**:1339310
13. Mrode RA. *Linear Models for the Prediction of Animal Breeding Values*. Cabi; 2014:
14. Cellon C, Amadeu RR, Olmstead JW. et al. Estimation of genetic parameters and prediction of breeding values in an autotetraploid blueberry breeding population with extensive pedigree data. *Euphytica.* 2018;**214**:1–13
15. Luby JJ, Finn CE. Quantitative inheritance of plant growth habit in blueberry progenies. *J Am Soc Hortic Sci.* 1986;**111**:609–11
16. Oh H, Pottorff M, Giongo L. et al. Exploring shelf-life predictability of appearance traits and fruit texture in blueberry. *Postharvest Biol Technol.* 2024;**208**:112643
17. Cappai F, Benevenuto J, Ferrão L. et al. Molecular and genetic bases of fruit firmness variation in blueberry—a review. *Agronomy.* 2018;**8**:174
18. Shi Y, Li B, Su G. et al. Transcriptional regulation of fleshy fruit texture. *J Integr Plant Biol.* 2022;**64**:1649–72
19. Muñoz-Bertomeu J, Miedes E, Lorences EP. Expression of xyloglucan endotransglucosylase/hydrolase (XTH) genes and XET activity in ethylene treated apple and tomato fruits. *J Plant Physiol.* 2013;**170**:1194–201
20. Zhai Z, Feng C, Wang Y. et al. Genome-wide identification of the xyloglucan endotransglucosylase/hydrolase (XTH) and polygalacturonase (PG) genes and characterization of their role in fruit softening of sweet cherry. *Int J Mol Sci.* 2021;**22**:12331
21. Wang D, Yeats TH, Uluisik S. et al. Fruit softening: revisiting the role of pectin. *Trends Plant Sci.* 2018;**23**:302–10
22. Uluisik S, Chapman NH, Smith R. et al. Genetic improvement of tomato by targeted control of fruit softening. *Nat Biotechnol.* 2016;**34**:950–2
23. Phan TD, Bo W, West G. et al. Silencing of the major salt-dependent isoform of pectinesterase in tomato alters fruit softening. *Plant Physiol.* 2007;**144**:1960–7
24. Gao Y, Zhang Y-P, Fan Z-Q. et al. Mutagenesis of *SlNAC4* by CRISPR/Cas9 alters gene expression and softening of ripening tomato fruit. *Textes anciens.* 2021;**1**:1–12
25. Manning K, Tör M, Poole M. et al. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat Genet.* 2006;**38**:948–52
26. Li J, Guo T, Liang M. Exogenous BR delayed peach fruit softening by inhibiting pectin degradation enzyme genes. *Front Plant Sci.* 2023;**14**:1226921
27. Lai X, Zhu X, Chen H. et al. The MaC2H2-like zinc finger protein is involved in ripening and ripening disorders caused by chilling stress via the regulation of softening-related genes in 'Fenjiao' banana. *Postharvest Biol Technol.* 2022;**186**:111817
28. Saracoglu O, Ozturk B, Yildiz K. et al. Pre-harvest methyl jasmonate treatments delayed ripening and improved quality of sweet cherry fruits. *Sci Hortic.* 2017;**226**:19–23
29. Wei J, Wen X, Tang L. Effect of methyl jasmonic acid on peach fruit ripening progress. *Sci Hortic (Amsterdam).* 2017;**220**:206–13
30. Zarid M, García-Carpintero V, Esteras C. et al. Transcriptomic analysis of a near-isogenic line of melon with high fruit flesh firmness during ripening. *J Sci Food Agric.* 2021;**101**:754–77
31. Zhao Y, Christensen SK, Fankhauser C. et al. A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science.* 2001;**291**:306–9
32. Tatsuki M, Soeno K, Shimada Y. et al. Insertion of a transposon-like sequence in the 5'-flanking region of the YUCCA gene causes the stony hard phenotype. *Plant J.* 2018;**96**:815–27
33. Deng H, Chen Y, Liu Z. et al. SLERF. F12 modulates the transition to ripening in tomato fruit by recruiting the co-repressor TOPLESS and histone deacetylases to repress key ripening genes. *Plant Cell.* 2022;**34**:1250–72
34. Guo J-E, Hu Z, Li F. et al. Silencing of histone deacetylase SIHDT3 delays fruit ripening and suppresses carotenoid accumulation in tomato. *Plant Sci.* 2017;**265**:29–38
35. Lin H, Ma L, Guo Q. et al. Berry texture QTL and candidate gene analysis in grape (*Vitis vinifera* L.). *Hortic Res.* 2023;**10**:uhad226
36. Ma L, Sun L, Guo Y. et al. Transcriptome analysis of table grapes (*Vitis vinifera* L.) identified a gene network module associated with berry firmness. *PLoS One.* 2020;**15**:e0237526
37. Rothkegel K, Espinoza A, Sanhueza D. et al. Identification of DNA methylation and transcriptomic profiles associated with fruit Mealiness in *Prunus persica* (L.) Batsch. *Front Plant Sci.* 2021;**12**:684130
38. Mengist MF, Bostan H, De Paola D. et al. Autopolyploid inheritance and a heterozygous reciprocal translocation shape chromosome genetic behavior in tetraploid blueberry (*Vaccinium corymbosum*). *New Phytol.* 2023;**237**:1024–39
39. Lee W-P, Stromberg MP, Ward A. et al. MOSAIK: a hash-based algorithm for accurate next-generation sequencing short-read mapping. *PLoS One.* 2014;**9**:e90581
40. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *arXiv Prepr arXiv12073907.* 2012;
41. Danecek P, Auton A, Abecasis G. et al. The variant call format and VCFtools. *Bioinformatics.* 2011;**27**:2156–8
42. Gerard D, Ferrão LFV, Garcia AAF. et al. Genotyping polyploids from messy sequencing data. *Genetics.* 2018;**210**:789–807
43. de Bem OI, Resende MFR, Ferrão LFV. et al. Genomic prediction of autotetraploids; influence of relationship matrices, allele dosage, and continuous genotyping calls in phenotype prediction. *G3 (Bethesda).* 2019;**9**:1189–98
44. Wright S. The genetical structure of populations. *Ann Eugenics.* 1949;**15**:323–54
45. Amadeu RR, Garcia AAF, Munoz PR. et al. AGHmatrix: genetic relationship matrices in R. *Bioinformatics.* 2023;**39**:btad445
46. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;**155**:945–59
47. Earl DA, VonHoldt BM. STRUCTURE HARVESTER: a web-site and program for visualizing STRUCTURE output and

- implementing the Evanno method. *Conserv Genet Resour.* 2012;**4**: 359–61
48. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. *J Stat Softw.* 2008;**25**:1–18
 49. Breiman L. Bagging predictors. *Mach Learn.* 1996;**24**:123–40
 50. Butler DG, Cullis BR, Gilmour AR. *et al.* {ASReml}-R reference manual. 2009;
 51. Hernandez CO, Wyatt LE, Mazourek MR. Genomic prediction and selection for fruit traits in winter squash. *G three.* 2020;**10**: 3601–10
 52. Yang J, Zaitlen NA, Goddard ME. *et al.* Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet.* 2014;**46**:100–6
 53. Törönen P, Medlar A, Holm L. PANNZER2: a rapid functional annotation web server. *Nucleic Acids Res.* 2018;**46**: W84–8
 54. Cantalapiedra CP, Hernández-Plaza A, Letunic I. *et al.* eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol.* 2021;**38**:5825–9