ORIGINAL ARTICLE

A low malic acid trait in cranberry fruit: genetics, molecular mapping, and interaction with a citric acid locus

Stephanie Kay Fong[1] · Joseph Kawash[2] · Yifei Wang[3] · Jennifer Johnson-Cicalese[4] · James Polashock[5] · Nicholi Vorsa[1] ✆

Abstract

The fruit of commercial cranberry (*Vaccinium macrocarpon* Ait.) cultivars have relatively high concentrations of malic acid (MA) and citric acid (CA), and, to a lesser extent, quinic acid. These acids contribute to the high titratable acidity (TA), a measure of tartness, of cranberry fruit, which typically ranges from 2.3 to 2.5% citric acid equivalents in commercial cultivars. Thus, considerable amounts of sugar are added ("added sugar") in products such as sweetened-dried cranberries and juices. Within our cranberry germplasm collection, a unique accession was identified with fruit having a TA ≈ 1.5%, where MA concentration was reduced to ~ 4 mg/g fresh weight (FW), compared to ~ 8 mg/g FW in current cultivars. Inbred crosses derived from this accession yielded progeny with a lower MA phenotype (~ 2 mg/g FW). Observed segregation indicated the accession was heterozygous (*Mala/mala*) for a low MA allele (*mala*) and very low MA progeny were homozygous (*mala/mala*). MA was reduced approximately 75% in these populations relative to standard cultivars. The homozygous *mala/mala* locus also depressed fruit CA and quinic acid concentrations. Quantitative trait loci mapping identified a region on chromosome 4 associated with low MA. The combined segregation of three half-sib populations derived from the low MA accession generated effective (within < 1 cM) Kompetitive allele-specific PCR (KASP) markers for use in breeding of cranberry cultivars with lower MA. Dihybrid populations were developed having a previously described low CA allele, *cita*, with the *mala* alleles of this study to explore the interaction of alleles at both loci.

✉ Nicholi Vorsa
vorsa@njaes.rutgers.edu

[1] Department of Plant Biology, Rutgers University, 59 Dudley Road, Foran Hall, New Brunswick, NJ 08901, USA

[2] ORISE Scholar and USDA-ARS, Genetic Improvement of Fruits and Vegetables Lab, Chatsworth, NJ 08019, USA

[3] USDA-ARS, Beltsville Human Nutrition Research Center, Methods and Application of Food Composition Laboratory, Beltsville, MD 20705, USA

[4] Philip E. Marucci Center for Blueberry and Cranberry Research and Extension, Rutgers University, Chatsworth, NJ 08019, USA

[5] USDA-ARS, Genetic Improvement of Fruits and Vegetables Lab, Chatsworth, NJ 08019, USA

Introduction

While cranberry (*Vaccinium macrocarpon* Ait.) fruit is well known for its health benefits, significant amounts of added sugars are required to balance the high acid content for palatability of cranberry products (juices, sweetened dried cranberries, sauces, etc.). In fact, most cranberry products typically contain up to 40% of added sugars (Ocean Spray Cranberries Inc. 2019). Fruit juices are formulated using measures of Brix (soluble solids) and titratable acidity (TA), as palatability is related to the sugar-acid ratio (Brix:TA) (Bates et al. 2001). However, too little acidity can cause bland flavor, as expression of fruit flavor requires a minimum level of acid, typically between 0.5 and 1.0% TA (Bates et al. 2001). Brix is a popular industry measure of sugar content; however, in fruit with high acidity, the value also includes constituents such as organic

acids. TA is measured as citric acid equivalents (CAE) in cranberry and is a measure of acidity or fruit tartness. Cranberry fruits are low in sugar (combined glucose, fructose, and sucrose of ~ 3%) relative to other fruits, such as blueberry, with typically > 6% total sugars (Ocean Spray Cranberries, Inc. 2019; Forney et al. 2012). Furthermore, fructose is typically equal to or greater in concentration than glucose in fruits consumed fresh; however, in cranberry, fructose is greatly reduced relative to glucose.

Cranberry fruit has three organic acids that contribute to TA: malic (MA), citric (CA), and to a lesser extent quinic acid (QA) (Cunningham et al. 2003; Wang et al. 2017), with MA typically ranging between 0.6 and 1.0%. Modeling with malic acid in peach showed that both malic and citric acids are significant contributors to TA (Lobit et al. 2002). Thus, reducing the MA content in cranberry should decrease TA. Additionally, if acidity of cranberry is reduced, it could be possible to increase the percentage of cranberry juice in cranberry juice products without increasing sugar content. This would potentially enhance the relative amounts of bioactive health compounds in cranberry products.

The perception of "tartness" and astringency depends on specific acids, various acid compositions, and the consumer (Rubico and McDaniel 1992). MA, a dicarboxylic acid, has a pKa of 3.4 and 5.2 (pKa_1 and pKa_2, respectively) and is approximately 90% as sour as CA (Kader 2008). A common and naturally occurring compound, MA contributes to the refreshing tart taste of various fruit species, such as apple, in which malic is the primary fruit acid (Zhang et al. 2010). MA concentrations in cranberry fruit tend to remain stable during fruit development, typically ranging 6–8 mg/g (Wang et al. 2017). CA is a tricarboxylic acid with a pKa of 3.13, 4.76, and 6.39 (pKa_1, pKa_2, and pKa_3, respectively). CA contributes to the sharp sour taste, most notably in citrus fruits, in which CA is the primary acid. CA is the principal acid in other *Vaccinium* species, e.g., highbush blueberry (Wang et al. 2019; Forney et al. 2012). Like MA, CA concentrations in cranberry fruit tend to remain stable during fruit development, typically ranging 6–12 mg/g (Wang et al. 2017). Additionally, a low CA trait has been well characterized at the *CITA* locus in cranberry (Fong et al. 2020). QA is a sugar acid with a pKa of 3.46 and contributes to the bitter astringent taste in cranberry fruit. QA is not found in many fruits, but is found in peach, *Prunus persica* (Etienne et al. 2002b; Bae et al. 2014) and kiwi fruit, *Actinidia* species (Nishiyama et al. 2008; Marsh et al. 2009). It is also reported to contribute to the bitter taste found in coffee (McCamey et al. 1990). QA concentrations in cranberry tend to decrease during fruit development (Wang et al. 2017).

A low MA trait was observed in apples (*Malus domestica*), measured by pH and TA, and was recessively inherited (Xu et al. 2012). Fine mapping of the low MA trait in apple revealed aluminum-activated malate transporter-like genes, which maintain malate homeostasis (Bai et al. 2012). In tomato, there was also evidence that acidity is controlled by aluminum-activated malate transporter-like genes (Ye et al. 2017). MA is also found in peaches (*Prunus persica*), with low acid varieties having both lower MA and CA concentrations (Moing et al. 1998). Molecular mapping has identified multiple markers linked with the low acid trait, which is primarily due to lower MA (Etienne et al. 2002b; Boudehri et al. 2009; Eduardo et al. 2014; Lambert et al. 2016; Zeballos et al. 2016). Several studies have identified candidate genes for the low acid trait in peaches, which include an auxin efflux carrier, malate dehydrogenase, and a tonoplast proton pump (Etienne et al. 2002a, b; Cao et al. 2016). Although cultivated sweet melon (*Cucumis melo* L.) has very low levels of CA and MA, genetic variation exists in the germplasm (Cohen et al. 2012). In Cohen et al. (2012), there was a lack of co-localization of candidate genes controlling acidity and the QTL identified.

In cranberry ($2x = 2n = 24$), there have been a number of genetic and genomic studies conducted. An updated genome sequence is available (Kawash, unpublished data), and multiple genetic maps and QTL studies have been published (Covarrubias-Pazaran et al. 2016; Daverdin et al. 2017; Diaz-Garcia et al. 2018; Georgi et al. 2013; Polashock et al. 2014; Schlautman et al. 2015 and 2017). These studies have constructed high-resolution linkage maps co-linear with the reference genome sequence, facilitating further genomic studies. Additionally, there were 3 QTL for TA across 3 linkage groups identified in Georgi et al. (2013). However, these three QTLs were not detected in Diaz-Garcia et al. (2018) where there were 17 QTL across 7 linkage groups identified. Additionally, we have characterized another locus (*CITA*) influencing fruit TA *via* CA concentration with an allele *cita*, which, when homozygous, results in a low CA phenotype (Fong et al. 2020). However, a negative correlation between CA and MA in these populations resulted in higher MA in the low CA progeny, attenuating the decrease in TA (Fong et al. 2020).

A native germplasm accession, NJ93-57, with reduced MA concentration was previously identified in the cranberry germplasm collection at the Marucci Blueberry and Cranberry Research and Extension Center, Chatsworth, NJ (Cunningham and Vorsa unpublished data). This study describes the inheritance of a low MA trait originating from NJ93-57. A series of breeding cycles indicated a partially recessive qualitative trait, modulated by the locus *MALA*. A low MA allele, *mala*, was derived from the heterozygous accession NJ93-57.

As the first characterization of the genetics of a qualitative low MA trait in cranberry fruit, the objectives of this study were to (1) describe the inheritance of the *mala* allele, its effect on TA and MA, and its relationship with CA and QA; (2) identify, develop, and validate molecular markers, e.g., Kompetitive allele-specific PCR (KASP), closely linked to *MALA* locus for identification of the *mala* allele in marker-

assisted selection; and (3) determine the effect of the *MALA* locus allele, *mala*, in genotypes having the *cita* allele at the *CITA* locus described in Fong et al. (2020).

