Waders to Wings: Evolving Phenotyping from Field to Flight in the American Cranberry (*Vaccinium macrocarpon* Ait.) to Accelerate Breeding

By

Andrew F. Maule

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Plant Breeding and Plant Genetics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2025

Date of final oral examination: 8/26/2025

The dissertation is approved by the following members of the Final Oral Committee:
Juan E. Zalapa, Professor, Plant and Agroecosystem Sciences
Matthew F. Digman, Associate Professor, Biological Systems Engineering
Natalia De Leon, Professor, Plant and Agroecosystem Sciences
Yiqun Weng, Professor, Plant and Agroecosystem Sciences

Dedication

To my parents –

My bedrock, my foundation, my love.

To Marie –

Mother-in-law, troll of the E. 3rd Street tolls, best friend.

To JLo –

Friend of my heart, partner in crimes, scourge of ALB in *Daucus*.

To Halley –

Friend of my soul, fellow terrier lover, menagerie extraordinaire.

To Dr. Ed Woolsey --

Local historian, zoologist turned horticulturalist, critic, and curmudgeon. Who else could teach proper potato-digging technique while acquainting students with Red Dot Potato Chips and a local tragedy?

Acknowledgements

To Dr. Jenyne Loarca –

Without your continuous, genuine, and sometimes stubborn encouragement, I would not have finished. You provided me with over 300,000 apt and creative metaphors for 'light at the end of the tunnel', even when it seemed that 'light' couldn't be found in the Hubble Telescope's Deep Field. You persisted in asking about goals and progress, even when I would respond combatively. You taught me valuable tools to help me advance. You were pivotal to helping me publish my first research chapter, digging me out of that self-made landslide of rubble. You reviewed my dissertation and helped me 'zoom out' in those critical moments where I was myopically 'zoomed in.' I look forward to our continued friendship and our enduring pursuit of knowledge and discovery in our careers, and to many productive collaborations.

To my dissertation committee –

First, I thank my advisor, Dr. Juan Zalapa. You always advocated for me, even when no one else did. You reminded me of my contribution to the lab when I felt my work was fruitless. You continued to support me financially when most would have cut me loose. You were a mountain of pithy expressions, and your big heart kept me afloat.

To Dr. Matthew Digman, your two-year computer loan provided the essential infrastructure that made completing this dissertation possible. Your sage life insights and constructive technical critiques have shaped me into a more thoughtful researcher and writer. Your food philosophy has inspired me toward a diet that is morally sound, healthy, and environmentally responsible.

To Dr. Natalia De Leon, you exemplify excellence in teaching across multiple dimensions. I regret not seeking your counsel more frequently, as your humble yet respectful approach to statistics and science reflects your depth of knowledge and wisdom. Your rigorous but measured questions and critiques, balanced with ample grace, motivates me to keep learning.

Finally, I thank Dr. Yiqun Weng for your valuable scientific critiques that have consistently strengthened my research. Your insights have enhanced both the rigor and clarity of this work.

To all those who helped me, in particular –

Matthew Phillips – Your field help was pivotal in my first year of data collection. You taught me Chicago is the objective epicenter of the Universe. I hope that wherever you are, you are flourishing and happy.

Dylan Schaap – You helped me with mountains of fieldwork, provided meticulous and diligent help with laboratory work, and entertained me in my 'DIY drone' pursuits.

Eric Wiesman – You helped keep the lab and my experiments afloat financially. Thanks for putting up with all my ludicrous experimental conceptions while having patience for the occasional lost vehicle fuel receipt.

Abstract

American cranberry (*Vaccinium macrocarpon* Ait.) faces significant breeding challenges due to limited domestication, extended lifecycle requirements, and vulnerability to late spring frosts during dormancy emergence. This research advances cranberry breeding through the integration of high-throughput phenotyping technologies with comprehensive genetic mapping approaches to address critical agronomic traits affecting yield stability and fruit quality.

