BMC Plant Biology

RESEARCH Open Access

Unraveling the genetic architecture of blueberry fruit quality traits: major loci control organic acid content while more complex genetic mechanisms control texture and sugar content

Heeduk Oh[1,2], Molla F. Mengist[1,3], Guoying Ma[1], Lara Giongo[4], Marti Pottorff[1], Jessica A. Spencer[2], Penelope Perkins-Veazie[1,2*] and Massimo Iorizzo[1,2*]

Abstract

Background Fruit quality traits, including taste, flavor, texture, and shelf-life, have emerged as important breeding priorities in blueberry (*Vaccinium corymbosum*). Organic acids and sugars play crucial roles in the perception of blueberry taste/flavor, where low and high consumer liking are correlated with high organic acids and high sugars, respectively. Blueberry texture and appearance are also critical for shelf-life quality and consumers' willingness-to-pay. As the genetic mechanisms that determine these fruit quality traits remain largely unknown, in this study, an F_1 mapping population was used to perform quantitative trait loci (QTL) mapping for pH, titratable acidity (TA), organic acids, total soluble solids (TSS), sugars, fruit size, and texture at harvest and/or post-storage and weight loss.

Results Twenty-eight QTLs were detected for acidity-related parameters (pH, TA, and organic acid content). Six QTLs for pH, TA, and citric acid, two for quinic acid, and two for shikimic acid with major effects were consistently detected across two years on the same genomic regions on chromosomes 3, 4, and 5, respectively. Putative candidate genes for these QTLs were also identified using comparative transcriptomic analysis. No QTL was detected for malic acid content, TSS, or individual sugar content. A total of 146 QTLs with minor effects were identified for texture- and size-related parameters. With a few exceptions, these QTLs were generally inconsistent over years and post-storage, indicating a highly quantitative nature.

Conclusions Our findings enhance the understanding of the genetic basis underlying fruit quality traits in blueberry and guide future work to exploit DNA-informed selection strategies in blueberry breeding programs. The major-effect QTLs identified for acidity-related fruit characteristics could be potential targets to develop DNA markers for marker-

*Correspondence:
Penelope Perkins-Veazie
penelope_perkins@ncsu.edu
Massimo Iorizzo
miorizz@ncsu.edu

Full list of author information is available at the end of the article

assisted selection (MAS). On the other hand, genomic selection may be a more suitable approach than MAS when targeting fruit texture, sugars, or size.

Keywords Blueberry, *Vaccinium corymbosum*, Fruit quality, QTL, Candidate genes, Organic acid, Texture

Background

Over the past few decades, blueberry (*Vaccinium corymbosum*) production has expanded substantially due to successful breeding efforts on developing cultivars with low to no chilling requirements [1], leading to increased consumption. The extensive market growth has slowed as product availability has increased, with industry and consumers becoming more selective about fruit quality [2, 3]. In this new scenario, fruit quality traits, including taste, flavor, texture, and shelf-life, have become new priorities for breeding programs and the production/distribution industry [3]. The blueberry industry needs cultivars with improved and more consistent fruit quality [2, 3]. Currently, commonly grown cultivars often produce fresh fruit with inconsistent texture and sensory profiles (e.g., firmness, crispness, sweetness), leading to consumer dissatisfaction [4]. Fruit quality inconsistency is a major limitation in maintaining or expanding high-value fresh markets for blueberry [2]. Additionally, as labor costs for hand-harvested fruit account for 50–80% of the production cost [5–7], the expansion of blueberry production needs successful mechanical harvesting for the fresh market. Many of the currently grown cultivars were developed over 20 years ago and produce blueberry fruit lacking the firmness required for machine harvest and storage, limiting market opportunities [3].

Traditional blueberry breeding approaches can take up to 20 years from the original cross to cultivar release. However, rapid advances in genotyping technologies and computational tools have allowed significant acceleration in crop genetics and breeding including in blueberry [8, 9]. A number of genetic studies have been conducted in blueberry [1], and some targeted fruit quality traits, such as pH [10, 11], titratable acidity (TA) [10], total soluble solids (TSS) [10], and firmness [11–13]. The outcomes of these studies indicated that genetic factors underlie these traits and opportunities exist to establish DNA tools for marker-assisted selection (MAS) or genomic selection.

Among fruit characteristics, organic acid and sugar profiles determine fruit taste, which plays a critical role in blueberry consumer acceptance [14]. However, the genetic basis underlying these compounds is still poorly understood. QTLs for pH, TA, and TSS have been previously reported in blueberry [10, 13, 15] but these are crude parameters used as a proxy to estimate the acidity or sweetness of the fruit. To our knowledge, there are no QTL studies for organic acid or sugar contents in blueberry. Recent work by our group [16] indicated that organic acid composition varies in the blueberry germplasm. Such variation in organic acid profiles can affect measurements of pH and TA, which are parameters traditionally measured in breeding programs to select for acidity. Also, sugar content may not have a strong correlation with TSS readings in blueberry [16], possibly due to the interference of anthocyanins and phenolic compounds [17]. This highlights the importance of understanding how specific organic acids and sugars contribute to generic parameters (e.g., pH, TA, and TSS) and their genetic basis.

Blueberry fruit texture critically influences postharvest quality, consumers' willingness to pay, and machine harvestability [3, 4, 18–20]. While previous genetic studies on blueberry texture were conducted by phenotyping only 1–3 mechanical properties [11–13], recent studies have demonstrated that texture in blueberry is a multicomponent trait that requires measurement of multiple mechanical texture parameters [21–26]. Accordingly, a new study has conducted a genome-wide association study (GWAS) to identify single nucleotide polymorphism (SNP) markers associated with 17 flat probe penetration parameters [27]. Numerous small effect QTLs were found to be related to mechanical texture, suggesting a complex genetic architecture for this trait. As fruit texture changes significantly during storage [21, 25], genetic studies following storage are needed, but to the best of our knowledge, no study has been done to shed light on the storability of fruit texture.

Therefore, to complement previous work, this study aimed to perform QTL mapping for metabolites associated with acidity (organic acids) or sweetness (sugars) and texture and appearance traits at harvest and post-storage stages to gain information on the genetic mechanism and genes controlling fruit quality traits in blueberry. Texture was profiled using diverse parameters derived from multiple methods to provide a comprehensive analysis of this trait. These traits were also evaluated in a large set of commercially released cultivars to relate the variation and inheritance of the fruit quality traits assessed in the mapping population for QTL analysis. Additionally, comparative transcriptome analysis was performed to initiate efforts to unveil genes controlling organic acids. The outcomes of this study establish foundational work to design DNA-based strategies to select for fruit quality traits in blueberry.

Materials and methods

Plant materials

An F_1 mapping population including 348 genotypes derived from 'Reveille' and 'Arlen' blueberry cultivars (R×A biparental mapping population) was used in this study. This population segregates for pH, TA, and organic acids [10] and the parents 'Reveille' and 'Arlen' have different texture profiles [21]. The mapping population was grown at the North Carolina Department of Agriculture and Consumer Services (NCDA&CS) Castle Hayne Horticultural Crops Research Station (34.3649°, −77.8386°), Castle Hayne, NC, following common management practices for irrigation, pruning, fertility, and pest control. This system helped control for phenotypic variation caused by non-genetic factors across years. Fruits were harvested for two consecutive years (2021–2022) when > 50% of the berries on each bush were ripe, into plastic clamshells, placed in coolers containing refreezable ice packs, transported by car (4 h) to the Plants for Human Health Institute (Kannapolis, NC, USA). Berries were stored at − 80 °C until evaluation of chemical parameters or at 2 °C until assessment of texture and appearance parameters. In addition, a diverse set of 53 commercially available cultivars (hereafter referred to as 'diversity set') was harvested in 2021, 2022, and 2023 to assess heritability and compare phenotypic variation. Both sets were used to evaluate texture, appearance, and chemistry traits at harvest. Material from the R×A population was also evaluated for post-storage texture and appearance traits.

Evaluation of chemical parameters

pH, TA, TSS

Frozen berries were placed in a 50 mL disposable plastic tube, thawed to room temperature, and ground with a tissue homogenizer (2010 Geno/Grinder, SPEX, Metuchen, NJ, USA). Two stainless steel balls (9 mm, Grainger, Lake Forest, IL, USA) were added to each tube and a program setting of 2 min at 1,200 strokes per min, 30 s rest, and 2 min of 1,200 strokes per min was applied. pH was determined by placing an electrode (Orion 8165, Thermo Fisher Scientific, Grand Island, NY, USA) in the puree and recording the value displayed on the pH meter (Orion Star A 2111, Thermo Fisher Scientific). TA, expressed as equivalent citric acid, was determined by diluting 0.5 g puree with 24.5 mL deionized water, shaken briefly by hand, and an aliquot of the mixture was applied to a digital acid refractometer (PAL Blueberry Acidity Meter, ATAGO, Bellevue, WA, USA). A 0.5 mL aliquot of each puree was used to determine TSS using a digital refractometer (PAL-1, ATAGO).

For further analyses, the remaining puree was frozen at − 20 °C, moved to − 80 °C, and then freeze-dried (SP VirTis General Purpose Freeze Dryer, SP Scientific, Warminster, PA, USA). Freeze-dried purees were ground to a fine powder as described for purees, but with a total grinding time of 2 min.

Soluble sugars

Fructose, glucose, and sucrose were estimated using near-infrared spectroscopy (NIRS) according to the method of Perkins-Veazie et al. (2022) [28]. The NIR prediction models built by Perkins-Veazie et al. (2022) [28] were very robust with R^2 values of 82.33, 96.14, 96.73, and 96.97 for sucrose, glucose, fructose, and total sugars, respectively, and residual prediction deviation (RPD) values of 2.41, 5.11, 5.53, and 5.77 for sucrose, glucose, fructose, and total sugars, respectively. In this study, NIR spectra for the R×A samples were obtained from the freeze-dried samples using a Fourier transform NIR (FT-NIR) Spectrometer (FT-NIR Multi Purpose Analyzer (MPA), Bruker Optics, Billerica, MA, USA), and the contents of fructose, glucose, and sucrose were estimated using the prediction models.

Organic acids

Organic acids were quantified using high-performance liquid chromatography (HPLC; Hitachi LaChrom, Hitachi Ltd., San Jose, CA, USA). Extraction was done from 0.02 g of freeze-dried sample with 1.5 mL distilled deionized water, vortexed for 1 min, sonicated for 5 min at room temperature (Ultrasonic Cleaner 3510 DTH, Branson, Danbury, CT, USA), and centrifuged for 15 min at 18,292 g at 4 °C in a microcentrifuge (5417R, Eppendorf, Pittsburgh, PA, USA). After filtering the supernatant through a 0.2 µm nylon syringe filter (F2513-2, Thermo Fisher Scientific), 20 µL was injected into a Hitachi Elite LaChrom (Hitachi Ltd.) equipped with a reversed-phase C18 column (Synergi 4 µm Hydro-RP 80 A°, 4.6 × 250 mm; Phenomenex Inc., Torrance, CA, USA), ultraviolet-vis diode array detector (DAD), controlled temperature autosampler (4 °C), and column compartment (30 °C). Identification and quantification of organic acids were performed using a mobile phase of 0.0065 N sulfuric acid (H_2SO_4) with a flow rate of 1 mL min^{-1}. Data were collected and processed using D-2000 software (Hitachi Ltd.). Content of each individual organic acid was calculated from calibration curves that were developed using citric, quinic, malic, and shikimic acid standards (Sigma Aldrich, St. Louis, MO, USA).