Materials and methods

Plant material

The germplasm accession NJ93-57 was collected from a native cranberry population in Suffolk County, NY, in 1993, and exhibited lower MA (4.05 mg/g) in a subsequent germplasm screen. An initial cross, NJ93-57 × cv. Mullica Queen (MQ), was made in 2004 to generate the CNJ04-52 population (Fig. 1). From this population, an individual with lower TA (CNJ04-52-46) was self-pollinated to generate the CNJ08-100 (hereafter MA100) population (Fig. 1 and S1). One additional full-sib of CNJ04-52-46 and CNJ04-52-54, also with lower TA, was also suspected of having the genes, i.e., alleles, for low MA concentrations. CNJ04-52-46 and CNJ04-52-54 were crossed with the low CA germplasm accession (NJ91-7-12) previously defined as homozygous for the low CA allele, *cita/cita* at the *CITA* locus (Fong et al. 2020), to generate CNJ08-98 and CNJ08-103 populations, respectively (Fig. 1). Progeny from these populations were evaluated for CA, MA, and QA concentrations. Individuals with the lowest MA concentrations (CNJ08-103-20 [6.0 mg/g], CNJ08-98-80 [5.5 mg/g], and CNJ08-98-3 [5.3 mg/g]) were selected for subsequent crosses (Fig. 1). In 2012, CNJ08-103-20 was self-pollinated to give population CNJ12-155 (CM155, named CM for containing both *cita* and *mala* alleles). Additionally,

the cross CNJ08-98-80 × CNJ08-103-20 gave population CNJ12-151 (CM151). CM155 and CM151 are half-sib populations with the common parent CNJ08-103-20 (Fig. 1 and Table S1, S2).

Additional F$_1$ populations (in addition to CNJ04-52) produced were CNJ04-13 (MQ × NJ91-7-12) and CNJ04-34 (cv. Crimson Queen (CQ) × NJ93-57). Third-generation cycle populations CNJ12-92 (CM92) and CNJ12-93 (CM93) were generated by intercrossing the second generation as shown in Fig. 1. CM92 and CM93 were derived from CNJ08-30-20 (a homozygous genotype *cita/cita*, from self-pollination of CNJ04-13-39), as a seed parent, crossed with the pollen parents CNJ08-90-7 and CNJ08-98-3, respectively (both derived from a backcross to NJ91-7-12, a homozygous *cita/cita* genotype) (Fig. 1 and Table S1, S2).

Crosses were made manually during April and May in the greenhouse in 2004–2012; flowers of the maternal parent were emasculated 1–2 days prior to anthesis and pollinated with pollen from the paternal parent 5–7 days post-emasculation. Seeds were germinated in the greenhouse after stratification at ~ 2 °C for approximately 3 months. Seedlings were transplanted into 108 cm^2 pots containing peat and sand at 1:1 v/v and maintained in the greenhouse and grown for at least 3 years. Populations flowered and set fruit during the years 2014–2018. During the spring/summer of 2017 and 2018, flowering plants were taken outside for bee open-pollination. After pollination and subsequent fruit set was completed, plants were brought back into the greenhouse for fruit development and ripening. Fruit was collected from each individual in each population once a year during late August and September in 2014–2018 for analysis of organic acids and
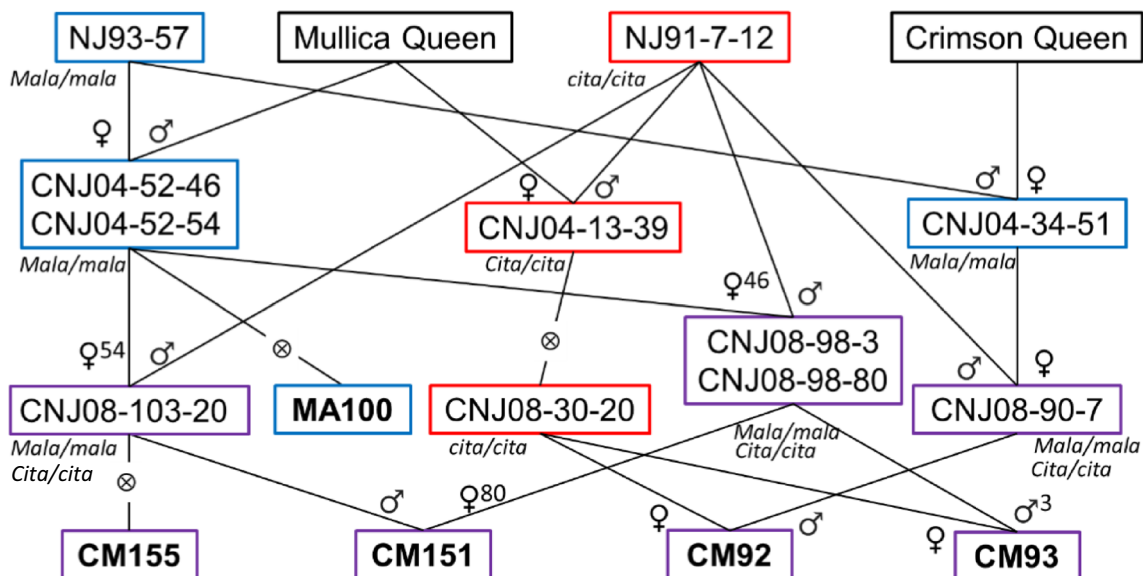

Fig. 1 Pedigree of populations (in bold) used in this study: **MA100**, **CM155**, **CM151**, **CM92**, and **CM93**. Red boxes indicate *cita* in background, blue boxes indicate *mala* in background, black boxes indicate neither *cita* nor *mala* in background, and purple boxes indicate both *cita* and *mala* alleles in background. Genotypes for *CITA* and/or *MALA* given when alleles *cita* and *mala* are present; otherwise, if alleles are not specified, then genotypes are *Cita/Cita* and/or *Mala/Mala*

TA. Leaf tissue was collected in the spring of 2017 for DNA extractions.

Organic acid and titratable acidity analysis

Fruit organic acids (MA, CA, QA) were extracted and analyzed with high-performance liquid chromatography (HPLC) as in Wang et al. (2017), with modifications as follows: Approximately 3–5 g of frozen fruit was used per sample (individual population, with 2–3 fruits per sample), depending on fruit weight available. The fruit was ground with a Precellys Evolution homogenizer (Bertin Corp., Rockville, MD, USA) using 2.8-mm ceramic beads at 7200 rpm for 1.5 min. Ten milliliters of distilled water for every gram of fruit was added to suspend the homogenized fruit. One milliliter of the slurry was centrifuged at 10,000 rpm for 5 min; the supernatant was heated to 90 ° C for 10 min, and then frozen at - 80 °C until analyzed with HPLC on a Dionex HPLC system with an AS50 Autosampler, AS50 Thermal Compartment, PDA-100 Detector, and a GP40 Gradient Pump (Dionex Corporation, Sunnyvale, CA, USA). The remainder of the slurry was used for TA and Brix (soluble solids) analysis as in Vorsa and Johnson-Cicalese (2012). TA was quantified by titrating to an endpoint of pH 8.2 with .05 N NaOH using a Metrohm Ti-Touch 916 (Metrohm AG, Riverview, FL, USA). The final percent TA was calculated using CAE. Brix was measured with an Atago PR-32 digital refractometer (Atago USA, Inc., Bellevue, WA, USA). HPLC and TA were performed on fruit from individuals in the CM151, CM155, CM92, and CM93 populations, collected over 2016–2018 crop years. Most progeny were sampled for 2 years (2016–2017), for two replications per progeny, with samples collected in 2018 used to fill in for individuals that did not produce fruit from the prior 2 years. HPLC was performed for MA100 fruit harvested in 2014–2017, while TA was performed only in 2016.

Genotyping

DNA was extracted from leaf tissue with a modified CTAB protocol and GBS libraries for MA100, CM151, and CM155 were generated as in Daverdin et al. (2017). The GBS libraries for all the populations used Msp1 and Pst1 as the restriction enzymes (New England Biolabs Inc., Ipswich, MA, USA). After the initial analysis, additional GBS libraries were generated for CM155 using the restriction enzymes Nde1 and Pst1 to yield more reads in the region of interest. Prepared GBS libraries were sent to Genewiz, Inc. (South Plainfield, NJ, USA) for sequencing on the Illumina Hi-seq platform (Illumina Inc., San Diego, CA, USA). The initial sequencing run was on a 2 × 100 bp configuration, with later sequencing runs on a 2 × 150 bp paired-end configuration.

All populations were also genotyped with SSRs (scf258d and scf153722) linked to the *CITA* locus for low CA trait as in Fong et al. (2020). After identification of SNP markers through QTL mapping, SNP genotyping was conducted using custom KASP (Kompetitive allele-specific PCR) assays designed by LGC Biosearch Technologies (Beverly, MA, USA) using sequence data provided. SSR primer sequences and SNP regions are shown in Table S3. Sensitivity and specificity were calculated for every marker as in Rosas et al. (2014), following classifications of true negatives, false negatives, false positives, and true negatives as in Fong et al. (2020).