A comprehensive meta-quantitative trait loci (QTL) analysis across three cranberry populations (CNJ02, CNJ04, GRYG) examined 29 primary traits encompassing fruit quality, yield, and chemical composition over multiple years. This analysis identified 170 major QTL in CNJ02 and 69 in CNJ04, with 22 stable meta-QTL validated across populations and studies. Importantly, this work demonstrated that modern high-throughput phenotyping methods achieve comparable accuracy to traditional breeding approaches while enabling greater efficiency and scalability.

Building upon these phenotyping advances, a novel assessment system using uncrewed aerial vehicles equipped with RGB cameras was developed to evaluate spring vegetative development patterns among cranberry genotypes. By quantifying leaf color changes associated with dormancy exit timing, a leaf maturity index (LMI) was created that integrates temporal anthocyanin pigmentation patterns. Advanced regression modeling using simulated annealing achieved high predictive accuracy ($R^2 = 0.88$) for anthocyanin levels, enabling rapid screening for frost tolerance characteristics.

The genetic architecture underlying spring developmental timing was subsequently mapped using the LMI trait, representing the first such analysis in cranberry. QTL mapping with high-density linkage maps containing 6,073 markers revealed 11 significant QTL in population CNJ02, where LMI exhibited high heritability ($h^2 = 0.74$). The quantitative genetic architecture suggests complex regulation by multiple genes of modest effect. Candidate gene analysis identified homologs of key dormancy regulators, including circadian clock components, photoperiod sensors, and flowering integrators, providing molecular targets for breeding frost-resilient cultivars.

Collectively, this research establishes a comprehensive framework integrating advanced phenotyping technologies with robust genetic analysis to accelerate cranberry breeding for improved yield stability and quality traits.

Table of Contents

Dedication ii

Acknowledgements iii

Abstract v

Table of Contents vii

List of Figures xii

List of Tables xix

List of Equations xxi

Chapter I Literature Review: Advancing Cranberry Breeding by Integrating Genetics with Aerial Imaging 1

- History and Economics 1
- Morphology and Trait Granularity 3
- Breeding Challenges and the Path to Improvement 3
- Genomic Advances in Cranberry Research 4
- Quantitative Trait Loci: Elucidating the Genetic Architecture of Cranberry Traits 6
- The Evolution of Phenotyping Approaches in Cranberry Evaluation 7
- Technological Integration in Modern Crop Phenomics 9
- Monitoring Spring Development to Address Frost Vulnerability 11
- Genomic Selection Applications for Cranberry Improvement 12
- Research Gaps and Future Directions 13
- Thesis Objectives 17
- References 19

Chapter II Of buds and bits: a meta-QTL study identifies stable QTL for berry quality and yield traits in cranberry mapping populations (*Vaccinium macrocarpon* Ait.) 23

Introduction 25

Materials and methods 28

Plant material and traits collected 28

Trait evaluation and transformation 31

Linkage maps 35

Estimating breeding values and heritability 36

QTL mapping 38

Meta-QTL 39

Results 41

Correlations 41

Heritabilities 42

Breeding value estimates 42

QTL analysis 45

Discussion 57

Trait correlations 58

Heritabilities 60

QTL summary 63

Conclusion 67

Data availability statement 67

Author contributions 67

Funding 68

Acknowledgments 68

Conflict of interest 68

Publisher's note 69

Supplementary Material 69

References 69

Chapter III Modeling Spring Greening Patterns among Cranberry Genotypes via Aerial Imaging of Leaf Color 73

Abstract 73

Introduction 74

Materials and Methods 78

Populations 78

Leaf Sampling and Processing 79

UAV Imaging 80

Vegetative Indices 81

Growing Degree Days 83

Correlations 84

Model Fit and Prediction 84

Selection Index 87

Results 88

Correlations 88

Model Fit 89

Developmental Progression 94

Genotype Performance 95

Discussion 99

Trait Patterns 100

Model Characteristics 102

Genotype Performance 105

Conclusion 106

References 109

Chapter IV Mapping the Genetic Basis of Temporal Segregation of Spring Greening in Cranberry 116