Evaluation of texture and appearance parameters

Texture and appearance traits in the R×A material were evaluated at harvest and six weeks post-storage (hereafter indicated as T_0 and T_6, respectively) while the diversity set was evaluated only at T_0. Texture analysis in R×A was performed using three methods (flat probe penetration, needle probe penetration, and double compression)

while the diversity set was evaluated with only one method (flat probe penetration).

Ten fully ripened berries, free from any indications of external defects, decay, or wrinkling, were placed into 188 mL plastic cups (Uline, Pleasant Prairie, WI, USA), and covered with lids that had five evenly spaced holes of 3 mm diameter. The cups were placed on shallow cardboard trays, covered with large transparent zip lock bags with the zipper open, and stored at 2 °C and 80% RH for 24 h or six weeks. Samples were aliquoted in a randomized complete block design for each genotype, storage time point (e.g., T_0 and T_6), and texture method (flat probe penetration, needle probe penetration, and double compression). Fruit weight was measured using the same berries at both time points, T_0 and T_6.

Berries were transferred to room temperature (~ 20 °C) an hour before texture and appearance evaluations. A TA.XTPlus Texture Analyzer (Stable Micro Systems, Hamilton, MA, USA) and the Exponent v.6 software (Stable Micro Systems) were used for texture profiling. A high precision scale (MS1602TS/00, Mettler Toledo, Columbus, OH, USA) was used to measure berry weight and a digital caliper (Mitutoyo Compact 4-Way 500-170-30, Mitutoyo, Kawasaki, Japan) was used to measure stem scar diameter.

Texture profiling via penetration test using a 2 mm flat probe

For the penetration test, 2 mm diameter probe with a flat end was used, with a pre-test speed of 1 mm s^{-1}, auto-trigger force of 0.05 N, test speed of 2 mm s^{-1}, stopping position of 90% strain, and post-test speed of 10 mm s^{-1}, and data collection rate of 200 points per second. Each berry was penetrated on the equatorial axis and 17 parameters were derived from the force-deformation curve (Supplementary Table S1). Texture profiling using the 2 mm flat probe was performed at both time points, T_0 and T_6, in both 2021 and 2022.

Texture profiling via penetration test using a 1.4 mm needle probe

With the needle probe, a pre-test speed of 200 mm min^{-1}, auto-trigger force of 0.01 N, test speed of 300 mm min^{-1}, and post-test speed of 1,000 mm min^{-1}, and data collection rate of 500 points per second was used. The needle probe was 1.4 mm in diameter and tapered from 4 mm to a sharp tip. The equatorial axis of the fruit samples were each penetrated to a 3 mm depth. Four parameters were derived from the obtained texture profile (Supplementary Table S1). Texture analysis with the needle probe was only performed at T_0 in 2021.

Texture profiling using texture profile analysis

Texture profile analysis (TPA) or double compression test was performed with a pre-test speed of 1 mm s^{-1}, auto-trigger force of 0.05 N, test speed of 1 mm s^{-1}, target strain of 30%, and post-test speed of 5 mm s^{-1}. The waiting time between the first and second compressions was 1 s. Data was collected at a rate of 200 points per second and 17 parameters were derived from the force-time curve (Supplementary Table S1). TPA was performed at both time points, T_0 and T_6, only in 2022.

Storage index

For texture and appearance parameters that were evaluated after storage, storage index (SI) was computed according to the method of Costa et al. (2012) [29] to quantify the changes in each parameter during storage. For example, the SI of fruit weight represents weight loss over storage. SI was calculated as the base-2 logarithm of the ratio between the observed values of each parameter at T_0 and T_6, using the formula $SI = \log_2 (T_6 / T_0)$.

Heritability

Broad-sense heritability (H^2) was estimated using variance components calculated from the restricted maximum likelihood (REML) as below:

$$H^2 = \frac{\partial_g^2}{\partial_g^2 + \frac{\partial_{gy}^2}{y} + \frac{\partial_e^2}{ry}}$$

where δ_g^2, δ_{gy}^2, and δ_e^2 are variance components of genotype, genotype × environment interaction, and residual, respectively, y is the number of environments (number of years in this study; $y = 2$), and r is the number of replications ($r = 3$).

QTL mapping

The linkage map for the R×A population constructed by Mengist et al. (2021) [10] was used in this study for QTL mapping. The linkage map was developed using 80 K SNP markers, which were mined using capture-seq method, and contains SNP dosage information and the phases of the eight parental haplotypes. QTL analysis was performed using the 'polyqtlR' R package [30]. Identity-by-descent (IBD) probabilities among offspring were estimated and were used for QTL interval mapping. Significance thresholds of the LOD scores were determined through a genome-wide permutation test with 1,000 permutations ($\alpha = 0.05$). After the initial detection of QTLs, the significant QTL peaks were used as co-factors for subsequent QTL analysis to search for additional QTLs. When no further QTL was identified, the most likely QTL model was determined for each significant QTL using Bayesian Information Criterion (BIC). The phenotypic variance explained (PVE) by the QTLs and the direction of QTL effect (positive or negative) were also calculated. Confidence intervals for QTL locations

were estimated using 1-, 1.5-, and 2-LOD support intervals, and the flanking makers were recorded. Major QTLs were confirmed using the 'qtlpoly' R package [31].

Expression analysis

RNA-seq and differential expression analysis

An RNA-seq experiment was performed to identify differentially expressed genes in regions spanning the QTLs for citric, quinic, and shikimic acids. For this experiment, 17 F_1 genotypes were selected from the R×A population in 2023 based on the two-year QTL mapping and organic acid quantification results from 2021 to 2022. Samples were selected to represent the haplotypes associated with contrasting levels of citric, quinic, or shikimic acids. For citric acid, three genotypes, RA185, RA188, RA333, and three genotypes, RA012, RA209, RA337, were used to represent the haplotypes controlling high and low citric acid content, respectively (Supplementary Figure S1a). Transcriptome data from these two sets of genotypes were compared to identify candidate genes underlying the QTLs mapped on linkage group (LG) 3, controlling pH, TA, and citric acid content. For quinic acid, three genotypes, RA062, RA081, RA333, and three genotypes, RA003, RA176, RA181, were used to represent the haplotypes controlling high and low quinic acid contents, respectively (Supplementary Figure S1b). Transcriptome data from these two sets of genotypes were compared to identify candidate genes underlying the QTLs mapped on LG4, controlling quinic acid content. For shikimic acid, three genotypes, RA047, RA097, RA304, and three genotypes, RA166, RA282, RA361, were used to represent the haplotypes controlling high and low shikimic acid contents, respectively (Supplementary Figure S1c). Transcriptome data from these two sets of genotypes were compared to identify candidate genes underlying the QTLs mapped on LG5, controlling shikimic acid content.

Fully ripened berries with no signs of external defects, decay, or wrinkling were harvested from 17 genotypes. Samples were flash-frozen in liquid nitrogen and stored at − 80 °C until RNA extraction. Total RNA was extracted from the fruit using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich, MO, USA). Library preparation and mRNA sequencing were performed by Novogene (Novogene Corporation Inc., CA, USA), using the 150 bp paired-end Illumina NovaSeq 6000 Sequencing System. The reads were trimmed with fastp v.0.23.2 [32] and were aligned independently to the W85-20_v2_p0 [33] and the Draper_v1 [34] reference genomes using STAR v.2.7.10a [35] and expression levels were quantified using Salmon v.1.9.0 [36]. The genes in the Draper_v1 genome represent all four haplotypes, while those in the W85-20_v2_p0 genome represent only the primary haplotype (namely, p0). DESeq2 v.1.38.3 was used for differential expression analysis [37]. Functional annotation of the genes was performed using eggNOG-mapper v.2.1.11 [38].

Validation of RNA-seq by quantitative real-time PCR

To validate the RNA-seq results, two putative candidate genes related to citric acid content were selected to conduct quantitative real-time PCR (qRT-PCR). First-strand cDNA was synthesized with 1 µg of total RNA using the Verso cDNA Synthesis Kit (Thermo Fisher Scientific). Primers for qRT-PCR were designed to span two exons to ensure no genomic DNA contamination (Supplementary Table S2). The absence of genomic DNA contamination in the cDNA samples was verified with PCR followed by gel electrophoresis analysis (data not shown). qRT-PCR was performed on a LightCycler 480 II (Roche Diagnostics, Indianapolis, IN, USA) using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems, Foster City, CA, USA). Conditions for the reactions were: 95 °C for 2 min, followed by 45 cycles of 95 °C for 15 s, 60 °C for 15 s, and 72 °C for 1 min, followed by a melting curve program from 60 °C to 95 °C with a heating rate of 0.15 °C s^{-1}. The *UBIQUITIN-CONJUGATING ENZYME* (*UBC28*; *Vcev1_p0.Chr1.02793*) was used as the reference gene [39] to calculate the relative expression of the candidate genes using the Pfaffl method [40]. Statistical differences were determined using SAS 9.4 (SAS Institute, Cary, NC, USA).

Results

Phenotypic variability and heritability

Extensive phenotypic variation among R×A genotypes was observed for all the chemical, textural, and appearance parameters evaluated in this study. Variations in chemistry parameters were similar to those characterized in the diversity set (for pH, TA, and TSS, see Supplementary Figure S2; for other parameters, data not shown). Most parameters in both sets followed near-normal distributions, pointing to a quantitative nature (Fig. 1, Supplementary Figure S3). Quinic and shikimic acid contents exhibited skewed distributions, meaning that these traits may be under oligogenic inheritance. The predominant organic acid in this population was citric acid (average 81.7%), followed by quinic (12.9%), malic (5.1%), and shikimic (0.3%) acids. The major sugars were fructose (average 50.1%) and glucose (47.6%), followed by sucrose (2.3%). Most texture parameters spanned the same variation as the diversity set, with near-normal distributions (data not shown). A few, such as the Young's Modulus (YM) parameters, had narrower variations in the R×A population compared to the diversity set.