QTL identification

Barcoded samples were de-multiplexed using STACKS and aligned to the cranberry reference genome (Kawash et al. unpublished) with bwa-mem (Catchen et al. 2011; Li and Durbin 2009). Samtools was used to call SNPs (Li 2011). Qualifying SNPs required a read support of 4 reads, and heterogeneity between 25 and 75%. Missing data was limited to only 10% of the population in a given marker, and markers that were homogenous through the population were also removed. R/qtl was used to calculate genetic distance between markers and identify QTL (Broman et al. 2003). Genome-wide significance of LOD scores was calculated at $p < 0.05$ through 1000 permutations. Gene prediction was performed using MAKER (Cantarel et al. 2008), using transcriptome information from Polashock et al. (2014) as training data. Function prediction of resulting genes from MAKER was performed using tblastx against the NCBI NR database and linked to candidate genes placed within the QTL region.

Statistical analyses

PROC GLM and CORR were used in SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) for means separations, regression analysis, and ANOVA (CM151 and CM155 populations pooled); CA and MA levels were analyzed as fixed variables, type III SS interaction error term used to test *CITA* and *MALA* loci main effects for MA, CA, QA, and TA, and year-to-year variation and correlations. Chi-square values were calculated using Chisq-.test in R.

Results

Relationship of MA, CA, QA, and TA

The MA population means for populations CM151 and CM155 were significantly different from each other (5.5 and 4.2 mg/g, respectively), but not from MA100 (5.1 mg/g) (Fig. S2a). The CM151 and CM155 populations, however, did have individuals with considerably lower MA (range of 0.4–

8.9 and 0.1–6.8 mg/g, respectively), as compared to MA100 (1.3–9.0 mg/g). The CA mean concentration and range were significantly lower in the CM151 and CM155 populations than in MA100 (Fig. S2b).

In all three populations, variation between harvest years was not significant for CA, MA, and TA. However, there was significant variation between years for QA (F-value = 45.1, $p < 0.05$). A significant correlation between 2017 and 2018 MA levels was observed for populations CM151 and CM155 combined ($r = 0.84$, $p < 0.0001$).

MA CM92 and CM93 population means (5.8 and 6.3 mg/g, respectively) were significantly higher than for CM155 (4.1 mg/g) (Fig. S2a). CA CM92 population mean (2.8 mg/g) was significantly lower than population means of MA100, CM151, CM155, and CM93. CA population means for CM155 and CM93 were significantly lower than for MA100 (Fig. S2b). The maximum CA concentration for CM92 was lower than CM93 (7.0 and 9.8 mg/g, respectively). TA CM92 and CM93population means (1.6% and 1.8% CAE, respectively) were intermediate to TA means for MA100 versus CM155 (Fig. S2c). QA population means were similar for

CM92 and CM93 (12.5 mg/g and 13.4 mg/g, respectively) and significantly higher compared to CM155 (Fig. S2d). There was significant year-to-year variation, but the year by genotype interaction was not significant (data not shown) for MA, CA, QA, and TA for CM92 and CM93A and low but significant correlation for MA between the 2 years was found ($r = 0.24$, $p < 0.01$ and $r = 0.27$, $p < 0.01$, respectively) for CM92 and CM93.

Concentrations of MA, CA, QA, and TA appeared to exhibit a bimodal distribution for MA100, CM151, and CM155 (Fig. 2a, b, c, d). In the CM92 and CM93 populations, the distributions for MA, QA, and TA generally appeared to follow a normal distribution while the distribution for CA could be considered bimodal (Fig. 2e, f, g, h).

QA, MA, and CA were significantly positively correlated with TA in MA100, CM151, and CM155 (Table 1). There was also a strong significant positive correlation between QA and MA for MA100, CM151, and CM155 populations ($r = 0.52$, 0.55, and 0.79, respectively). However, only MA100 showed a positive correlation between CA and MA ($r = 0.76$). For MA100 and CM155, there was a significant positive

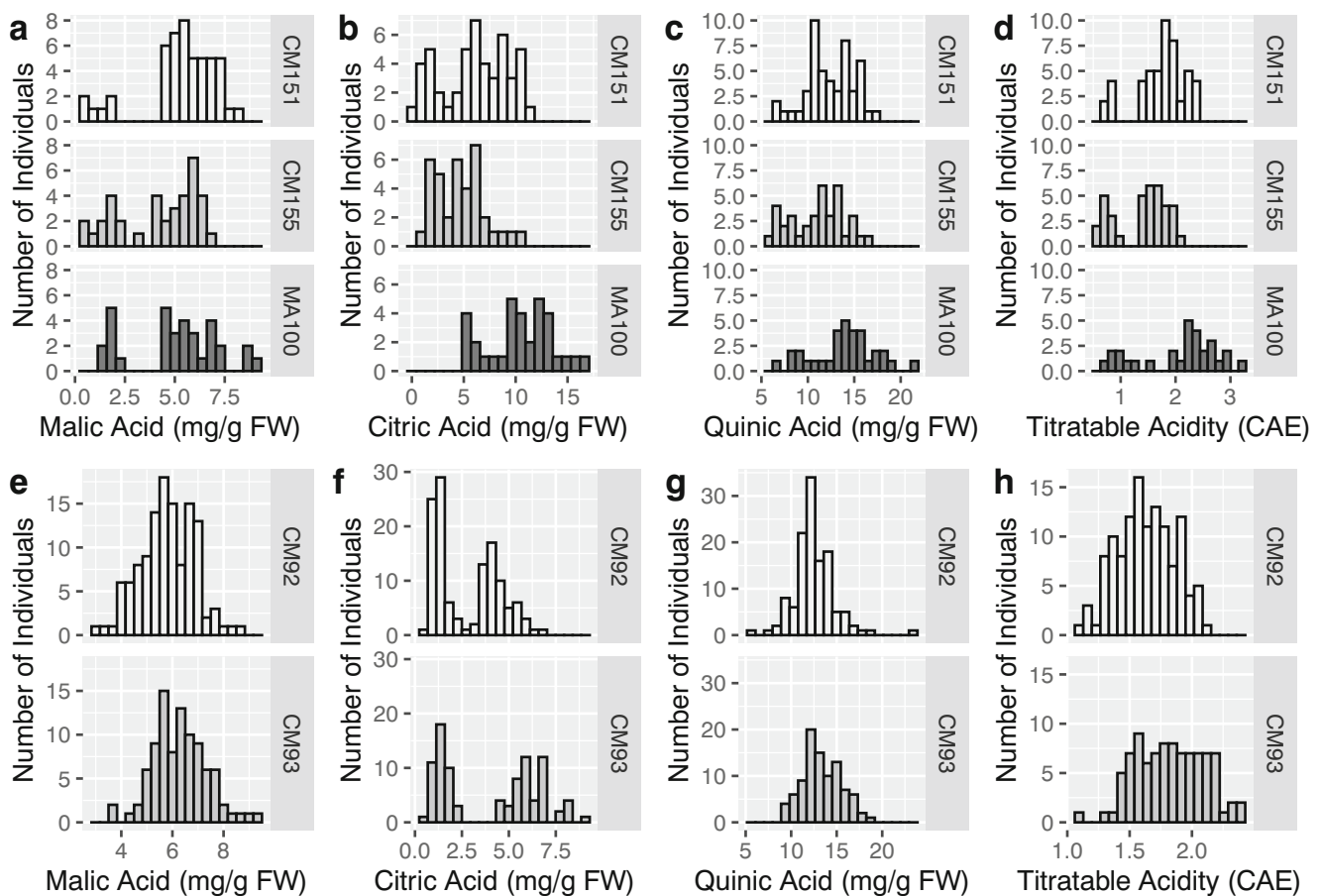**Fig. 2** Distributions of malic acid (**a**), citric acid (**b**), and quinic acid (**c**) concentrations and titratable acidity (**d**) for populations MA100 (mean from 2014–2016), CM155, and CM151 (means from 2016–2017).

Distributions of malic acid (**e**), citric acid (**f**), and quinic acid (**g**) concentrations and titratable acidity (**h**) for populations CM92 and CM93 (means from 2016-2017)

Table 1 Pearson's correlation coefficients and significance between quinic acid, malic acid, citric acid, TA, and Brix for MA100, CM151, and CM155. Due to the lack of significant variation between years, means of 2 years (CM151 and CM155) or 3 years (MA100) of data were used in the analysis. Quinic, malic, and citric acids were measured in mg/g fruit weight, TA was in % citric acid equivalents, and Brix was in % soluble solids

	Population	Quinic	Malic	Citric	TA	Brix
Quinic	MA100					
	CM151					
	CM155					
Malic	MA100	0.523**				
	CM151	0.553***				
	CM155	0.791***				
Citric	MA100	0.593**	0.758***			
	CM151	0.24	0.248			
	CM155	0.342*	0.148			
TA	MA100	0.523**	0.859***	0.817***		
	CM151	0.547***	0.743***	0.772***		
	CM155	0.828***	0.860***	0.587**		
Brix	CM151	0.160	0.442**	0.260	0.327*	
	CM155	0.178	0.204	0.345*	0.334	

*** $p < 0.005$, ** $p < 0.005$, * $p < 0.05$

correlation ($r = 0.59$ and 0.34, respectively) between QA and CA. CM155 also had a modest significant positive correlation ($r = 0.35$, respectively) between Brix and CA. CM151 had significant positive correlations between Brix and MA ($r = 0.44$), and TA ($r = 0.33$). CM92 and CM93 had significant differences between years for all five traits (however, there was no significant interaction between year and genotype); thus, the correlations were analyzed by year (Table S4). The only significant positive correlations consistently observed both years (2016 and 2017) were between MA and TA and between CA and TA, in both CM92 and CM93 (Table S3).