Abstract 116

Introduction 116

Materials and Methods 118

Trait 118

Populations 119

Linkage Maps 120

Breeding Values and Heritability 120

QTL Mapping 121

Results 122

Discussion 128

Model Fit and Heritabilities 128

QTL Patterns 129

Candidate Genes 130

Conclusion and Future Directions 133

References 135

Chapter V Conclusions & Recommendations for Future Research 140

From Upright Trait Assessment to High-Throughput Phenomics: Validation and New Applications in Breeding 140

Limitations and Areas for Improvement 140

Near-Term Research Priorities 143

Genetic Architecture Analysis and Validation 143

Multi-Trait Integration and Analysis 144

Longer-Term Research Objectives 144

Advanced Phenotyping and Data Integration 144

Population Genetics and Genomic Selection Implementation 145

Genomic Prediction Model Development 146

Community Resource Development and Integration 146

References 147

Appendix A Supplementary Data for Chapter III 150

Appendix B Supplementary Data for Chapter IV. 153

List of Figures

Figure II-1. Examples of categories of berry shape parameters, used to help classify cranberry shape (Franklin *et al.*, 1958). 31

Figure II-2. Example showing the methodology for generating representative genotype shape, or berry chimera, from ten-upright samples. This berry chimera is subsequently used as an image to generate digital image processing shape descriptors used in creating quantifying features for QTL mapping. 34

Figure II-3. Trait phenotype correlation heatmaps for traits in cranberry populations *CNJ02* (A and B) and *CNJ04* (C and D). Only traits found within non-singleton hierarchical clusters are shown in (A) and (C), where black squares delineate clusters. The 15 largest positive and 15 smallest negative correlations are shown in (B) and (D). Only significant pairwise trait correlations ($p < 0.05$) are displayed. 44

Figure II-4. Co-located, stable QTL in cranberry populations *CNJ02* (A) and *CNJ04* (B) found in at least three of four BLUP models. Only QTL with $\overline{R^2} \geq 10\%$ are shown. Shaded bars represent the 1.5LOD confidence interval for the original QTL, with the darker line segments indicating the location of the QTL. The linkage groups are divided into four columns, one for each fitted BLUP model — from left to right: 2011, 2012, 2013, and *all-years* (A); and 2011, 2012, 2014, and *all-years* (B). Each QTL interval is color-coded by its corresponding trait. The labels and interval lines to the right of linkage groups show the stable metaQTL associated with their respective genomic regions. 49

Figure II-5. Noteworthy multi-trait, co-located QTL for cranberry populations CNJ02 (A), CNJ04 (B), and both CNJ02 and CNJ04 (C), based on results from this study. 51

Figure II-6. Cross study co-located meta-QTL found comparing QTL from this study with QTL from Diaz-Garcia et al. (2018b) (A), Diaz-Garcia et al. (2018a) (B), Schlautman et al. (2015) (C), and a combined synthesis of Diaz-Garcia et al. (2018a) and Schlautman et al. (2015) (D) in cranberry populations CNJ02, CNJ04, and GRYG. 54

Figure II-7. Mean heritability between cranberry populations CNJ02 and CNJ04 of all-year model (A) and across within-year models (B). Only the top 50% of traits based on mean within-year model heritability are displayed. 56

Figure II-8. Heritabilities and Spearman's rank correlation coefficients (ρ) for cranberry traits in populations (A) *CNJ02* and (B) *CNJ04*. Correlation coefficients represent the associations between genotype BLUPs and their associated phenotypes for the all-years mixed model. Cross study co-located metaQTL found comparing QTL from this study with QTL from Diaz-Garcia *et al.* (2018b) (A), Diaz-Garcia *et al.* (2018a) (B), Schlautman *et al.* (2015) (C), and a combined synthesis of Diaz-Garcia *et al.* (2018a) and Schlautman *et al.* (2015) (D) in cranberry populations *CNJ02*, *CNJ04*, and *GRYG*. 57