Broad-sense heritability (H^2) estimation revealed that pH, TA, and organic acid content are high to moderately heritable traits, where H^2 ranged from 44% for malic acid content to 91% for percent quinic acid concentration

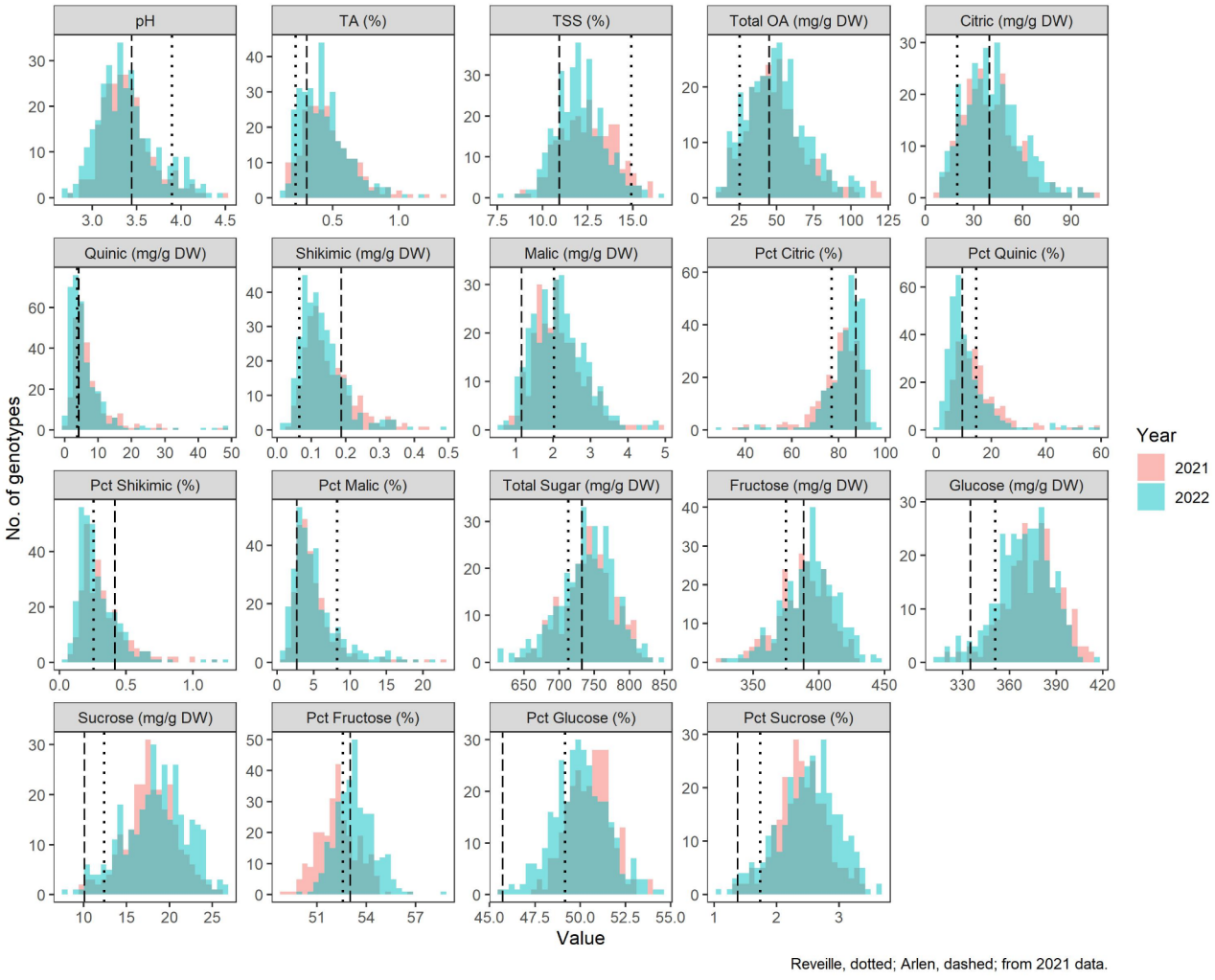

Fig. 1 Distribution of pH, titratable acidity (TA), total soluble solids (TSS), contents of organic acids and sugars, and percent concentration (Pct) of organic acids and sugars observed in a 'Reveille' × 'Arlen' F_1 population across two years (2021–2022). Dotted and dashed lines indicate the phenotype of the parents 'Reveille' and 'Arlen', respectively

(Fig. 2). Except for malic acid, H^2 was higher than 70% for pH, TA, and the other organic acids. TSS and sugar contents had moderate to low heritability, ranging from 31 to 51%. All the flat probe texture parameters had relatively high heritability (≥ 60%), demonstrating that texture is a highly heritable trait (Supplementary Figure S4). High heritability of appearance traits was found for Ssd, Wg, and Dia, with H^2 of 73, 68, and 70%, respectively. Similar levels of H^2 were observed in the diversity set except for individual sugars, which had slightly higher H^2 values (data not shown). The heritability of the needle probe and TPA texture parameters were not estimated since data was collected for only one year.

Pearson correlation analysis was performed to explore the relationship between parameters (Fig. 3, Supplementary Figure S5). In R×A, strong positive correlations were found between TA, total organic acid content, and citric acid content, and these were negatively correlated with pH. Thus, citric acid explained most of the phenotypic variation of pH and TA in this population. Total organic acid content, citric acid content, and TA were negatively correlated with TSS and sugar contents. Fructose, glucose, total sugar content, and TSS were significantly ($p < 0.05$) and positively correlated. For texture, separate clusters of positively correlated flat probe parameters were identified (Supplementary Figure S5). One cluster included BrSt, DFM, and LDFM (see Supplementary Table S1 for abbreviations), which had negative correlations with fruit size. On the other hand, the other texture parameters were grouped in another cluster, where most of them were positively correlated with size. Interestingly, FLD had an especially high correlation with size, meaning that it is a texture parameter that may be highly affected by the berry size. The TPA texture parameters were evaluated in 2022 only. Most TPA parameters were significantly ($p < 0.05$) and positively correlated with each

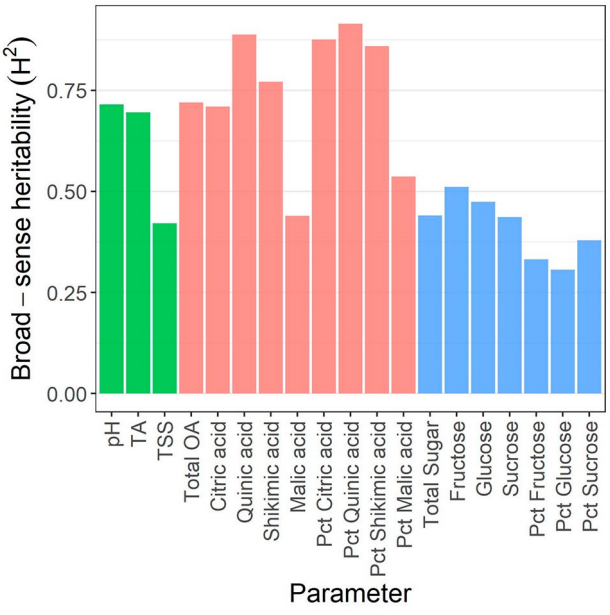

Fig. 2 Broad-sense heritability (H^2) of pH, titratable acidity (TA), total soluble solids (TSS), contents of organic acids and sugars, and percent concentration (Pct) of organic acids and sugars observed in a 'Reveille' × 'Arlen' F_1 population across two years (2021–2022)

other and with the flat probe parameters. No strong correlations were found between fruit chemistry and texture parameters. The correlation patterns among fruit characteristics observed in R×A generally coincided with those in the diversity set (data not shown).

In summary, similar phenotypic variability, H^2, and correlation patterns were observed between the R×A population used in this study for QTL mapping and the diversity set, representing a wide range of phenotypes. This demonstrated sufficient segregation for the fruit quality traits in the R×A population, warranting further genetic investigation of the traits.

QTLs for pH, TA, and organic acid content

A total of 30 QTLs were detected for chemistry characteristics in this study (Supplementary Table S3, Supplementary Figures S6-10). Details of the detected QTLs are outlined in Supplementary Table S3, including peak marker, physical location, LOD score, PVE, co-factors, and intervals. Out of these QTLs, 28 were associated with pH, TA, and organic acid content. QTLs on LGs 3, 4, and 5 were consistently detected with the 2021 and 2022 data with high LOD scores and PVE values (Fig. 4, Supplementary Figure S6). A QTL controlling pH, TA, total organic acid content, and citric acid content was identified on LG3, which was consistent across the two years and explained 15.6–20.1% of the phenotypic variance of each trait. The LG3 QTL was mainly linked to homologs H7 and H8 from 'Reveille' (Fig. 4a). Alleles on H7 and H8 had negative and positive additive effects, respectively, on TA, total organic acid content, and citric acid content. The effects were opposite for pH (Supplementary Figure S11a-d), which coincided with the negative correlations between pH and the other traits (Fig. 3).

The QTL mapped on LG4 controlled quinic acid content. The PVE was estimated between 27.6 and 32.9% for 2021 and 2022, respectively. H1 from 'Arlen' and H5 and H6 from 'Reveille' had negative dominant effects on quinic acid content (Fig. 4b, Supplementary Figure S11e). The QTL on LG5 controlled shikimic acid content and explained 17.2 and 18.8% of the phenotypic variances for 2021 and 2022, respectively. H1 from 'Arlen' had a positive additive effect on shikimic acid content (Fig. 4c, Supplementary Figure S11f). These major effect QTLs on LGs 3, 4, and 5 were also consistently detected when QTL mapping was performed using the 'qtlpoly' R package, confirming the stability of these QTLs (data not shown).

QTLs with lower LOD scores and PVE values were detected for these chemistry parameters on other LGs, which were not consistent over years (Supplementary Table S3, Supplementary Figure S6). A QTL for TA was detected on LG8 in 2021, explaining 13.4% of the phenotypic variance. A QTL on LG9 was identified for TA, pH, total organic acid content, and citric acid content in 2021, where the PVE was estimated between 10.9 and 14.0%. A QTL for shikimic acid content was detected on LG4 in 2022, which had PVE of 16.6% and was located in the same region as the LG4 QTL for quinic acid content. No QTL was detected for malic acid content.

The percent concentration of each organic acid, relative to the total organic acid content, was also determined to identify QTLs for the contribution of specific organic acids to the overall organic acid profile. A QTL on LG4 was associated with the percent concentrations of citric acid and quinic acid in both years. It was detected in the same region as the LG4 QTLs for quinic acid and shikimic acid contents. Additionally, QTLs for percent concentrations of shikimic acid and malic acid were detected on LG3, which was in the same region as the LG3 QTLs for TA, pH, total organic acid content, and citric acid content.

QTLs for sugar, texture, and size

For sugars, only two minor QTLs were detected for percent concentration: a QTL for percent fructose on LG10 in 2021 and a QTL for percent sucrose on LG2 in 2022. The QTLs had LOD scores of 6.4 and 6.8 and PVE of 11.2 and 9.8%, respectively. No QTLs that were consistent across years were detected. Also, no QTLs were detected in either year for TSS, total sugar content, or fructose, glucose, sucrose contents.