Multiple regression analysis for populations MA100, CM151, and CM155 indicated that variation in MA and CA contributed the most to the variation in TA. For MA100, CA and MA accounted for 81% of the variation in TA. For CM151, CA and MA accounted for 92% of the variation in TA. For CM155, CA and MA accounted for 96% of the variation in TA.

Inheritance of low MA trait

The distribution of MA for MA100, CM151, and CM155 suggested that there is a threshold at < 2.5 mg/g MA (phenotypes between 2.5 and 3.0 mg/g essentially lacking) to be considered low MA (Fig. 2a). All three populations exhibited bimodal MA distributions, indicating the presence of a single locus affecting the low MA phenotype, herein referred to as the *MALA* locus (Fig. 2a). In the populations, there was segregation for two phenotypic levels, low MA at < 2.5 mg/g and "normal" MA at > 2.5 mg/g (Table 2). MA100 segregation was consistent with expected 3:1 ($\chi^2 = 0.01$, $p = 0.92$) for normal MA to low MA (< 2 mg/g) (Table 2). CM155 also fit 3:1 segregation for MA ($\chi^2 = 0.88$, $p = 0.35$) (Table 2). However, the observed segregation for MA concentration in CM151 was significantly different from the expected 3:1 ratio, with a deficiency of the low MA phenotype ($\chi^2 = 4.3$, $p = 0.04$).

When CM151 and CM155 are analyzed as dihybrid crosses involving two loci (*CITA* and *MALA*), segregation did not significantly deviate from a 9:3:3:1 segregation ratio for a dihybrid cross with two independently segregating traits (CM151: $\chi^2 = 7.0$, $p = 0.07$, CM155: $\chi^2 = 6.6$, $p = 0.09$) (Table 3). In CM92 and CM93 (*Cita/cita* x *cita/cita* populations), the distribution of CA fit the expected bimodal distribution, while the distribution for MA suggests a unimodal distribution (Fig. 2e). Both CM92 and CM93 did not

Table 2 Observed and expected segregation ratios and chi-square analysis for each population (MA100, CM151, CM155) segregating for fruit malic acid concentration (mg/g)

Cross	Observed phenotype (mg/g)		Expected phenotype (mg/g)		X^2	DF	p value
	MA < 2	MA > 4	MA < 2	MA > 4			
MA100	8	25	8.25	24.75	0.07	1	0.79
CM151	6	43	12.25	36.75	4.62	1	0.03
CM155	12	26	9.5	28.5	0.74	1	0.39
Total of X^2 s					5.43	3	0.14
Pooled X^2	26	94	30	90	0.81	1	0.85
Heterogeneity X^2					4.62	2	0.10

Table 3 Observed and expected segregation ratios and chi-square analysis for populations CM151 and CM155 as dihybrid crosses segregating for both fruit malic and citric acid concentrations (mg/g)

Cross	Observed phenotype (mg/g)				Expected phenotype (mg/g)				X^2	DF	p value
	MA < 2 CA > 2.5	MA > 4 CA < 2	MA < 2 CA < 2	MA > 4 CA > 2.5	MA < 2 CA > 2.5	MA > 4 CA < 2	MA < 2 CA < 2	MA > 4 CA > 2.5			
CM151	3	8	3	35	9	9	3	28	7.03	3	0.07
CM155	12	7	0	19	7	7	2	21	6.61	3	0.09
Total of X^2 s									13.64	6	0.03
Pooled X^2	15	15	3	54	16	16	5	49	2.42	3	0.49
Heterogeneity X^2									11.22	5	0.05

significantly deviate from a 1:1 segregation ratio expected with the *CITA* locus (Table 4) for low CA (< 2.5 mg/g) versus high (> 3 mg/g) (CM92: $\chi^2 = 0.13$, $p = 0.73$, CM93: $\chi^2 = 1.6$, $p = 0.21$).

Genetic mapping

The initial sequencing of the GBS libraries (MA100 and the first half of CM155, $n = 56$ total) with a 2×100-bp format resulted in a total of 43,517 Mbases with a mean quality score of 34.9. The second sequencing run (CM151 and the second half of CM155, $n = 75$ total) with a 2×150-bp format resulted in a total of 108,584 Mbases with a mean quality score of 37.2. After de-multiplexing with STACKS, the libraries yielded an average of 7 million reads per individual. The first sequencing run had 431 million reads and 6.4 million reads per sample. The second sequencing run had 719 million reads and 7.6 million reads per sample. After aligning the reads to the current cranberry reference genome (unpublished), calling SNPs using Samtools, and filtering, a total of 13,698 SNP markers were identified.

The SNP markers for MA (MA_271 and MA_476) were initially identified using a previous reference genome version (Table S3). Marker positions were therefore updated as the reference genome was further refined using Oxford Nanopore (Oxford, UK) long read and Illumina (San Diego, CA, USA) paired-end short-read technologies. These SNPs were converted into KASP assays for high-throughput

analysis of progeny. CM151 and CM155 were also genotyped with SNP CA_609 (also converted into a KASP assay) to follow segregation at the *CITA* locus for CA concentration (Fong et al. 2020) (Table S3).

For MA100, QTLs were identified using scanone in R/qtl and peaks that surpassed the LOD threshold of 9.4 were deemed significant based on 1000 permutations (Fig. S4a). The QTL was located at marker 4_36181956 on chromosome 4 of Schlautman et al. (2017). Further analysis with R/qtl determined the QTL interval to be 267 kb in size from 36,106,859 to 36,371,940 (Fig. S4a). This QTL on chromosome 4 had a LOD score of 11.6 and accounted for 81% of the phenotypic variance in the population. For CM151 and CM155, the data were combined for QTL analysis. Peaks having a LOD threshold of 10.21 or higher were deemed significant. The QTL was located at marker 4_36106859 (Fig. S4b). Further analysis determined the QTL interval to be 437 kb in size from 36,063,067 to 36,499,933. This QTL was also on chromosome 4 with a LOD score of 24.0 and accounted for 72% of the phenotypic variance in the populations. All significant peaks occurred on chromosome 4. Gene prediction of the QTL region yielded a number of uncharacterized proteins along with metal tolerance protein and 7-dehydrocholesterol reductase.

These two QTLs identified by SNP markers were consistent with Mendelian segregation and co-segregated (98%) with the low MA trait (Table 5). For marker 4_36106859 (MA_859) and marker 4_36181956 (MA_956), the TT

Table 4 Observed and expected segregation ratios and chi-square analysis for populations CM92 and CM93 segregating for fruit citric acid concentrations (mg/g)

Cross	Observed phenotype (mg/g)		Expected phenotype (mg/g)		X^2	DF	p value
	CA < 2.5	CA > 3	CA < 2.5	CA > 3			
CM92	60	64	62	62	0.129	1	0.719
CM93	39	51	45	45	1.6	1	0.206
Total of X^2 s					1.729	2	0.42
Pooled X^2	99	115	107	107	1.2	1	0.27
Heterogeneity X^2					0.529	1	0.47

genotype co-segregated with low MA (< 2.5 mg/g FW) and was defined as *mala/mala*. These two markers are 11 kb apart. Additionally, when the MA phenotypes are grouped by the SNP genotypes, there are significant differences between the SNP genotype groups (Fig. 3a,b). Homozygous (*mala/mala*) exhibited MA concentrations below 2.5 mg/g with non-overlapping ranges to that of *Mala/-*. Heterozygotes (*Mala/mala*) had intermediate mean concentrations (3.0–7.8 mg/g) and homozygous (*Mala/Mala*) had the highest mean concentrations (4.8–9.0 mg/g) (Fig. 3a). The phenotypic ranges of *Mala/Mala* and *Mala/mala* genotypes overlapped for MA. The phenotypes indicated that the low MA trait was partially dominant, as the heterozygous (*Mala/mala*) genotype has a MA phenotype of approximately 5 mg/g, close to the midparent value of 5.4 mg/g MA. QA and CA levels were also significantly reduced by *MALA* genotypes having the *mala* allele as defined by the marker (Fig. 3c, d), with *mala/mala* genotype having the lowest levels.

KASP and SSR marker utilization and validation

The relationship of the low CA and low MA traits was evaluated in two additional populations with different genetic backgrounds, CM92 and CM93, to determine the sensitivity and specificity of the identified markers (Table 5). The heterozygous individuals were determined based on expected phenotypic segregation. For CA, the sensitivity and specificity was over 90% for all 3 markers, with scf258d having the best sensitivity and specificity. For MA, while CM92 and CM93

did not segregate for low MA phenotype, the SNP markers were used to identify the efficacy of the markers in identifying false positives. As Table 5 shows, MA_271 had many false positives, resulting in a specificity of 81%. However, MA_476 showed consistency against false positives and had a high sensitivity and specificity for detecting the *mala* alleles successfully (Table 5).