Figure III-1. A. Geographic map showing location of study's three cranberry breeding populations and the proximal weather station with daily temperature data. B. Spring temporal patterns for mean reference A535 (dark red) and CDD (dark blue) for years 2018-19. Dashed lines represent the corresponding best fit curve for both datasets. Points represent the mean A535 values measured on a given month and day. 79

Figure III-2. UAV cranberry imaging process topology. Breeding field plots for population CNJ02 and CNJ04 were imaged at Cranberry Creek Cranberries (site 1) and field plots for population GRYG were assessed at Tomah Corporation (site 2). Datasets represent a given capture date and are numbered incrementally from Z001. Sessions represent distinct, redundant, and comprehensive captures of all field plots at a given location. 80

Figure III-3. Pearson's correlations across three cranberry populations. A. Pairwise CDD-by-trait correlations ranked by strength of association. B. Pairwise A535-by-trait correlations ranked by strength of association. A.-B. Non-significant ($p \geq 0.05$) correlations are labeled with "NS". Correlations are displayed and sorted as absolute values and are colored according to their correlation sign – positive correlations are red, while negative correlations are blue. The correlations are scaled exponentially to better differentiate stronger associations. C. Heatmap of absolute pairwise correlations between VIs. 89

Figure III-4. Fit benchmarks across five regression model types in three cranberry populations. Each column's error bar represents the 95% confidence interval of the respective benchmark after 10-fold cross-validation repeated 200 times. The horizontal black line with corresponding text indicates the median value and the column extents representing the 25% and 75% quartiles. R^2 is the model coefficient of determination, *RMSE* is the root mean square error, and MAE is the mean absolute error. Number of features (*m*), displayed in the legend for each model type, represents the number of predictor variables used to fit the model. Predictors are synthetic principal components for PCR and PLSR or relevant SRs and VIs for RR, OLSE, and SA methods. 90

Figure III-5. Model fit parameter estimates for simulated annealing (SA) selected vegetative indices (VIs) across three cranberry breeding populations. A. SA regression coefficient estimates ($\hat{\beta}$) for the VIs listed in the y-axis along with the coefficient p-values. The horizontal bars represent the 95% confidence interval of estimates for each predictor coefficient. B. ANOVA effect size estimates ($\widehat{\eta_p^2}$) of each VI, where larger relative values indicate higher meaningful effect contributions. Figure was generated using the ggcoefstats() function of the ggstatsplot R package (Patil, 2021). 92

Figure III-6. Coefficient of variation (CV) of regression A535 predictions across repeated photos of the same plot genotype within an imaging session. Plot was generated using the ggwithinstats() function of the ggstatsplot R package (Patil, 2021). CV was calculated across all three cranberry populations of this study. Statistical groupings between regression methods were determined using

Holm-adjusted (Holm, 1979) pairwise p-values, with group membership designated using letter characters and distinctions set at $p < 0.05$. 93

Figure III-7. SA regression of three cranberry populations - fitted vs. observed plots. Panel A represents the regression fit for populations CNJ02 and CNJ04. Panel B illustrates the regression fit for population GRYG. Ellipses circumscribe all datapoints associated with a given UAV flight date. The colors associated with the ellipses and datapoints are graded using a linear function mapping CDD inputs to informative color outputs: low CDD values map to dark red and high CDD values map to dark green. Each ellipse is labeled with its associated date in month-day-year format and its associated CDD. 94

Figure III-8. Sample UAV images from three cranberry populations representing gradient of top quintile to bottom quintile of SA model predicted A535. Each column represents a sample of two plot images within that column's associated quintile. Quintiles are ordered highest to lowest, from left to right. Quintiles are labeled top (quintile 1), Q75 (upper quartile), median (quintile 0.5), Q25 (lower quartile), and bottom (quintile 0). Each plot is labeled with its predicted A535, the date at which the plot image was captured (in month-day-year format), and the calculated CDD on that image date. 95