For texture and size parameters, 146 QTLs were identified in total on LGs 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11, which explained 5.9–14.1% of the phenotypic variances

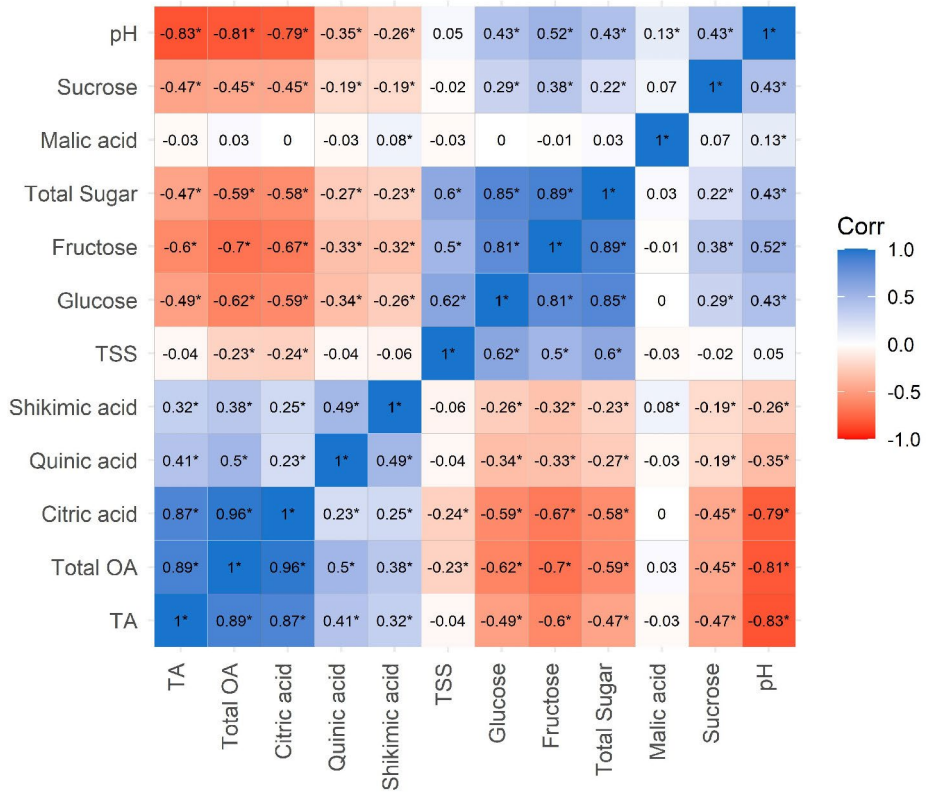

Fig. 3 Correlations among pH, titratable acidity (TA), total soluble solids (TSS), contents of organic acids and sugars, and percent concentration (Pct) of organic acids and sugars observed in a 'Reveille' × 'Arlen' F_1 population across two years (2021–2022)

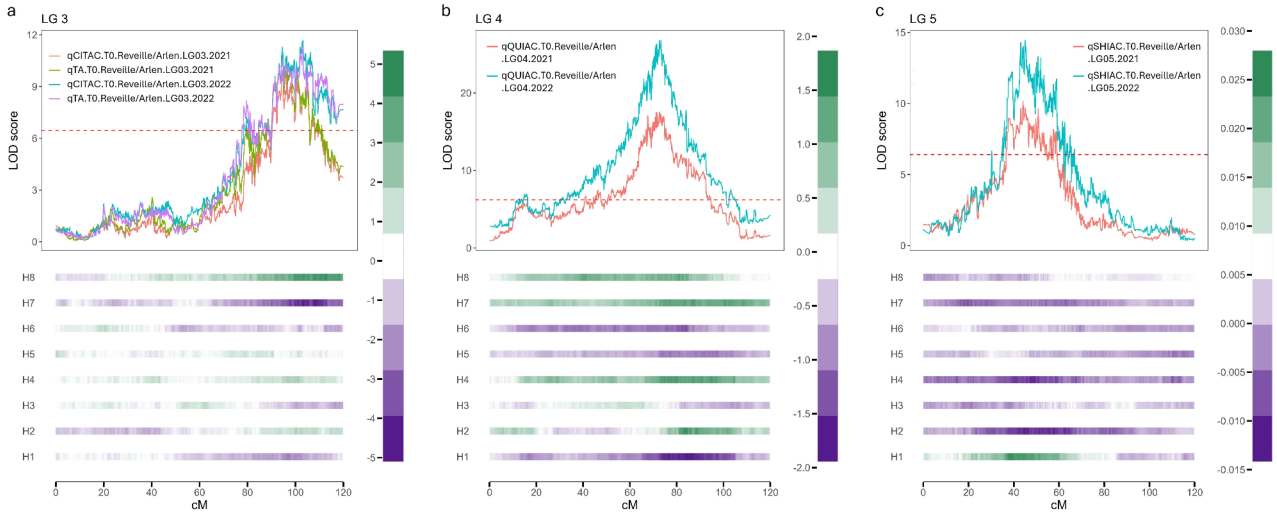

Fig. 4 Major-effect quantitative trait loci (QTLs) identified for total organic acid (OA) and citric acid contents on linkage group (LG) 3 (**a**), quinic acid content on LG4 (**b**), and shikimic acid on LG5 (**c**) in a 'Reveille' × 'Arlen' F_1 population. The heatmap illustrates the effect of each homolog relative to the overall phenotypic mean performance. H1-H8 represents the eight homologs with H1-H4 inherited from the parent 'Arlen' and H5-H8 inherited from the parent 'Reveille'. Dotted line indicates LOD threshold (1,000 permutations, $\alpha = 0.05$) for significant QTLs

(Supplementary Table S3, Supplementary Figures S7-10). A few QTLs were consistent over time points (T_0 and T_6) or across years (2021 and 2022). This included a QTL on LG10 for FLD and QTLs for YM20_BrSt and YM1to2 (Supplementary Figure S8). However, no QTLs related to the SI values (representing change during storage) were consistent over years. Also, no QTL was consistent across years for the size parameters of fruit weight and diameter.

RNA-seq to identify candidate genes for organic acids

Potential candidate genes for pH, TA, and citric, quinic, shikimic acids were explored. RNA-seq experiments were conducted using genotypes with contrasting haplotypes across the regions spanning the conserved physical locations of the major-effect QTLs on LGs 3 (between 44,412,643 and 50,885,848 bp), 4 (between 17,303,781 and 48,065,094 bp), and 5 (between 12,330,354 and 24,329,493 bp) (see Materials and Methods). An average of 48 million reads (paired-end 150 bp) were generated via the Illumina NovaSeq 6000 Sequencing System. The mapping rate of the reads for each genotype and reference genome are listed in Supplementary Tables S4-5.

Using comparative transcriptome analysis, differentially expressed genes (DEGs) were identified in high vs. low citric acid content, high vs. low quinic acid content, and high vs. low shikimic acid content genotypes (Supplementary Figure S1). Using the reads mapped onto the W85-20_v2_p0 reference genome, 770, 200, and 187 DEGs were identified in the high vs. low citric acid content, high vs. low quinic acid content, and high vs. low shikimic acid content comparisons, respectively. Out of the DEGs, 414 and 356 genes, 94 and 106 genes, and 106 and 81 genes were up- and down-regulated in the high vs. low citric acid content, high vs. low quinic acid content, and high vs. low shikimic acid content comparisons, respectively.

Although the W85-20_v2 genome is the newest high-quality, phased, chromosome-scale genome available for blueberry, this genome represents a wild diploid species (*Vaccinium caesariense*) also known as diploid blueberry. As there is evolutionary divergence between diploid and tetraploid cultivated blueberry, it is possible that genes not present in the W85-20 genotype could be missed. Therefore, the Draper_v1 genome, representing a cultivated tetraploid blueberry cultivar, was used as an additional reference genome. Using the reads mapped onto the Draper_v1 reference genome across the four haplotypes, 2,939, 642, and 674 DEGs were identified in the high vs. low citric acid content, high vs. low quinic acid content, and high vs. low shikimic acid content comparisons, respectively. Out of the DEGs, 1,618 and 1,321 genes, 325 and 317 genes, and 384 and 290 genes were up- and down-regulated in the high vs. low citric acid content, high vs. low quinic acid content, and high vs. low shikimic acid content comparisons, respectively. The larger number of DEGs identified with the Draper genome is largely because four haplotypes were used for the analysis while only one haplotype was used with the W85-20 genome. Indeed, considering the number of DEGs per haplotype reported within the QTL regions (Supplementary Tables S6-S7), the number of DEGs identified using the W85-20_v2_p0 (59 DEGs) and Draper_v1 (ranging between 49 and 59 DEGs) reference genomes were relatively similar.

To further pinpoint candidate genes putatively associated with citric, quinic, and shikimic acids, functional annotation was conducted for genes spanning the QTL regions associated with these traits. The regions that were significantly (1,000 permutations, $\alpha = 0.05$) associated with these traits across the two years were considered for this analysis. For the DEGs identified using the W85-20_v2_p0 reference genome, 22, 32, and 5 DEGs for citric, quinic, and shikimic acids were detected, respectively (Fig. 5, Supplementary Table S6). For the DEGs identified using the Draper_v1 reference genome, 110, 77, and 24 DEGs were within each significant QTL interval for citric, quinic, and shikimic acids, respectively (Supplementary Table S7).

In the LG3 QTL interval associated with pH, TA, and citric acid, the DEGs *Vcev1_p0.Chr3.08885* and *Vcev1_p0.Chr3.08969* were annotated as class-III pyridoxal-phosphate-dependent aminotransferase and Cys/Met metabolism PLP-dependent enzyme, respectively. These enzymes are known to be involved in amino acid metabolism utilizing 2-oxoglutarate, which is a primary metabolite in the citric acid cycle (also known as the tricarboxylic acid (TCA) cycle or the Krebs cycle). *Vcev1_p0.Chr3.08885* was down-regulated in high citric acid genotypes, while *Vcev1_p0.Chr3.08969* was up-regulated. Within the QTL intervals associated with quinic (LG4) and shikimic acids (LG5), no DEGs are known to be associated with organic acid biosynthesis and metabolism.

To validate the RNA-seq results, two DEGs in the LG3 QTL region were selected for qRT-PCR: *Vcev1_p0.Chr03.08687* and *Vcev1_p0.Chr03.08715*. These genes were differentially expressed between genotypes with contrasting citric acid levels. qRT-PCR results confirmed significant differences in gene expression between high vs. low citric acid genotypes (Supplementary Figure S12). The expression patterns were similar to the RNA-seq results, in which both genes were up-regulated in the high citric acid genotypes.