Interaction of *cita* and *mala* alleles

Overall, in populations CM92, CM93, CM151, and CM151, individuals homozygous for *cita/cita* have significantly lower CA concentrations (< 2 mg/g) (Fig. 4a). The presence of one allele of *mala* has a significant effect on reducing overall TA and CA (Fig. 4a, c). However, the presence of two *cita* alleles (*cita/cita*) appears to have an additive effect on increasing MA concentrations (Fig. 4b). Even though the presence of *cita* alleles increases MA concentrations, cumulatively, the *cita* and *mala* alleles seem to have an additive effect on reducing overall TA (Fig. 4c).

Analysis of genotypes at the *CITA* and *MALA* loci and their interaction revealed that *MALA* significantly influenced QA concentrations ($F = 15.2$, $p < 0.0001$). MA concentrations were significantly influenced by both *CITA* ($F = 3.8$, $p = 0.03$) and *MALA* ($F = 88$, $p < 0.0001$), but there was no significant interaction between the two loci. CA concentrations were significantly influenced by both *CITA* ($F = 72.67$, $p < 0.0001$) and *MALA* ($F = 21.41$, $p < 0.0001$), and there was an interaction between the two loci ($F = 2.97$, $p = 0.02$).

Table 5 Phenotypes and marker genotypes for the citric acid trait (SSRs and KASPs) and malic acid trait (KASPs) in CM92, CM93, CM155, and CM151. Sensitivity and specificity were calculated based on the ability to detect *cita/cita* or *mala/mala* individuals. CA: low < 2.5 mg/g, moderate 3–6 mg/g, high > 6; MA: low < 2.5 mg/g, moderate 3.5–5 mg/g, high > 5 mg/g

| Markers	CM151			CM155			CM92			CM93			Sensitivity	Specificity
scf258d	217/217	227/227	217/227	217/217	227/227	217/227	217/217	217/227		217/217	217/227		0.940	0.994
Low CA	9	0	0	7	0	2	55	4		38	1			
Moderate CA	0	2	18	0	2	15	0	53		1	45			
High CA	0	9	3	0	5	5	0	0		0	0			
scf157322							276/276	272/276		276/276	270/276		0.938	0.971
Low CA							54	4		37	2			
Moderate CA							0	54		3	48			
SNP CA_609	G/G	A/G	A/A	G/G	A/G	A/A	G/G	A/G		G/G	A/G		0.912	0.982
Low CA	9	0	0	6	2	0	53	5		36	3			
Moderate CA	0	20	3	0	16	2	2	52		1	49			
High CA	0	3	12	0	5	5	0	0		0	0			
SNP MA_271	T/T	T/C	C/C	T/T	T/C	C/C	T/T	T/C	C/C	T/T	T/C	C/C	0.941	0.811
Low MA	5	1	0	11	0	0	0	0	0	0	0	0		
Moderate MA	1	22	5	0	12	5	21	28	8	14	15	16		
High MA	0	6	9	0	5	3	9	26	22	6	6	31		
SNP MA_476	A/A	A/C	C/C	A/A	A/C	C/C		A/C	C/C		A/C	C/C	0.941	1.000
Low MA	5	1	0	11	0	0		0	0		0	0			
Moderate MA	0	23	5	0	12	5		43	14		13	32			
High MA	0	6	9	0	5	3		18	39		29	16			
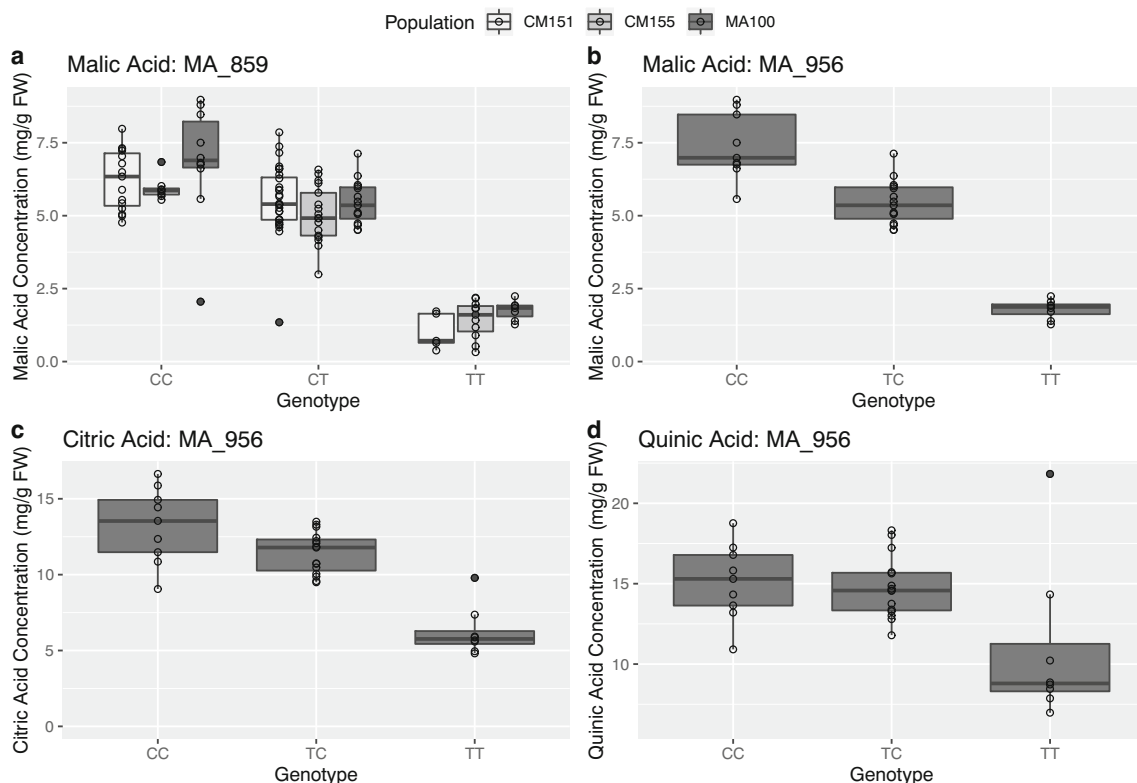
Fig. 3 Distribution of organic acid phenotypes (means from 2016 and 2017) in CM151, CM155, and MA100 along with SNP genotypes. The TT genotypes in both SNPs MA_859 and MA_956 are associated with low malic acid (< 2 mg/g). **a** Malic acid phenotypes in CM151, CM155, and MA100 for SNP MA_859. **b** Malic acid phenotypes in MA100 for SNP MA_956. **c** Citric acid phenotypes in MA100 for SNP MA_956. **d** Quinic acid phenotypes in MA100 for SNP MA_956

Overall, TA was significantly influenced by both *CITA* (F = 5.9, p = 0.004) and *MALA* (F = 112.25, p < 0.0001), with no significant interaction between the two loci. There was no significant effect of the heterozygous *Mala/mala* versus homozygous *Mala/Mala* genotype on CA concentrations (Fig. 5b). However, overall, there was a significant effect of one copy of the *cita* allele (*cita/-*) on CA concentration, consistent with co-dominance. The homozygous *mala* genotype (*mala/mala*) significantly decreased the CA concentration. At the *MALA* locus, there was a significant effect of the *mala* allele in the heterozygous genotype on decreasing MA concentrations, except when the *CITA* locus was homozygous (*cita/cita*) (Fig. 5a). In fact, *cita* appears epistatic to *mala*, where the presence of at least one *cita* allele, and enhanced when homozygous, increases the MA concentration when *mala* was heterozygous. Additionally, *mala* had an additive effect in the presence of *cita* in decreasing CA concentrations. In the presence of *mala/mala*, there was a decrease in CA concentrations. Overall, both *mala* and *cita* seem to have an additive effect in reducing TA (Fig. 5d).

Utilizing all the phenotype and genotype data, a multiple regression analysis was performed to determine the total proportion of observed TA variation that could be accounted for by these three acids. For CM151 and CM155 combined, the variables most significant for TA were CA, MA, QA, and genotype, with a R^2 = 0.92. For CM92, the variables most significant were MA, QA, and genotype, with a R^2 = 0.65. For CM93, the variables most significant were CA, MA, and genotype, with a R^2 of 0.72. When all the data were pooled, CA, MA, and genotype were most significant, with a R^2 = 0.86.