Figure III-9. Cranberry population CNJ02 leaf-color development for top- (row *2*) and bottom-performing (row *3*) genotypes in 2018. Column *A* contains functional graphs of the predicted A535 versus CDD for each genotype. These plots contain the exponential decay curve fit functions for the genotype datapoints (solid line), population datapoints (dashed line), and color-coded area between the two curve types. Yellow indicates an area that positively contributes to the selection index score, while blue indicates a negative contribution. The dashed vertical line at CDD equal to 200 heat units demarcates the inflection point at which the integral area difference function is inverted (Equation III-3). Each graph is labeled with its respective genotype exponential decay fit function and the fit curve's associated R^2, while the population's decay fit curve is described in row *1*, column *A*. Row *1*, columns *B-E* show the probability distributions of the predicted A535 for a given genotype, with labeled vertical dashed lines indicating the location of the top- (‡) and bottom-performing (§) genotypes. Each distribution graph is colored according to its A535 bin value: redder values indicating higher A535 and greener indicating lower A535. Each column of *B-E* corresponds to a given CDD value and date (month-day-year), with a representative plot image for each genotype at that timepoint. Columns *B-E* are sorted with increasing CDD from left to right to indicate leaf developmental progression. Each representative plot image is labeled below with its predicted A535. All dates are formatted as month/day/year. 96

Figure III-10. Cranberry population CNJ02 leaf phenology progression for top- (row *2*) and bottom-performing (row *3*) genotypes in 2019. Column *A* contains functional graphs of the predicted A535 versus CDD for each genotype. These plots contain the exponential decay curve fit functions for the genotype datapoints (solid line) and population datapoints (dashed line), along with color-coded area between the two curve types. Yellow indicates an area that positively contributes to the selection index score, while blue indicates a negative contribution. The dashed vertical line at CDD equal to 200 heat units demarcates the inflection point at which the integral area difference function is inverted (Equation III-3). Each graph is labeled with its respective genotype exponential decay fit function and the fit curve's associated R^2, while the population's decay fit curve is described in row *1*, column *A*. Row *1*, columns *B-E* show the probability distributions of the predicted A535 for a given genotype, with labeled vertical dashed lines indicating the location of the top- (‡) and bottom-performing (§) genotypes. Each distribution graph is colored according to its A535 bin value: redder values indicating higher A535 and greener indicating lower A535. Each column of *B-E* corresponds to a given CDD value and date (month-day-year), with a representative plot image for each genotype at that timepoint. Columns *B-E* are sorted with increasing CDD from left to right to indicate leaf developmental progression. Each representative plot image is labeled below with its predicted A535. All dates are formatted as month/day/year. 97

Figure III-11. Cranberry population GRYG leaf phenology progression for top- (row *2*) and bottom-performing (row *3*) genotypes in 2018. Column *A* contains functional graphs of the predicted A535 versus CDD for each genotype. These plots contain the exponential decay curve fit functions for the genotype datapoints (solid line) and population datapoints (dashed line), along with color-coded area between the two curve types. Yellow indicates an area that positively contributes to the selection index score, while blue indicates a negative contribution. The dashed vertical line at CDD equal to 200 heat units demarcates the inflection point at which the integral area difference function is inverted (Equation III-3). Each graph is labeled with its respective genotype exponential decay fit function and the fit curve's associated R^2, while the population's decay fit curve is described in row *1*, column *A*. Row *1*, columns *B-E* show the probability distributions of the predicted A535 for a given genotype, with labeled vertical dashed lines indicating the location of the top- (‡) and bottom-performing (§) genotypes. Each distribution graph is colored according to its A535 bin value: redder values indicating higher A535 and greener indicating lower A535. Each column of *B-E* corresponds to a given CDD value and date (month-day-year), with a representative plot image for each genotype at that timepoint. Columns *B-E* are sorted with increasing CDD from left to right in order to indicate leaf developmental progression. Each representative plot image is labeled below with its predicted A535. All dates are formatted as month/day/year. 98

Figure IV-1. Distributions and summary statistics of the original *LMI* trait values and their fitted BLUPs for cranberry population CNJ02. 124