Discussion

Three major QTLs control acidity and organic acid content in blueberry

Fruit acidity is a crucial component of the organoleptic quality of blueberries [41]. Three major QTLs associated

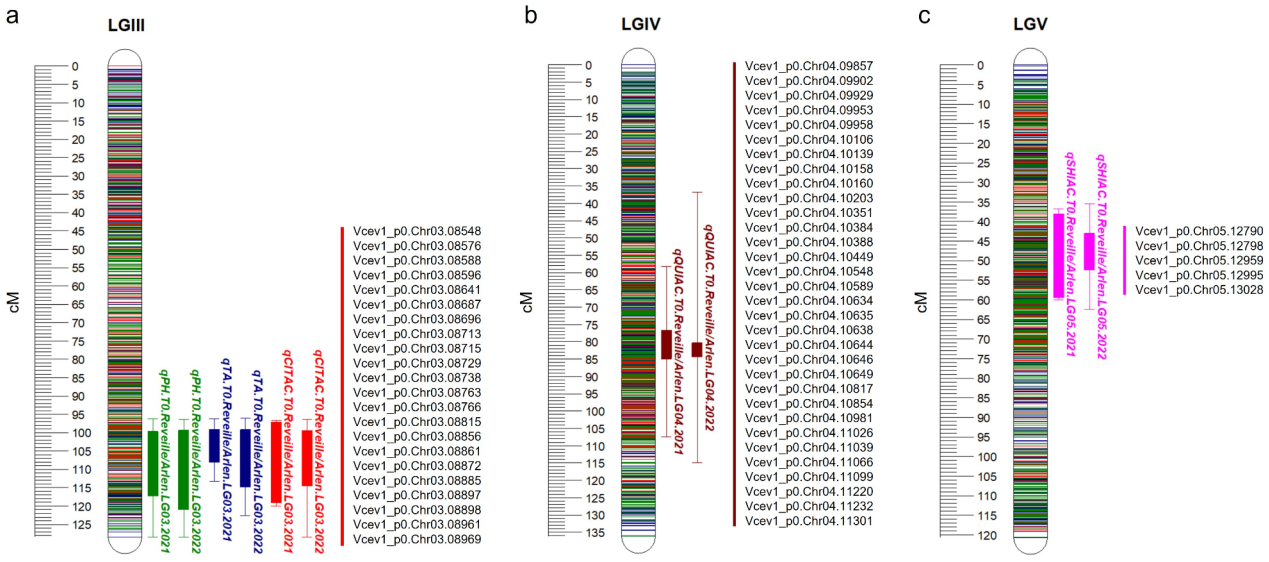

Fig. 5 Quantitative trait loci (QTL) detected for chemical parameters, including pH, titratable acidity (TA), and contents of citric, quinic, and shikimic acids, and differentially expressed genes (DEG) in each QTL region identified from RNA-seq analysis on linkage groups (LG) 3 (**a**), 4 (**b**), and 5 (**c**). The boxplots represent the 95% permutation support interval of the QTLs (interval where LOD score exceeded the threshold (1,000 permutations, $\alpha = 0.05$)). The solid box within each boxplot represents the two-LOD support interval. DEGs identified from RNA-seq analysis using genotypes with contrasting organic acid contents (Supplementary Figure S1) that were located within the 95% permutation support interval of the respective QTL are listed. Description of the DEGs (e.g., chromosomal location, function annotation, and differential expression analysis results, etc.) is listed in Supplementary Tables S6-7

with acidity and organic acid levels were detected on LGs 3, 4, and 5 (Fig. 4, Supplementary Figure S6). The LG3 QTLs identified for pH, TA, total organic acid content, and citric acid content were co-located with the QTLs found for pH or TA in previous reports [10, 13, 15]. The LG3 QTLs were all estimated to have additive effects on these quality traits, aligning with previous findings [10]. SNP markers with the highest LOD scores for each trait – based on single marker analysis results – confirmed the additive nature of the QTLs and provided information on the allele dosage effect (Supplementary Figure S11a-d). The SNP marker NCSU_Chr03_47259524 had the highest LOD score for pH, TA, and total organic acid content, for which the average allele dosage effect was +0.20, −0.10% points, and −11.27 mg g^{-1} DW per allele, respectively. The direction of the effects of the QTLs (positive or negative) for pH and TA coincided with previous findings [15]. For citric acid content, the SNP marker NCSU_Chr03_49554919 scored the highest LOD and the average allele dosage effect was +11.41 mg g^{-1} DW. The average allele dosage effects of these SNP markers accounted for roughly 10% of the total phenotypic variation in the respective fruit characteristics. Relatively high PVE values of the LG3 QTLs (15.6–20.1%), similar to previous reports [10, 15], and high broad-sense heritability of these fruit characteristics (>70%) indicate that response to selection can be achieved via DNA-informed breeding strategies.

The LG4 QTL was associated with quinic acid content in both years, explaining 27.6–32.9% of the phenotypic variance, and was estimated to have a dominant action. Further investigation on this peak via single marker analysis revealed that the SNP marker NCSU_Chr04_36016196, which had the highest LOD score, had a recessive effect on quinic acid (Supplementary Figure S11e). The homozygote genotypes with 0 allele dose had an average quinic acid content of 18.3 mg g^{-1} DW, followed by 6.0, 4.0, and 3.1 mg g^{-1} DW for genotypes with 1, 2, and 3 dosages, respectively.

While citric acid is highly correlated with fruit acidity, the role of quinic acid in sensory perception is still obscure in blueberry. In fresh blueberry juice, quinic acid had low correlations with sourness ($r = 0.49$) and bitter taste ($r = 0.53$) [42]. However, quinic acid was reported to be negatively correlated with sweet taste ($r = -0.76$) of the juice [42] and with the juice 'blueberry' flavor ($r = -0.75$), composed of aromatic volatiles associated with fresh blueberries [43]. The role of quinic acid in the taste or flavor of fresh blueberry fruit will need to be established in future work.

A QTL, co-located on LG4 with the QTLs for quinic acid, was identified for shikimic acid content in 2022 and accounted for 16.6% of the phenotypic variance (Supplementary Table S3, Supplementary Figure S6). The co-location of these QTLs for quinic and shikimic acids may be due to the shared metabolic pathways between the two organic acids. Quinic acid has been suggested to act as a reserve compound for phenolic biosynthesis in fruit [44] and stored quinic acid can re-enter the shikimate pathway through the action of quinate dehydrogenase

or quinic hydrolase, which may lead to increased biosynthesis of shikimic acid [45]. Alternatively, quinic acid can be directly involved in synthesizing compounds such as chlorogenic acids or acyl-quinic acids, which are conjugates of quinic and cinnamic acids [46–48]. Quinic acid can also serve as a precursor for synthesizing caffeoylquinic acids (CQAs), which are specialized bioactive metabolites derived from the phenylpropanoid biosynthesis pathway [49]. As this QTL was not detected in 2021, a strong environmental effect on the intertwined metabolisms between quinic and shikimic acids may be present.

The LG5 QTL had a major effect on shikimic acid content across both years, explaining 17.2–18.8% of the phenotypic variances (Supplementary Table S3, Supplementary Figure S6). This QTL was estimated to have an additive action and single marker analysis indicated that the SNP marker with the highest LOD score, NCSU_Chr05_21287627, had an average dosage effect of + 0.05 mg g^{-1} DW (Supplementary Figure S11f). Shikimic acid is a minor constituent of the organic acid profile in blueberry, likely limiting its role in altering fruit acidity level. However, chorismate, the terminal metabolite of the shikimic acid pathway, serves as an important intermediate branch point metabolite for the biosynthesis of several aromatic amino acids [50]. Therefore, alteration of shikimic acid may have a crucial influence on the aromatic perception and flavor of blueberries.

The acidity-related fruit characteristics that are controlled by major-effect QTLs and have high heritability (e.g., pH, TA, total organic acid, and citric, quinic, shikimic acid contents) can be potential targets for marker-assisted selection (MAS). To our knowledge, this is the first study to perform QTL mapping for organic acids, which enabled us to relate these QTLs to those detected for pH and TA. Previous studies in blueberry using different genetic backgrounds identified a QTL for pH and TA in chromosome 3. Ferrao et al. (2018) [13] identified QTLs for pH spanning the QTL region identified in this paper. Mengist et al. (2021, 2022) [10, 15] identified a QTL on the distal part of the long arm of chromosome 3 (unknown physical position) that might overlap with the QTLs detected in this study. These observations validate the significance of this QTL across the blueberry germplasms and make it an ideal target region to design DNA markers for MAS. In contrast, since no other genetic studies assessed the genetic mechanism controlling quinic and shikimic acids, future work is needed to validate those QTLs in the blueberry germplasm and assess the importance of these regions to design DNA markers for MAS. Implementing MAS can significantly increase the efficiency and accuracy of selection in breeding programs aiming for improved fruit taste/flavor. Genotyping assays that allow detection of the dosage of SNP markers, such as the Kompetitive Allele Specific PCR (KASP) assay [51], could be developed to facilitate MAS for these traits. As fruit taste and flavor play crucial roles in consumer-liking [52] and willingness-to-pay [14] in blueberries, the QTLs identified in this work should be considered for the application of DNA-informed selection.

Although malic acid is known to be an important component of acidity along with citric acid in many types of fleshy fruit [41], malic acid accounts for only a small proportion in highbush blueberries [53]. Indeed, malic acid was a minor constituent of the overall organic acid profile in this population as well, composing less than 10% of the total organic acids in most genotypes (Fig. 1). Moreover, the broad-sense heritability estimate of malic acid content was relatively low compared to those of other organic acids (Fig. 2). Consequently, we were not able to detect any QTLs for malic acid content despite the large phenotypic variation (6.4-fold and 7.9-fold in 2021 and 2022, respectively). These results indicate that malic acid may be controlled by a large number of genes and/or is highly affected by environmental factors. It is also possible that the level of variation captured in this study may not be sufficient to detect QTLs for malic acid. Future work should explore other populations segregating for malic acid or germplasms with higher malic acid levels. Notably, two QTLs were detected for the percent concentration of malic acid in 2022 on LGs 3 and 5, explaining 14.2 and 9.3% of the phenotypic variances, respectively (Supplementary Table S3, Supplementary Figure S6). The LG3 QTL for percent concentration of malic acid was co-located with the major-effect QTL controlling citric acid content. This indicates that the proportion of the total organic acids that malic acid represents is likely controlled more by the citric acid content than the malic acid content.

Candidate genes controlling organic acid contents were identified

Within the genomic regions spanning the major QTLs on LGs 3, 4, and 5, we identified putative candidate genes involved in the organic acid metabolism or transport (Fig. 5, Supplementary Table S6). RNA-seq analysis of high vs. low organic acid content genotypes revealed that only two DEGs within the citric acid QTL (LG3) interval, *Vcev1_p0.Chr3.08885* and *Vcev1_p0.Chr3.08969*, are involved in citric acid biosynthesis and metabolism. Since post-translational modification or any other mutations that are not associated with the gene expression levels can control these QTLs, the possible role of not differentially expressed genes (non-DEGs) should be considered.

Within the LG3 QTL interval, several genes that were non-DEGs were associated with the citric acid accumulation. *Vcev1_p0.Chr3.08774* and *Vcev1_p0.Chr3.09032* were predicted to be involved in the citric acid cycle.