Discussion

We have characterized a low MA trait in cranberry fruit derived from a native germplasm accession (NJ93-57) and determined its interaction with genotypes at the *CITA* locus. NJ93-57 was determined to be heterozygous (*Mala/mala*), with a MA concentration of ~ 4 mg/g. In an F_2 population, derived from an F_1 Mullica Queen × NJ93-57, progeny with a MA concentration as low as ~ 2 mg/g were recovered. This is the lowest level MA phenotype reported in cranberry and is a result of a *mala/mala* homozygous genotype, with the *mala* being largely recessive, and the wild-type allele partially dominant. The high correlation of MA levels across years indicates a strong qualitative genetic effect of the *mala/mala* genotype. However, significant differences in MA and TA between populations with the low the MA trait, e.g., CM151 versus CM155, indicate general genetic background also influences acidity to some extent. To characterize and map the low MA
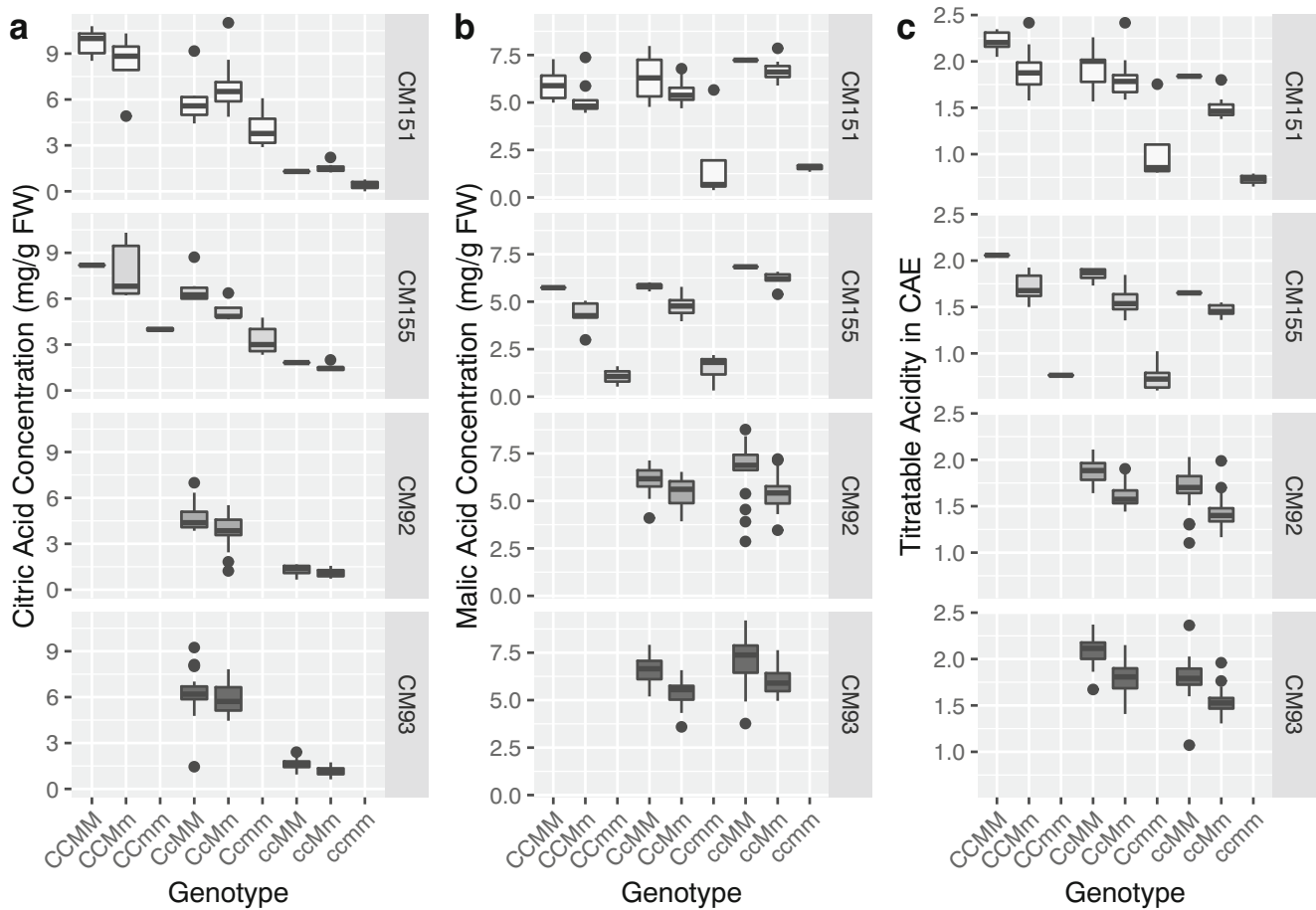
Fig. 4 Distribution of phenotypes (means from 2016 and 2017): citric acid (**a**), malic acid (**b**), and TA (**c**) based on the consensus genotype from genotyping with SSR and KASP markers. CC is *Cita/Cita*, Cc is *Cita/cita*, cc is *cita/cita*. MM is *Mala/Mala*, Mm is *Mala/mala*, mm is *mala/mala*

trait, *mala*, three populations with a total of 119 unique individuals were phenotyped and genotyped. The three populations segregated for low MA trait (~ 2 mg/g FW) consistent with a single, co-dominant gene in a Mendelian pattern and we named the locus *MALA*. The initial KASP markers (MA_271 and MA_476) developed for genotyping the *MALA* locus were identified from a SNP map with MA_271 being placed on a smaller fragment that was not initially contiguous with the MA_476. Subsequently, with an improved reference map, MA_271 was placed adjacent to MA_476. Two additional SNP markers (MA_859 and MA_956), 11 kb apart and within 1 cM, co-segregated with *MALA* locus, increasing the resolution of this region and aiding in possible gene identification.

To assess the utility of these SNP markers and determine the effect of the *mala* allele in the heterozygous genotype (*Mala/Mala* vs. *Mala/mala*), in the three *CITA* genotypes (*Cita/Cita*, *Cita/cita*, and *cita/cita*), two populations (CM92 and CM93) with a total of 214 individuals were genotyped with markers associated with *mala* and *cita* alleles (Fong et al. 2020). As the loci for *mala* and

cita are on separate chromosomes (chromosomes 4 and 1, respectively), the observed segregation pattern revealed, as expected, that *cita* and *mala* indeed segregated independently. For *mala*, the KASP marker MA_476 had the best sensitivity and specificity for successfully determining lower MA individuals and importantly discriminating heterozygotes with the *mala* allele from homozygotes lacking the allele. For *cita*, the scf258d SSR marker was considered best, as found in Fong et al. (2020).

There were significant differences in population means for citric, malic, and quinic acids and TA for all five populations analyzed in the present study, likely reflecting heterogeneity in the segregation of genotypes at *MALA* and *CITA* loci between various populations. Lacking the low CA allele, *cita*, at the *CITA* locus, population MA100 had a higher average CA concentration and a higher average TA, in contrast to the other four populations, which is likely due to the strong effect of the *cita* allele over the *mala* allele in reducing CA, and thus TA. Populations segregating for only two *CITA* genotypes, *cita/cita* and *Cita/cita*, and lacking the homozygous wild-type genotype, *Cita/Cita* (CM92 and CM93) exhibited lower CA
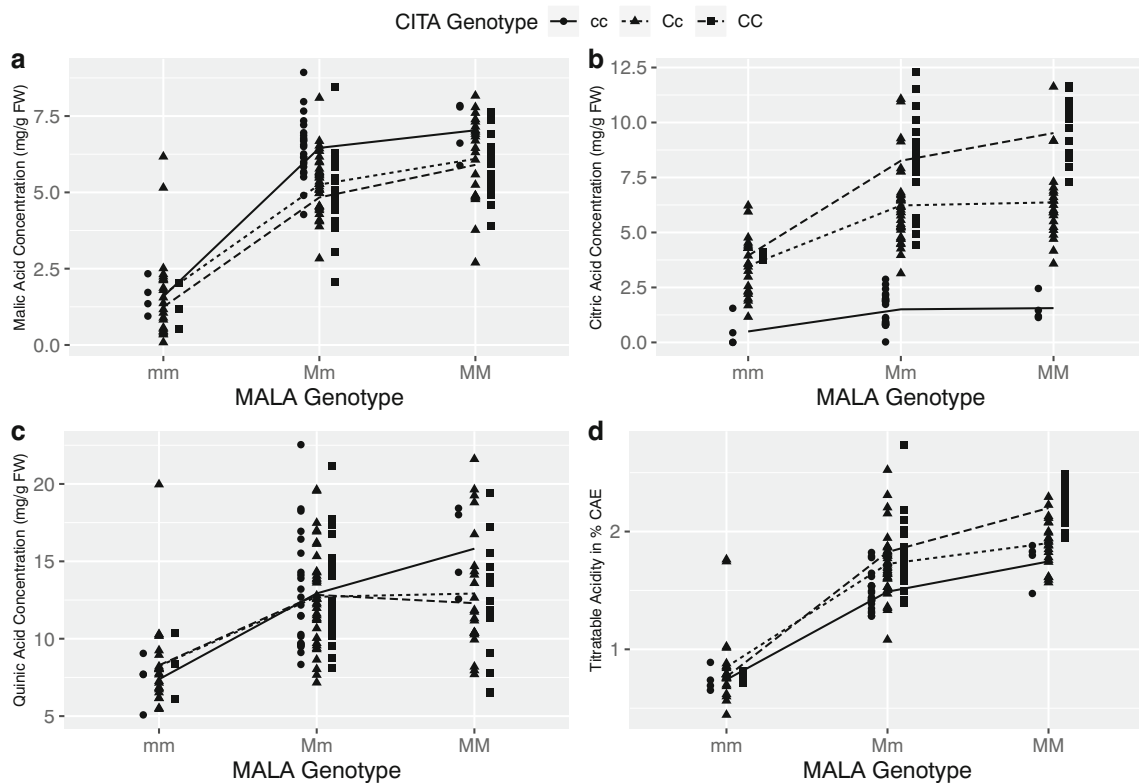
Fig. 5 The effect of *mala* and *cita* on organic acid concentration (mg/g fresh weight) and titratable acidity (%citric acid equivalents) of CM151 and CM155 populations (mean of 2016 and 2017): malic acid (**a**), citric acid (**b**), quinic acid (**c**), and titratable acidity (**d**). CC is *Cita/Cita*, Cc is *Cita/cita*, cc is *cita/cita*. MM is *Mala/Mala*, Mm is *Mala/mala*, mm is *mala/mala*

concentrations than those homozygous for wild-type alleles at both *CITA* and *MALA*, e.g., CM151 and CM155, which segregated for all three genotypes at both *CITA* and *MALA*.