Figure IV-2. Significant QTL found for trait *LMI* in cranberry population CNJ02. Effect size plots for the four highest-ranking QTL by R^2 are displayed along the figure margins. Crossbar plots of mean effect sizes across four possible QTL genotypes are shown at the top and bottom of the figure, with their associated standard error. Bar plots of the average maternal, paternal, and interaction effect sizes are shown at the left and right margins. The outer ring of the circos plot shows the QTL (circles) and their associated LOD-1.5 interval, with QTL sizes proportional to their respective R^2. The center ring displays the linkage map marker locations, and the innermost circos sector shows the percent variance explained by the QTL model. The center plot demonstrates the CNJ02 population's pedigree, where each node is labeled with its cultivar abbreviation, with edge arrows indicating the crossing direction. Cultivar abbreviations are as follows: HO = Howes, SL = Searles, BL = Ben Lear, MF = McFarlin, PO = Potters Favorite, LM = LeMunyon, 35 = #35, CQ = Crimson Queen, ST = Stevens, MQ = Mullica Queen. Red edges represent the ovule parent, dark blue the pollen parent. Edge letters, where specified, are the QTL haplotypes associated with the respective grandparent. 127

Supplementary Figure B-1. Leaf maturity index (LMI) trait QQ (**A**) and distribution (**B**) plots of three cranberry breeding populations CNJ02, CNJ04, and GRYG. 153

List of Tables

Table II-1A. Trait groupings for each set of co-located QTL in present study. 47

Table II-2. Cranberry population CNJ02 trait and BLUP summary for all-year model with heritability above 0.47. Only models with significant (*p* < 0.05) genotypic effects are displayed. 55

Table II-3. Cranberry population CNJ04 trait and BLUP summary for all-year model with heritability above 0.47. Only models with significant (*p* < 0.05) genotypic effects are displayed. 60

Table II-4. Co-located, stable QTL in cranberry population CNJ02 found in at least three of four BLUP models. Only QTL with $R^2 \geq 10$ % are shown. Table is arranged in descending order by mean marker variance. 61

Table II-5. Co-located, stable QTL in cranberry population CNJ04 found in at least three of four BLUP models. Only QTL with $R^2 \geq 10$ % are shown. Table is arranged in descending order by mean marker variance. All QTL were reported below were identified with both scanone and stepwiseqtl methods. 62

Table III-1. Vegetative indices (VIs) used to study cranberry breeding field plots for cranberry populations CNJ02, CNJ04, and GRYG. 81

Table III-2. Reference data types and descriptions for cranberry leaf tissue chlorophyll and anthocyanin levels. This reference data is derived from documented anthocyanin and chlorophyll pigment spectral absorbance curves (Wintermans and De Mots, 1965; Fuleki and Francis, 1968; Workmaster, 2001). Values are unitless but recorded as absorbance units (AU) at defined spectral wavelength (nm), or are ratios between two absorbance readings. 84

Table III-3. Top 5 percent of genotypes (up to maximum of 10 genotypes) per cranberry population using the selection index defined in Equation III-3. The genotypes are ranked from higher to lower from top to bottom, with the associated selection index score to the right of each genotype. 98

Table III-4. Bottom 5 percent of genotypes (up to maximum of 10 genotypes) per cranberry population using the selection index defined in Equation III-3. The genotypes are ranked from lower to higher from top to bottom, with the associated selection index score to the right of each genotype. 99

Table IV-1. Cranberry population CNJ02 consensus QTL for *LMI*, sorted from highest to lowest R_q^2. 125

Supplementary Table A-1. Pairwise correlations of all VIs used when fitting regression models. These correlations are derived from all UAS images of cranberry plots, including those without ground truth reference data. 150

Supplementary Table A-2. Pairwise correlations of VIs, spectrophotometer absorbances, and estimated CDD on reference data gathered from cranberry plots. 151

Supplementary Table B-1. Gene counts in the genomic neighborhood of QTL for the LMI trait in cranberry population CN02. Neighborhood is defined as the genomic regions associated with the LOD 1.5 intervals of QTL. These counts are for all gene models that have homology to any characterized genes or motifs in the scientific literature. Gene models were taken from annotations in the Stevens (Diaz-Garcia et al., 2021) and Ben Lear (Kawash et al., 2022) cranberry genome assemblies. 153