Vcev1_p0.Chr3.08774 was annotated as pyruvate dehydrogenase, which catalyzes the overall conversion of pyruvate to acetyl-CoA and CO_2. *Vcev1_p0.Chr3.09032* was annotated as a malic enzyme that is known to convert malate to pyruvate or oxaloacetate. Additionally, *Vcev1_p0.Chr3.08722*, also in the LG3 QTL interval, was annotated as a malate synthase, known to catalyze the condensation of acetyl-CoA with glyoxylate to form (S)-malate in the glyoxylate cycle. Lastly, *Vcev1_p0.Chr3.08881* was annotated as a plasma membrane H^+ ATPase, which is known to pump protons across the cellular membrane, regulating the organic acid accumulation in apple [54] and citrus [55, 56].

Citric acid was the predominant organic acid in the R×A population, making up more than 80% of the total organic acid content on average (Fig. 1). Additionally, it was highly correlated to pH, TA, and total organic acid content (Fig. 3). These findings indicate that citric acid explains most of the genetic and phenotypic variation of pH, TA, and total organic acid content in this population. Moreover, citric acid has been identified as the predominant organic acid in highbush blueberries [53, 57, 58]. Given these results, the putative candidate genes that were identified within the LG3 QTL interval emerge as particularly intriguing.

There were no DEGs in the LG4 and LG5 QTL intervals that were noteworthy based on the annotations. However, several non-DEGs that could be putative candidate genes were identified. In the LG4 QTL interval, *Vcev1_p0.Chr4.11308* was annotated as shikimate kinase, known to be involved in the shikimate pathway. In the LG5 QTL interval, *Vcev1_p0.Chr5.12667* was annotated as phospho-2-dehydro-3-deoxyheptonate aldolase and *Vcev1_p0.Chr5.12990* was annotated as dehydratase shikimate dehydrogenase, which are involved in the shikimate pathway as well.

The non-DEG putative candidate genes mentioned above could be differentially expressed at different fruit developmental stages. Expression levels of the genes related to organic acid metabolism have been reported to be differentially regulated with fruit maturation stages in blueberry [59] and other fruits [60–64]. Future metabolomic or transcriptomic analyses that include different fruit ripening stages could provide a better understanding of the mode of action of these genes, and thus, further work is needed to fully validate them as candidate genes. Assessing the expression levels of DEGs identified for citric acid in other genetic backgrounds could help validate these candidate genes and design experiments for functional characterization such as gene transformation using silencing (e.g., virus-induced gene silencing (VIGS)). Gene regulation could also be affected at the post-transcriptional level and alternative approaches to identify the best candidate genes controlling citric, quinic, and shikimic acids could involve proteomic analyses in samples harboring the dominant and recessive alleles.

No major-effect QTL was detected for sugars, texture, or size

Sugars play a considerable role in consumer liking for blueberries. Higher sugar content or TSS generally leads to increased sweetness resulting in better consumer liking [52]. Despite the importance, no QTL was detected in this study for TSS, total sugar content, or fructose, glucose, sucrose contents. The heritability for these traits was relatively low ($H^2 \leq 51\%$), indicating that in blueberry, sugar accumulation is a complex trait and is influenced by environmental factors. Several reports have suggested that TSS or sugar content in fruit may be controlled by complex genetic mechanisms [65–67]. Also, the phenotypic variability of these characteristics in both the R×A population and the diversity set was narrow, having less than two-fold variation in TSS and major sugars (Fig. 1). This observation confirms that, in blueberry, sugar accumulation may not be a qualitative trait. Major QTLs associated with TSS or sugars have been reported in other fruit where the phenotypic variation was wider (up to 7-fold variation), such as apple [68], citrus [69], grape [70–72], melon [67, 73, 74], peach [75], and watermelon [76]. The absence of QTLs for sugar content in blueberry fruit in this study suggests that identifying candidate genes or developing DNA markers for MAS strategies might be challenging and genomic selection may be a more suitable approach. MAS targeting acidity parameters or organic acid content could be a more suitable strategy for breeding programs that aim to improve taste or flavor. Nevertheless, the possibility of detecting QTLs for sugars in populations with wider phenotypic variation should not be excluded.

Texture is an important fruit quality trait that influences machine harvestability, shelf-life quality, and consumers' acceptance in blueberry [3, 4, 18–20]. A total of 130 QTLs were detected for texture parameters (Supplementary Table S3, Supplementary Figures S8-10). Some texture QTLs were consistent over years (2021–2022) and/or time points (T_0 and T_6), indicating that these loci could be potential targets for marker development. For example, a QTL for 'force linear distance' (FLD), a flat probe parameter, was detected on LG10 both at T_0 and T_6 in both 2021 and 2022. FLD is a parameter that has been reported to be useful for predicting wrinkling occurrence during storage [21] and is highly correlated with sensory attributes such as springiness and hardness [26]. Several other flat probe parameters, such as 'area force linear distance' (AFLD), 'force at 1 mm' (F1mm), and 'Young's Modulus' (YM) parameters (e.g., YM10, YM20_BrSt, YM80_BrSt, YM100_BrSt, YM1to2), were consistently detected over years on LG10 for measurements at T_0.

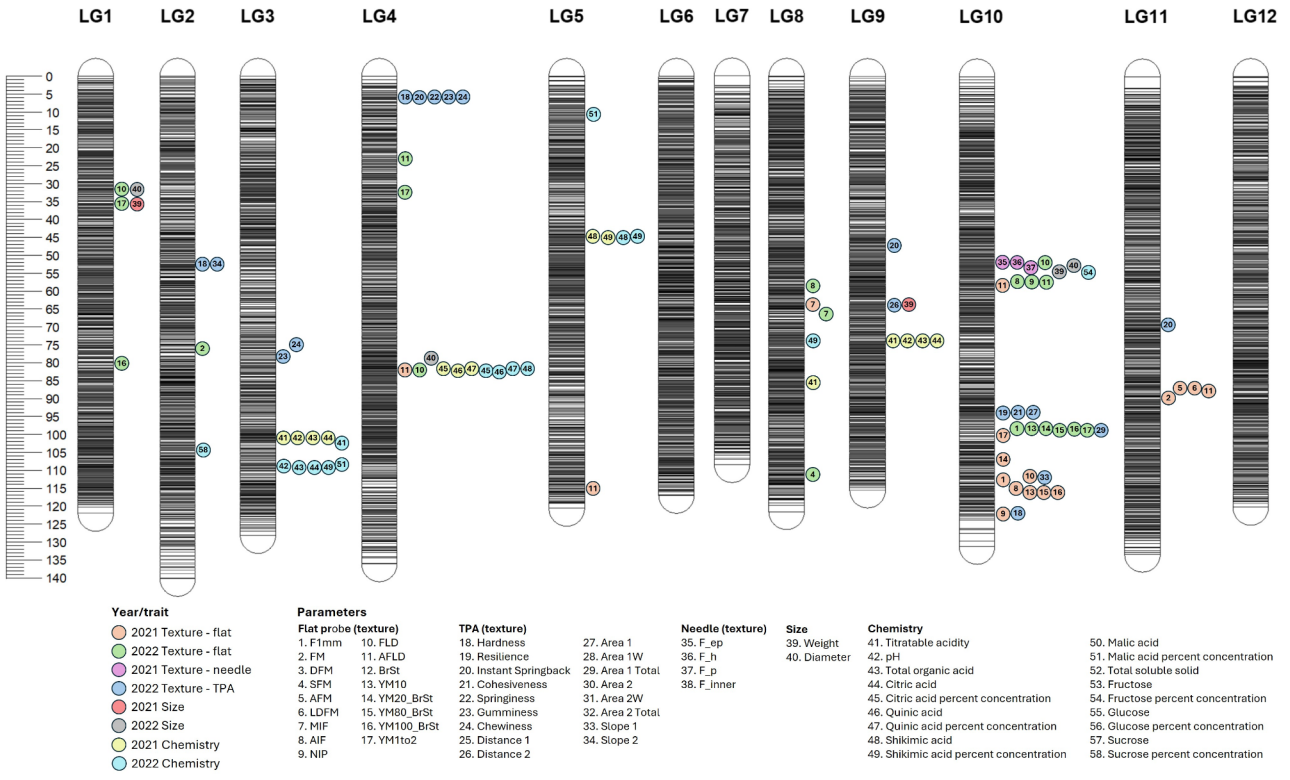

Fig. 6 Quantitative trait loci (QTL) detected for texture, size, and chemistry parameters measured at harvest (T_0). The circles represent the peak location of each QTL. The colors and numbers inside the circles represent the year/trait and the parameters, respectively

These parameters are closely associated with the resistance to external force before skin break [21] and were also highly related to the sensory perception of firmness [26]. Additionally, the F1mm and YM parameters were determined to be important parameters for predicting shelf-life texture change [21, 25]. Selecting for high F1mm and YMs at harvest could contribute to better post-storage mechanical texture.

Several QTL clusters for fruit texture were identified on LGs 1, 4, 6, and 10, on which 17, 14, 23, and 49 QTLs were mapped, respectively, including all years, time points, and parameters (Supplementary Table S3, Supplementary Figures S8-10). In these clusters, co-location between QTLs for different texture parameters that were not highly correlated with each other was observed at several locations. For example, a hotspot on LG10 included QTLs associated with flat probe parameters (F1mm, FLD, AIF, NIP, YM) and TPA parameters (hardness, resilience, cohesiveness, area 1, area 1 total, and slope 1) measured at T_0 (Fig. 6), which were not highly correlated with each other in all cases. This suggests that some texture QTLs might have pleiotropic effects controlling several parameters, perhaps due to the relationship between texture parameters and multiple biological processes that occur during ripening, such as cell wall modification, changes in turgor pressure, hormonal regulation, or changes in the biochemical constitution [77–82]. Fine-mapping these QTL clusters may provide further insight into these genomic regions. Also, the PVE values for all the texture parameter QTLs were relatively low (less than 15%) despite the relatively high heritability (≥ 60%) of the texture parameters, indicating that texture is a highly quantitative trait possibly controlled by multiple minor-effect QTLs. Ferrao et al., 2024 [27] evaluated blueberries using the flat probe parameters at T_0, as in this study. Several minor-effect QTLs with very few stable across years were found, similar to our findings. This suggests that using genomic selection might be a more suitable approach for breeding programs when targeting fruit texture.

Fruit size is an important trait for breeders as larger berry size leads to less water loss and wrinkling during shelf-life [21, 25], which could substantially contribute to consumers' acceptance and marketability. In this study, fruit weight and diameter were used as proxies for fruit size since they are highly correlated with fruit volume [83]. For these size-related parameters (i.e., fruit weight and diameter), we did not detect any QTLs that were consistently significant across years despite the relatively high heritability of fruit size ($H^2 \approx 70\%$). Small-effect QTLs on LGs 1, 4, 9, and 10 were detected in only one year (Supplementary Table S3, Supplementary Figure

S7), suggesting that size is a trait controlled by multiple minor-effect QTLs and possibly largely affected by environmental factors. As with texture and sugar accumulation, genomic selection could be a more suitable strategy to select for size. Also, testing QTLs in diverse environments to elucidate the influence of environment or genotype × environment interactions could benefit breeding strategies in the future.