QA variation did not appear to be related to the genetic composition of the population, but QA concentrations were correlated with TA across populations, which is consistent with the observation that the *mala* allele effect depresses all three acids. However, QA levels had significant year-to-year variation in all populations, indicating that there is likely an environmental effect on QA concentrations. Wang et al. (2017) found, as cranberry fruit matured, QA concentrations decreased. Thus, the fruit harvested at variable fruit maturation stages may have contributed to variation in QA concentrations, as has also been found in kiwi fruit species (Marsh et al. 2009). Cranberry fruit ripening appears continuous, not having a definitive demarcation point as is found in other fruit species, e.g., blueberry. It is well known that genetics has an influence on fruit development and ripening in cranberry.

As in Fong et al. (2020), a strong positive correlation between MA and CA with TA was found, as a result of MA and CA being the primary contributors to TA (Lobit et al. 2002). Even though QA is a weaker acid, the wild-type allele(s) at the *MALA* locus yield higher QA levels resulting in a positive correlation between QA and TA in populations segregating

for the alleles *Mala* and *mala*. This is also possibly due to the level of QA being proportionally higher relative to CA and MA in certain populations, i.e., MA100, CM151, and CM155, and thus contributed more to TA. A similar effect was found for the correlations between MA and QA.

The CM155 population was missing individuals with the *cita/cita mala/mala* genotype. However, when tested for dihybrid 9:3:3:1 segregation of both CA and MA, and 3:1 for low MA, CM155 segregation did not deviate significantly from expected. The CM151 population was missing progeny having the *Cita/Cita mala/mala* genotype, but segregation did not significantly deviate from the expected dihybrid ratio of 9:3:3:1 for these two traits, indicating *MALA* and *CITA* lack allelic interactions that would strongly affect fitness at these loci. However, CM151 deviated significantly from expected 3:1 segregation for the low MA trait, with fewer than expected *mala/mala* genotype. The low MA phenotype is associated with a "dwarf-like" vegetative morphology, as well as delayed germination (unpublished data) (see Fig. S3ab). The dwarf-like morphology, and *mala/mala* genotype, is disproportionally more frequent in the latest germinating seedlings of a population and may result in higher lethality of the *mala/mala* genotype, resulting in the segregation ratios being skewed in certain backgrounds. To date, no recombination has been observed

between the dwarf-like phenotype and the low malic (*mala/mala*) acid trait indicating they are either pleiotropic or very tightly linked.

The *MALA* locus was localized to a 143 kb region interval of chromosome 4 (S5). A BLAST search of this region revealed several candidate genes, most interestingly a cation/calcium exchange-like protein coupled with an aldehyde dehydrogenase family 7, EXO70 superfamily, sm-like protein, and a phosphoinositide phosphatase SAC2 isoform. Variation in EXO70 superfamily, the sm-like protein, and the phosphoinositide phosphatase SAC2 isoform have been suggested as causes of reduced growth and smaller plant sizes (Xiong et al. 2001; Hála et al. 2008; Nováková et al. 2014). This would suggest that the low MA trait is pleotropic with the dwarf-like phenotype. Additionally, malate synthase was identified 4.6 Mbp upstream of the QTL region, which, when knocked out in *Arabidopsis thaliana*, resulted in reduced and slower plant growth. Malate synthase is also related to the glycosylate pathway which is associated with slower seed germination (Cornah et al. 2004). As mentioned earlier, retardation in germination has been observed in *mala/mala* individuals (Vorsa, unpublished data).

The BLAST search for genes identified with fruit acidity by others (Etienne et al. 2002a, b, Cohen et al. 2012, Bai et al. 2012) to the cranberry reference genome yielded several genes (ketoglutarate dehydrogenase, citrate lyase, ATP-citrate lyase, NAD-malate dehydrogenase, succinyl-CoA ligase, vacuolar ATP dependent H+ transporter, and malate synthase) on chromosome 4 where *MALA* is located. However, there were none in the vicinity of the *MALA* QTL. The gene found closest to the QTL was MU44768 for malate synthase (Cohen et al. 2012) which was 4.6 Mbp upstream of the *MALA* QTL. The lack of predicted genes in the QTL region as well as lack of co-localization with candidate genes from various studies indicates that there could be another mechanism not previously discovered controlling MA concentrations in cranberry fruit.

A low CA trait in cranberry has been characterized (Fong et al. 2020), complementing this study which focuses on characterizing the phenotypic variation and genetics of the low MA trait in cranberry. Evaluation of MA throughout the cranberry fruit development period revealed that MA increases early after fruit set and levels off during fruit ripening (Wang et al. 2017). In addition, QTLs were identified for MA concentration on chromosomes 1, 6, and 12 in various breeding populations segregating for fruit rot resistance (Fong 2019) suggesting quantitative variation could be exploited to a degree. In populations with the low CA trait (*cita/cita*), there was an inverse relationship between citric and malic acids; CA concentration was negatively correlated with MA concentration attenuating the effect on TA (Fong et al. 2020).

The *mala* allele had a significant effect on QA and CA as well as TA. In comparison to *cita*, homozygous *mala* decreases TA more than *cita* (Fong et al. 2020), since *mala* also

decreased CA and to some extent QA concentration as well. Dihybrid crosses with *MALA* and *CITA*, with segregation for all possible combinations of *cita* and *mala*, allowed the analysis of the interaction of these two traits. An epistatic effect was observed in the presence of at least one copy of the *cita* allele, which appears to increase MA concentration, thereby damping the effect in reducing TA. However, overall, when the interactions for TA are analyzed, there was a significant effect of the *cita* allele in decreasing overall TA. Although *cita* may cause an increase in MA, it greatly decreases CA, thereby decreasing TA. The *mala* allele in the *Mala/mala cita/cita* genotype does not appear to dampen the negative correlation of CA and MA significantly. Although the *mala/mala* genotype yields TA < 1%, there is a caveat, as the homozygous *mala* plants have a dwarf-like growth habit which is likely not commercially viable (Fig. S3).

To generate potentially commercial viable individuals, two populations were generated with the cross *Cita/cita Mala/mala* x *cita/cita Mala/Mala* and thus do not segregate for the dwarfed phenotype. This resulted in segregation of 1:1 for low CA (*cita/cita*) to heterozygous CA (*Cita/cita*) and heterozygous MA (*Mala/mala*) and normal MA (*Mala/Mala*). Overall, *cita/cita Mala/mala* individuals have an average TA ≈ 1.50. The presence of *mala* and *cita* alleles had an additive effect on reducing TA. And in contrast to the populations in Fong et al. (2020), there was no negative correlation between citric and malic acid in the presence of *cita*. Other than the dwarf-like vegetative phenotype associated with the homozygous *mala/mala* genotype, there was no apparent effect on fruit size nor fruit morphology.

Conclusion

Cranberry fruit has been well studied for its various health benefits due to the high level of human health-promoting compounds such as flavonols (Wang et al. 2017). However, their "superfruit" status is diminished due to the added-sugars necessary to balance cranberry's high acid content (Bates et al. 2001). Here, a low MA trait was characterized and a genetic locus that confers a decreased TA to below 1% was identified. TA less than 1 is in the range of fruit that is consumed fresh, such as strawberries (Kallio et al. 2000). Other than variation in acidity, there were no other apparent fruit quality traits affected. The SNP markers (MA_859 and MA_956) identified have been developed into KASP assays to begin screening different populations, as well as the germplasm collection for low MA. Further sequencing of this region to develop a haplotype marker set would be advantageous to use where the two SNPs identified are not sufficient. Cranberry is a woody perennial, producing fruit 2–3 years after germination of seed (Vorsa and Johnson-Cicalese 2012). Due to this long generation time, marker-assisted selection (MAS) is useful to

decrease selection time and allow seedlings of no value to be culled, saving space, and money. Identifying the genes controlling citric and malic acid accumulation would allow greater ability for MAS and gene editing applications in the future.