Supplementary Table B-2. Putative cranberry candidate genes relevant to dormancy processes that are in the genetic neighborhood of QTL for the LMI trait in cranberry population CNJ02. Sequences of the flanking 1.5 LOD interval markers were aligned to the Stevens (Diaz-Garcia et al., 2021) and Ben Lear (Kawash et al., 2022) annotated cranberry genome assemblies. 154

List of Equations

Equation II-1. Mixed model for estimating BLUPs within years. 36

Equation II-2. Mixed model for estimating BLUPs across years. 37

Equation II-3. Narrow sense heritability calculation for within-year mixed model. 38

Equation II-4. Narrow sense heritability calculation for across-years mixed model. 38

Equation III-1. Calculation for adjusted growing degree day (*GDD*) and cumulative degree days (*CDD*) for American cranberry. 83

Equation III-2. General linear regression model form. 85

Equation III-3. Leaf maturity index (LMI) calculation, where $s_g(x)$ is the score at CDD x, $\widehat{r_g}(x)$ is the genotype-fit curve A535 value at CDD x, $\widehat{r_p}(x)$ is the population-fit curve A535 value at CDD x, x_c is the critical frost period threshold CDD, S_g is the selection index of genotype g, x_s and x_f are the starting and ending CDD values that bound integration of $s_g(x)$. 87

Equation IV-1. Leaf maturity index (LMI) calculation, where $s_g(x)$ is the score at CDD x, $\widehat{r_g}(x)$ is the genotype-fit curve A535 value at CDD x, $\widehat{r_p}(x)$ is the population-fit curve A535 value at CDD x, x_c is the critical frost period threshold CDD, S_g is the selection index of genotype g, x_s and x_f are the starting and ending CDD values that bound integration of $s_g(x)$. 119

Equation IV-2. Mixed model for estimating BLUPs within years. 120

Equation IV-3. Narrow sense heritability calculation for within-year mixed model. 121

Chapter I Literature Review: Advancing Cranberry Breeding by Integrating Genetics with Aerial Imaging

History and Economics

The American cranberry (*Vaccinium macrocarpon* Ait.) represents a significant species within the *Vaccinium* genus, which encompasses close to 500 species that have adapted to various marginal habitats (Vander Kloet, 1988; Wagner et al., 1990). Members of this genus demonstrate remarkable adaptation to ecological niches including acidic peatlands, volcanic environments, and forest canopies (Vander Kloet, 1988; Vander Kloet and Avery, 2010). This evolutionary adaptation to challenging environments positions cranberry as a crop with potential for cultivation in areas unsuitable for conventional agriculture, thus potentially expanding agricultural production beyond traditional arable land.

Prior to its formal breeding, indigenous communities across North America utilized cranberries for millennia, recognizing their nutritional, healing (antiseptic), and storage properties. The natural preservative qualities of cranberries were used by indigenous groups to prevent spoilage of high calorie pemmican, an original paleo energy bar consisting of dried fruit, animal fat, and meat. This pemmican not only supplied the calories needed to match high energy demands to survive winters but also provided essential nutrients such as vitamin C to stave off scurvy (Eck, 1990). These properties provided cranberries with incredible social value, acting as gastronomic symbols of peace during communal feasts (Norwood, 1936). Later, European colonists learned the value of this native crop from indigenous peoples, providing needed nutrition in the first winters that would otherwise have decimated their populations in early settlements. Early American ships and later whaling ships even supplied their ships with barrels of cranberries as both a source of Vitamin C and as a valuable trade commodity (Hall, 1952; Eck, 1990).