Conclusion

Our comprehensive study assessed the genetic basis of organic acids, sugars, texture, size, and shelf-life in blueberry. All traits had moderate to high heritability, indicating that strong genetic factors interplay with the environment to control these traits. Major-effect QTLs controlling organic acid content and a number of underlying putative candidate genes were unveiled in this study. Additionally, QTLs for fruit texture were identified but had a lower effect, and no consistent QTL was identified for sugar content or size. Overall, the results of this study indicate that organic acids have a relatively simple genetic inheritance in blueberry, making these traits more suitable for MAS. In contrast, traits like size, texture, and sugar content have a more complex genetic architecture, making them more suitable for genomic selection. Our findings provide valuable information to facilitate DNA-informed selection in breeding programs.

Abbreviations

QTL	Quantitative trait locus
TA	Titratable acidity
TSS	Total soluble solids
MAS	Marker-assisted selection
R×A	'Reveille' × 'Arlen'
NCDA&CS	North Carolina Department of Agriculture and Consumer Services
NIRS	Near-infrared spectroscopy
FT-NIR	Fourier transform NIR
HPLC	High-pressure liquid chromatography
PLSR	Partial least squares regression
RPD	Residual prediction deviation
DAD	Diode array detector
TPA	Texture profile analysis
SI	Storage index
H^2	Broad-sense heritability
REML	Restricted maximum likelihood
SNP	Single nucleotide polymorphism
IBD	Identity-by-descent
BIC	Bayesian Information Criterion
PVE	Phenotypic variance explained
LG	Linkage group
qRT-PCR	Quantitative real-time PCR
UBC28	UBIQUITIN-CONJUGATING ENZYME
DEG	Differentially expressed gene
TCA cycle	Tricarboxylic acid cycle
CQA	Caffeoylquinic acid
KASP assay	Kompetitive Allele Specific PCR

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06061-4.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

We would like to thank Dr. Hudson Ashrafi, Dr. Mike Mainland, John Garner, Jessica Spencer, Jonathan Franck, and the NCDA&CS, Castle Hayne, NC, for providing access to and harvesting the blueberry material. We thank Joyce Edwards, Erin Deaton, Lorie Beale, Charles Warlick, and Brianna Haynes for their technical support. We thank Dr. Marcelo Mollinari for his valuable input during QTL analysis. We thank Marc Johnson (Texture Technologies Corp) and Randy Koch (Texture Guy, LLC) for providing help with establishing the instrumental texture measurement methodologies. Also, we would like to thank the reviewers for their invaluable comments.

Author contributions

HO, PPV, and MI conceived the idea and designed the experiments. HO and GM performed phenotyping. MFM, GM, LG, MP, PPV, and MI assisted with establishing the methodologies. HO analyzed and interpreted the data, and drafted the manuscript. MFM, GM, LG, MP, JS, PPV, and MI revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was funded by the United States Department of Agriculture National Institute of Food and Agriculture, award number 2019-51181-30015, project "VacciniumCAP: Leveraging genetic and genomic resources to enable development of blueberry and cranberry cultivars with improved fruit quality attributes.". MI was also supported by the United States Department of Agriculture National Institute of Food and Agriculture, Hatch project 1008691.

Data availability

Genotypic and phenotypic data are made available in the supplementary data (Supplementary Tables S8 and S9, respectively). The linkage map used in this study is available on the genome database for vaccinium (GDV; https://www.vaccinium.org/bio_data/1659687). RNA-seq raw data is available through NCBI short reads archive (SRA) project number PRJNA1148463 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1148463).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Plants for Human Health Institute, North Carolina State University, Kannapolis, NC 28081, USA
[2]Department of Horticulture, North Carolina State University, Raleigh, NC 27607, USA
[3]Agricultural Research Station, Virginia State University, Petersburg, VA 23806, USA
[4]Fondazione Edmund Mach, Research and Innovation Centre, San Michele a/A, Trento, Italy

Received: 11 September 2024 / Accepted: 6 January 2025
Published online: 10 January 2025

References

1. Edger PP, Iorizzo M, Bassil NV, Benevenuto J, Ferrão LFV, Giongo L, et al. There and back again; historical perspective and future directions for Vaccinium breeding and research studies. Hortic Res. 2022;9:1–297.

2. Brazelton C, Kayla Y, Bauer N. 2016 Global Blueberry Statistics and Intelligence Report. 2017.
3. Gallardo RK, Zhang Q, Dossett M, Polashock JJ, Rodriguez-Saona C, Vorsa N, et al. Breeding trait priorities of the blueberry industry in the United States and Canada. HortScience. 2018;53:1021–8.
4. Gilbert JL, Olmstead JW, Colquhoun TA, Levin LA, Clark DG, Moskowitz HR. Consumer-assisted selection of blueberry fruit quality traits. HortScience. 2014;49:864–73.
5. Sutton S, Sterns J. Blueberry economics: The costs of establishing and producing conventional blueberries in the Willamette Valley. 2020.
6. Safley CD, Cline WO, Mainland CM. Evaluating the profitability of blueberry production. 2012.
7. Rodgers A, Morgan K, Harri A. Technology adoption and risk preferences: the case of machine harvesting by Southeastern blueberry producers. J Food Distrib Res. 2017;48:1–21.
8. Zeng Z, Mollinari M, Pereira S, Olukolu BA, Yencho GC. Polyploid goes to genomics. 21st Century Pathol. 2022;2:113.
9. Varshney RK, Bohra A, Roorkiwal M, Barmukh R, Cowling WA, Chitikineni A, et al. Fast-forward breeding for a food-secure world. Trends Genet. 2021;37:1124–36.
10. Mengist MF, Bostan H, Young E, Kay KL, Gillitt N, Ballington J, et al. High-density linkage map construction and identification of loci regulating fruit quality traits in blueberry. Hortic Res. 2021;8:169.
11. Cappai F, Amadeu RR, Benevenuto J, Cullen R, Garcia A, Grossman A, et al. High-resolution linkage map and QTL analyses of fruit firmness in autotetraploid blueberry. Front Plant Sci. 2020;11:1–11.
12. Qi X, Ogden EL, Bostan H, Sargent DJ, Ward J, Gilbert J, et al. High-density linkage map construction and QTL identification in a diploid blueberry mapping population. Front Plant Sci. 2021;12:1–13.
13. Ferrão LFV, Benevenuto J, Oliveira I, de Cellon B, Olmstead C, Kirst J et al. M,. Insights into the genetic basis of blueberry fruit-related traits using diploid and polyploid models in a GWAS context. Front Ecol Evol. 2018;6 July.
14. Canales E, Gallardo RK, Iorizzo M, Munoz P, Ferrão LF, Luby C, et al. Willingness to pay for blueberries: sensory attributes, fruit quality traits, and consumers' characteristics. HortScience. 2024;59:1207–18.
15. Mengist MF, Grace MH, Mackey T, Munoz B, Pucker B, Bassil N et al. Dissecting the genetic basis of bioactive metabolites and fruit quality traits in blueberries (*Vaccinium corymbosum* L). Front Plant Sci. 2022;13 September.
16. Oh H, Mengist MF, Ma G, Spencer J, Perkins-Veazie P, Iorizzo M. Identification of QTLs for non-volatile chemical compounds in blueberry. HortScience. 2023;58:S196.
17. Magwaza LS, Opara UL. Analytical methods for determination of sugars and sweetness of horticultural products-A review. Sci Hortic (Amsterdam). 2015;184:179–92.
18. Moggia C, Graell J, Lara I, González G, Lobos GA. Firmness at harvest impacts postharvest fruit softening and internal browning development in mechanically damaged and non-damaged highbush blueberries (*Vaccinium corymbosum* L). Front Plant Sci. 2017;8:1–11.
19. NeSmith DS, Nunez-Barrios A, Prussia SE, Aggarwal D. Postharvest berry quality of six rabbiteye blueberry cultivars in response to temperature. J Am Pomol Soc. 2005;59:13–7.
20. Olmstead JW, Finn CE. Breeding highbush blueberry cultivars adapted to machine harvest for the fresh market. Horttechnology. 2014;24:290–4.
21. Oh H, Pottorff M, Giongo L, Mainland CM, Iorizzo M, Perkins-Veazie P. Exploring shelf-life predictability of appearance traits and fruit texture in blueberry. Postharvest Biol Technol. 2024;208 July 2023:112643.
22. Giongo L, Poncetta P, Loretti P, Costa F. Texture profiling of blueberries (*Vaccinium* spp.) during fruit development, ripening and storage. Postharvest Biol Technol. 2013;76:34–9.
23. Giongo L, Ajelli M, Pottorff M, Perkins-Veazie P, Iorizzo M. Comparative multi-parameters approach to dissect texture subcomponents of highbush blueberry cultivars at harvest and postharvest. Postharvest Biol Technol. 2022;183:August2021–111696.
24. Rivera S, Kerckhoffs H, Sofkova-Bobcheva S, Hutchins D, East A. Influence of harvest maturity and storage technology on mechanical properties of blueberries. Postharvest Biol Technol. 2022;191:111961.
25. Mengist MF, Pottorff M, Mackey T, Ferrao F, Casorzo G, Lila MA, et al. Assessing predictability of post-storage texture and appearance characteristics in blueberry at breeding population level. Postharvest Biol Technol. 2024;214:112964.
26. Oh H, Stapleton L, Giongo L, Johanningsmeier S, Mollinari M, Mainland CM, et al. Prediction of blueberry sensory texture attributes by integrating multiple instrumental measurements. Postharvest Biol Technol. 2024;218:113160.
27. Ferrão LFV, Azevedo C, Benevenuto J, Mengist MF, Luby C, Pottorff M, et al. Inference of the genetic basis of fruit texture in highbush blueberries using genome-wide association analyses. Hortic Res. 2024. https://doi.org/10.1093/hr/uhae233.
28. Perkins-Veazie P, Ma G, Oh H, Trandel-Hayse M, Bassil N, Luby C, et al. Development of a high throughput method to evaluate soluble sugar content of large sets of blueberry fruit. HortScience. 2022;57(1):S161. 9 Supplement (Part.
29. Costa F, Cappellin L, Fontanari M, Longhi S, Guerra W, Magnago P, et al. Texture dynamics during postharvest cold storage ripening in apple (*Malus* × *Domestica* Borkh). Postharvest Biol Technol. 2012;69:54–63.
30. Bourke PM, Voorrips RE, Hackett CA, Van Geest G, Willemsen JH, Arens P, et al. Detecting quantitative trait loci and exploring chromosomal pairing in autopolyploids using polyqtlR. Bioinformatics. 2021;37:3822–9.
31. da Silva Pereira G, Gemenet DC, Mollinari M, Olukolu BA, Wood JC, Diaz F, et al. Multiple QTL mapping in autopolyploids: a random-effect model approach with application in a hexaploid sweetpotato full-sib population. Genetics. 2020;215:579–95.
32. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90.
33. Mengist MF, Bostan H, De Paola D, Teresi SJ, Platts AE, Cremona G, et al. Autopolyploid inheritance and a heterozygous reciprocal translocation shape chromosome genetic behavior in tetraploid blueberry (*Vaccinium corymbosum*). New Phytol. 2023;237:1024–39.
34. Colle M, Leisner CP, Wai CM, Ou S, Bird KA, Wang J, et al. Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. Gigascience. 2019;8:1–15.
35. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.
36. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–9.
37. Love MI, Huber W, Anders S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:1–21.
38. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Von Mering C, et al. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol Biol Evol. 2017;34:2115–22.
39. Vashisth T, Johnson LK, Malladi A. An efficient RNA isolation procedure and identification of reference genes for normalization of gene expression in blueberry. Plant Cell Rep. 2011;30:2167–76.
40. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001;29:e45–45.
41. Etienne A, Génard M, Lobit P, Mbeguié-A-Mbéguié D, Bugaud C. What controls fleshy fruit acidity? A review of malate and citrate accumulation in fruit cells. J Exp Bot. 2013;64:1451–69.
42. Bett-Garber KL, Lea JM, Watson MA, Grimm CC, Lloyd SW, Beaulieu JC, et al. Flavor of fresh blueberry juice and the comparison to amount of sugars, acids, anthocyanidins, and physicochemical measurements. J Food Sci. 2015;80:S818–27.
43. Bett-Garber KL, Lea JM. Development of flavor lexicon for freshly pressed and processed blueberry juice. J Sens Stud. 2013;28:161–70.
44. Walker RP, Famiani F. Organic acids in fruits. Horticultural Reviews. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2018. pp. 371–430.
45. Leuschner C, Herrmann KM, Schultz G. The metabolism of quinate in pea roots: purification and partial characterization of a quinate hydrolyase. Plant Physiol. 1995;108:319–25.
46. Clifford MN, Jaganath IB, Ludwig IA, Crozier A. Chlorogenic acids and the acyl-quinic acids: Discovery, biosynthesis, bioavailability and bioactivity. Nat Prod Rep. 2017;34:1391–421.
47. Koshiro Y, Jackson MC, Nagai C, Ashihara H. Changes in the content of sugars and organic acids during ripening of *Coffea arabica* and *Coffea canephora* fruits. Eur Chem Bull. 2015;4:378–83.
48. Clifford MN, Kerimi A, Williamson G. Bioavailability and metabolism of chlorogenic acids (acyl-quinic acids) in humans. Compr Rev Food Sci Food Saf. 2020;19:1299–352.
49. Alcázar Magaña A, Kamimura N, Soumyanath A, Stevens JF, Maier CS. Caffeoylquinic acids: chemistry, biosynthesis, occurrence, analytical challenges, and bioactivity. Plant J. 2021;107:1299–319.