References

Bae H, Yun SK, Jun JH et al (2014) Assessment of organic acid and sugar composition in apricot, plumcot, plum, and peach during fruit development. J Appl Bot Food Qual 87:24–29. https://doi.org/10.5073/JABFQ.2014.087.004

Bai Y, Dougherty L, Li M et al (2012) A natural mutation-led truncation in one of the two aluminum-activated malate transporter-like genes at the Ma locus is associated with low fruit acidity in apple. Mol Gen Genomics 287:663–678. https://doi.org/10.1007/s00438-012-0707-7

Bates RP, Morris JR, Crandall PG (2001) Principles and practices of small- and medium- scale fruit juice processing. Food and Agriculture Organization, Rome, p 146

Boudehri K, Bendahmane A, Cardinet G et al (2009) Phenotypic and fine genetic characterization of the D locus controlling fruit acidity in peach. BMC Plant Biol 9:59. https://doi.org/10.1186/1471-2229-9-59

Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890. https://doi.org/10.1093/bioinformatics/btg112

Cantarel BL, Korf I, Robb SMC et al (2008) MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res 18:188–196. https://doi.org/10.1101/gr.6743907

Cao K, Zhou Z, Wang Q et al (2016) Genome-wide association study of 12 agronomic traits in peach. Nat Commun 7:13246. https://doi.org/10.1038/ncomms13246

Catchen JM, Amores A, Hohenlohe P et al (2011) Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes. Genomes, Genetics 1:171–182. https://doi.org/10.1534/g3.111.000240

Cohen S, Tzuri G, Harel-Beja R et al (2012) Co-mapping studies of QTLs for fruit acidity and candidate genes of organic acid metabolism and proton transport in sweet melon (Cucumis melo L.). Theor Appl Genet 125:343–353. https://doi.org/10.1007/s00122-012-1837-3

Cornah JE, Germain V, Ward JL et al (2004) Lipid utilization, gluconeogenesis, and seedling growth in Arabidopsis mutants lacking the glycosylate cycle enzyme malate synthase. J Biol Chem 279:42916–42923. https://doi.org/10.1074/jbc.M407380200

Covarrubias-Pazaran G, Diaz-Garcia L, Schlautman B et al (2016) Exploiting genotyping by sequencing to characterize the genomic structure of the American cranberry through high-density linkage

mapping. BMC Genomics 17:451. https://doi.org/10.1186/s12864-016-2802-3

Cunningham DG, Vannozzi SA, Turk R, Roderick R, O'Shea E, Brilliant K (2003) Cranberry phytochemicals and their health benefits. In: Nutraceutical Beverages, ACS Symposium Series. American Chemical Society, Washington DC, pp 35–51. https://doi.org/10.1021/bk-2004-0871.ch004

Daverdin G, Johnson-Cicalese J, Zalapa J et al (2017) Identification and mapping of fruit rot resistance QTL in American cranberry using GBS. Mol Breed 37:38. https://doi.org/10.1007/s11032-017-0639-3

Diaz-Garcia L, Schlautman B, Covarrubias-Pazaran G et al (2018) Massive phenotyping of multiple cranberry populations reveals novel QTLs for fruit anthocyanin content and other important chemical traits. Mol Gen Genomics 293:1379–1392. https://doi.org/10.1007/s00438-018-1464-z

Eduardo I, López-Girona E, Batlie I et al (2014) Development of diagnostic markers for selection of the subacid trait in peach. Tree Genet Genomes 10:1695–1709. https://doi.org/10.1007/s11295-014-0789-y

Etienne C, Moing A, Dirlewanger E et al (2002a) 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 114:259–270. https://doi.org/10.1034/j.1399-3054.2002.1140212.x

Etienne C, Rothan C, Moing A et al (2002b) Candidate genes and QTLs for sugar and organic acid content in peach [Prunus persica (L.) Batsch]. Theor Appl Genet 105:145–159. https://doi.org/10.1007/s00122-001-0841-9

Fong SK (2019) The genetics of organic acids variation in cranberry fruit. PhD Dissertation, Rutgers University

Fong SK, Kawash J, Wang Y et al (2020) A low citric acid trait in cranberry: genetics and molecular mapping of a locus impacting fruit acidity. Tree Genet Genomes 16:42. https://doi.org/10.1007/s11295-020-01432-4

Forney CF, Kalt W, Jordan MA et al (2012) Blueberry and cranberry fruit composition during development. J Berry Res 2:169–177. https://doi.org/10.3233/JBR-2012-034

Georgi L, Johnson-Cicalese J, Honig J et al (2013) The first genetic map of the American cranberry: exploration of synteny conservation and quantitative trait loci. Theor Appl Genet 126:673–692. https://doi.org/10.1007/s00122-012-2010-8

Hála M, Cole R, Synek L et al (2008) An exocyst complex functions in plant cell growth in Arabidopsis and tobacco. Plant Cell 20:1330–1345. https://doi.org/10.1105/tpc.108.059105

Kader AA (2008) Flavor quality of fruits and vegetables. J Sci Food Agric 88:1863–1868. https://doi.org/10.1002/jsfa.3293

Kallio H, Hakala M, Pelkkikangas A-M, Lapveteläinen A (2000) Sugars and acids of strawberry varieties. Eur Food Res Technol 212:81–85. https://doi.org/10.1007/s002170000244

Lambert P, Campoy JA, Pacheco I et al (2016) Identifying SNP markers tightly associated with six major genes in peach [Prunus persica (L.) Batsch] using a high-density SNP array with an objective of marker-assisted selection (MAS). Tree Genet Genomes 12:121. https://doi.org/10.1007/s11295-016-1080-1

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lobit P, Soing P, Génard M, Habib R (2002) Theoretical analysis of relationships between composition, pH, and titratable acidity of peach fruit. J Plant Nutr 25:2775–2792. https://doi.org/10.1081/PLN-120015538

Marsh KB, Boldingh HL, Shilton RS, Laing WA (2009) Changes in quinic acid metabolism during fruit development in three kiwifruit species. Funct Plant Biol 36:463–470. https://doi.org/10.1071/FP08240

McCamey DA, Thrope TM, McCarthy JP (1990) Coffee bitterness. In: Rouseff RL (ed) Bitterness in foods and beverages. Elsevier Science Publishers B.V, Amsterdam, The Netherlands, pp 169–182

Moing A, Svanella L, Rolin D et al (1998) Compositional changes during the fruit development of two peach cultivars differing in juice acidity. J Am Soc Hortic Sci 123:770–775. https://doi.org/10.21273/JASHS.123.5.770

Nishiyama I, Fukuda T, Shimohashi A, Oota T (2008) Sugar and organic acid composition in the fruit juice of different *Actinidia* varieties. Food Sci Technol Res 14:67–73. https://doi.org/10.3136/fstr.14.67

Nováková P, Hirsch S, Feraru E et al (2014) SAC phosphoinositide phosphatases at the tonoplast mediate vacuolar function in Arabidopsis. Proc Natl Acad Sci USA 111:2818–2823. https://doi.org/10.1073/pnas.1324264111

Original Dried Cranberries (2019) Ocean Spray®. https://www.oceanspray.com:443/sitecore/content/OS/Brand Sites/OceanSpray/Master/Intl Master/Data/Products/Craisins Dried Cranberries/Craisins Original Dried Cranberries. Accessed 10 Mar 2019

Polashock J, Zelzion E, Fajardo D et al (2014) The American cranberry: first insights into the whole genome of a species adapted to bog habitat. BMC Plant Biol 14:165. https://doi.org/10.1186/1471-2229-14-165

Rosas JE, Bonnecarrère V, Pérez de Vida F (2014) One-step, codominant detection of imidazolinone resistance mutations in weedy rice (*Oryza sativa* L.). Electron J Biotechnol 17:95–101. https://doi.org/10.1016/j.ejbt.2014.02.003

Rubico SM, McDaniel MR (1992) Sensory evaluation of acids by free-choice profiling. Chem Senses 17:273–289. https://doi.org/10.1093/chemse/17.3.273

Schlautman B, Covarrubias-Pazaran G, Diaz-Garcia LA et al (2015) Development of a high-density cranberry SSR linkage map for comparative genetic analysis and trait detection. Mol Breed 35:177. https://doi.org/10.1007/s11032-015-0367-5

Schlautman B, Covarrubias-Pazaran G, Diaz-Garcia L et al (2017) Construction of a high-density American cranberry (*Vaccinium macrocarpon* ait.) composite map using genotyping-by-sequencing for multi-pedigree linkage mapping. G3 (Bethesda) 7:1177–1189. https://doi.org/10.1534/g3.116.037556

Vorsa N, Johnson-Cicalese J (2012) American cranberry. In: Badenes ML, Byrne DH (eds) Fruit Breeding. Springer US, Boston, MA, pp 191–223

Wang Y, Johnson-Cicalese J, Singh AP, Vorsa N (2017) Characterization and quantification of flavonoids and organic acids over fruit development in American cranberry (*Vaccinium macrocarpon*) cultivars using HPLC and APCI-MS/MS. Plant Sci 262:91–102. https://doi.org/10.1016/j.plantsci.2017.06.004

Wang Y, Fong SK, Singh AP et al (2019) Variation of anthocyanins, proanthocyanidins, flavonols, and organic acids in cultivated and wild diploid blueberry species. HortScience 54:576–585. https://doi.org/10.21273/HORTSCI13491-18

Xiong L, Gong Z, Rock CD et al (2001) Modulation of abscisic acid signal transduction and biosynthesis by an sm-like protein in Arabidopsis. Dev Cell 1:771–781. https://doi.org/10.1016/S1534-5807(01)00087-9

Xu K, Wang A, Brown S (2012) Genetic characterization of the Ma locus with pH and titratable acidity in apple. Mol Breed 30:899–912. https://doi.org/10.1007/s11032-011-9674-7

Ye J, Wang X, Hu T et al (2017) An InDel in the promoter of *Al-ACTIVATED MALATE TRANSPORTER9* selected during tomato domestication determines fruit malate contents and aluminum tolerance. Plant Cell 29:2249–2268. https://doi.org/10.1105/tpc.17.00211

Zeballos JL, Abidi W, Giménez R et al (2016) Mapping QTLs associated with fruit quality traits in peach [*Prunus persica* (L.) Batsch] using SNP maps. Tree Genet Genomes 12:37. https://doi.org/10.1007/s11295-016-0996-9

Zhang Y, Li P, Cheng L (2010) Developmental changes of carbohydrates, organic acids, amino acids, and phenolic compounds in 'Honeycrisp' apple flesh. Food Chem 123:1013–1018. https://doi.org/10.1016/j.foodchem.2010.05.053