50. Tzin V, Galili G. New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. Mol Plant. 2010;3:956–72.
51. Semagn K, Babu R, Hearne S, Olsen M. Single nucleotide polymorphism genotyping using Kompetitive Allele specific PCR (KASP): overview of the technology and its application in crop improvement. Mol Breed. 2014;33:1–14.
52. Gilbert JL, Guthart MJ, Gezan SA, De Carvalho MP, Schwieterman ML, Colquhoun TA, et al. Identifying breeding priorities for blueberry flavor using biochemical, sensory, and genotype by environment analyses. PLoS ONE. 2015;10:1–21.
53. Retamales JB, Hancock JF. Blueberries. 2nd edition. CABI; 2018.
54. Ma B, Liao L, Fang T, Peng Q, Ogutu C, Zhou H, et al. A *Ma10* gene encoding P-type ATPase is involved in fruit organic acid accumulation in apple. Plant Biotechnol J. 2019;17:674–86.
55. Shi CY, Song RQ, Hu XM, Liu X, Jin LF, Liu YZ. Citrus PH5-like H+-ATPase genes: Identification and transcript analysis to investigate their possible relationship with citrate accumulation in fruits. Front Plant Sci. 2015;6 MAR.
56. Shi CY, Hussain SB, Yang H, Bai YX, Khan MA, Liu YZ. CsPH8, a P-type proton pump gene, plays a key role in the diversity of citric acid accumulation in citrus fruits. Plant Sci. 2019;289:110288.
57. Forney CF, Kalt W, Jordan MA, Vinqvist-Tymchuk MR, Fillmore SAE. Compositional changes in blueberry and cranberry fruit during ripening. Acta Hortic. 2012;926:331–8.
58. Forney CF, Kalt W, Jordan MA, Vinqvist-Tymchuk MR, Fillmore SAE. Blueberry and cranberry fruit composition during development. J Berry Res. 2012;2:169–77.
59. Li X, Li C, Sun J, Jackson A. Dynamic changes of enzymes involved in sugar and organic acid level modification during blueberry fruit maturation. Food Chem. 2020;309(September 2019):125617.
60. Cheng H, Kong W, Tang T, Ren K, Zhang K, Wei H et al. Identification of key gene networks controlling soluble sugar and organic acid metabolism during oriental melon fruit development by integrated analysis of metabolic and transcriptomic analyses. Front Plant Sci. 2022;13 May.
61. Umer MJ, Bin Safdar L, Gebremeskel H, Zhao S, Yuan P, Zhu H et al. Identification of key gene networks controlling organic acid and sugar metabolism during watermelon fruit development by integrating metabolic phenotypes and gene expression profiles. Hortic Res. 2020;7.
62. Etienne C, Moing A, Dirlewanger E, Raymond P, Monet R, Rothan C. Isolation and characterization of six peach cDNAs encoding key proteins in organic acid metabolism and solute accumulation: involvement in regulating peach fruit acidity. Physiol Plant. 2002;114:259–70.
63. Li N, Wang J, Wang B, Huang S, Hu J, Yang T, et al. Identification of the carbohydrate and organic acid metabolism genes responsible for Brix in tomato fruit by transcriptome and metabolome analysis. Front Genet. 2021;12:1–16.
64. Lin Q, Wang C, Dong W, Jiang Q, Wang D, Li S, et al. Transcriptome and metabolome analyses of sugar and organic acid metabolism in Ponkan (*Citrus reticulata*) fruit during fruit maturation. Gene. 2015;554:64–74.
65. Núñez-Lillo G, Lillo-Carmona V, Pérez-Donoso AG, Pedreschi R, Campos-Vargas R, Meneses C. Fruit sugar hub: gene regulatory network associated with soluble solids content (SSC) in Prunus persica. Biol Res. 2024;57:1–11.
66. Vizzotto G, Falchi R. Genetics of sugar and starch metabolism. The Kiwifruit Genome. Springer; 2016. pp. 189–204.
67. Argyris JM, Díaz A, Ruggieri V, Fernández M, Jahrmann T, Gibon Y, et al. QTL analyses in multiple populations employed for the fine mapping and identification of candidate genes at a locus affecting sugar accumulation in melon (*Cucumis melo* L). Front Plant Sci. 2017;8:1–20.
68. Ma B, Zhao S, Wu B, Wang D, Peng Q, Owiti A, et al. Construction of a high density linkage map and its application in the identification of QTLs for soluble sugar and organic acid components in apple. Tree Genet Genomes. 2016;12:1–10.
69. Khefifi H, Dumont D, Costantino G, Doligez A, Brito AC, Bérard A et al. Mapping of QTLs for citrus quality traits throughout the fruit maturation process on clementine (*Citrus reticulata* × *C. sinensis*) and mandarin (*C. Reticulata* Blanco) genetic maps. Tree Genet Genomes. 2022;18.
70. Chen J, Wang N, Fang LC, Liang ZC, Li SH, Wu BH. Construction of a high-density genetic map and QTLs mapping for sugars and acids in grape berries. BMC Plant Biol. 2015;15:1–14.
71. Mamani M, LÓpez ME, Correa J, Ravest G, Hinrichsen P. Identification of stable quantitative trait loci and candidate genes for sweetness and acidity in tablegrape using a highly saturated single-nucleotide polymorphism-based linkage map. Aust J Grape Wine Res. 2021;27:308–24.
72. Bayo-Canha A, Costantini L, Fernández-Fernández JI, Martínez-Cutillas A, Ruiz-García L. QTLs related to berry acidity identified in a wine grapevine population grown in warm weather. Plant Mol Biol Rep. 2019;37:157–69.
73. Zhao H, Zhang T, Meng X, Song J, Zhang C, Gao P. Genetic mapping and QTL analysis of fruit traits in melon (*Cucumis melo* L). Curr Issues Mol Biol. 2023;45:3419–33.
74. Obando-Ulloa JM, Eduardo I, Monforte AJ, Fernández-Trujillo JP. Identification of QTLs related to sugar and organic acid composition in melon using near-isogenic lines. Sci Hortic (Amsterdam). 2009;121:425–33.
75. Zeballos JL, Abidi W, Giménez R, Monforte AJ, Moreno MÁ, Gogorcena Y. Mapping QTLs associated with fruit quality traits in peach [*Prunus persica* (L.) Batsch] using SNP maps. Tree Genet Genomes. 2016;12.
76. Fall LA, Perkins-Veazie P, Ma G, McGregor C. QTLs associated with flesh quality traits in an elite × elite watermelon population. Euphytica. 2019;215:1–14.
77. Paniagua AC, East AR, Hindmarsh JP, Heyes JA. Moisture loss is the major cause of firmness change during postharvest storage of blueberry. Postharvest Biol Technol. 2013;79:13–9.
78. Saladié M, Matas AJ, Isaacson T, Jenks MA, Goodwin SM, Niklas KJ, et al. A reevaluation of the key factors that influence tomato fruit softening and integrity. Plant Physiol. 2007;144:1012–28.
79. Montecchiarini ML, Silva-Sanzana C, Valderramo L, Alemano S, Gollán A, Rivadeneira MF, et al. Biochemical differences in the skin of two blueberries (Vaccinium corymbosum) varieties with contrasting firmness: implication of ions, metabolites and cell wall related proteins in two developmental stages. Plant Physiol Biochem. 2021;162:483–95.
80. Wang S, Zhou Q, Zhou X, Wei B, Ji S. The effect of ethylene absorbent treatment on the softening of blueberry fruit. Food Chem. 2018;246:286–94.
81. Zhou Q, Zhang F, Ji S, Dai H, Zhou X, Wei B, et al. Abscisic acid accelerates postharvest blueberry fruit softening by promoting cell wall metabolism. Sci Hortic (Amsterdam). 2021;288(February):110325.
82. Oh HD, Yu DJ, Chung SW, Chea S, Lee HJ. Abscisic acid stimulates anthocyanin accumulation in 'Jersey' highbush blueberry fruits during ripening. Food Chem. 2018;244:403–7.
83. Mengist MF, Grace MH, Xiong J, Kay CD, Bassil N, Hummer K et al. Diversity in metabolites and fruit quality traits in blueberry enables ploidy and species differentiation and establishes a strategy for future genetic studies. Front Plant Sci. 2020;11 April.

Publisher's note