The historical narrative of cranberry did not expand to include managed cultivation until 1810, when Henry Hall of Dennis, Massachusetts, applied sand to cranberry propagules after observing increased productivity in cranberry stands exposed to wind depositions of sand (Hall, 1941). A two-decade “cranberry fever” started in the 1850s, driven by new markets and high prices and defined by intensified planting and cultivation of notable native selections along the northeastern United States seaboard. By the early 1900s, over 132 native varieties from Massachusetts, New Jersey, Wisconsin, and Michigan had been selected for intensive cultivation (Vorsa and Zalapa, 2019). The 20th century was defined by rapid and continued growth of the Wisconsin cranberry industry, establishment of new plantings in western states, emergent pest and disease pressures, increased scientific research to address industry problems, rapid improvements in effective management practices, innovations that advanced harvesting efficiency, and punctuated but effective breeding efforts leading to more productive cultivars (Eck, 1990). The late 20th and early 21st centuries were defined by second-generation releases of new cultivars targeting color, quality, early maturity, and perennial productivity and research unraveling/uncovering/detailing. New issues include market saturation driven by surplus outpacing demand, berry quality losses from abiotic and biotic stressors driven by climate change, and new breeding initiatives focused on improving berry quality attributes needed to produce higher value products like sweetened dried cranberries (Vorsa and Zalapa, 2019). Breeding for resistance to abiotic and biotic stressors reduces management costs while increasing effective yield through higher quality.

Among commercially cultivated *Vaccinium* species, cranberry shares economic significance with blueberries and lingonberries (Chandler et al., 1947). Cranberry constitutes a significant agricultural commodity with global production reaching approximately 700,000 metric tonnes in 2023 (United Nations FAO, 2023). The United States contributes just over 50% of world production, followed by Canada and Chile (~22-23% each), with the remaining production scattered across several eastern European and western Asian countries. Within the United States, the unprocessed cranberry market generated just over $300 million in 2023, a figure that increases substantially when value-added products are considered (Alston, Julian M. et al., 2014; USDA National Agricultural Statistics Service, 2023).

Morphology and Trait Granularity

Cranberry possesses distinctive morphological and physiological characteristics that differentiate it from most cultivated fruit crops. The plant develops as a low, trailing woody vine that produces specialized vertical stems termed "uprights." These uprights are classified into two functional categories with distinct developmental patterns: vegetative uprights, which develop exclusively leaf-producing buds, and reproductive uprights, which bear both leaf and flower buds. The balance between these upright types significantly influences productivity, with bud induction and differentiation regulated through complex interactions between genetic factors and environmental conditions (Bolivar-Medina et al., 2019).

Traditionally uprights served as the elemental unit for breeding and physiological research. Modern breeding has started to move away from the granular scale of the 'upright' and towards 'whole plot' assessments, driven by increased throughput and efficiencies provided by technological phenotyping tools (Diaz-Garcia et al., 2016, 2018b; Maule et al., 2024; Loarca et al., 2024; Akiva et al., 2020, 2022; Johnson et al., 2023). More efficient trait collection is necessary to assess large numbers of crosses at scale, which is important to both the effectiveness of a breeding program and for building genomic prediction models. In contrast, physiological research continues to focus on 'upright' as the base unit of scientific study, given that much of the physiological characteristics of cranberry plants seem to be driven by the low-growing, lawn-like architecture of the plant, with short-distance travel between nutrients in roots and locally relevant hormonal and carbohydrate impacts within an upright. Scaled up, these physiological effects can be massive, but understanding their mechanisms and effects is best done at the granularity of an upright.

Breeding Challenges and the Path to Improvement

Cranberry cultivation requires distinct management practices and breeding considerations that differentiate it from most other fruit crops, presenting unique challenges and opportunities for agricultural advancement. Despite a century of commercial cultivation, cranberry domestication remains limited compared to many agricultural crops. Most commercial cultivars represent only first or second-generation selections from wild populations, with third-generation varieties only recently entering production systems (Vorsa and Zalapa, 2019).

This restricted domestication progress results from several interconnected factors that limit breeding efficiency. Historical cranberry improvement programs have experienced inconsistent research funding, creating periods of advancement followed by stagnation (Eck, 1990; Vorsa and Zalapa, 2019). Cranberry's perennial nature necessitates substantial space allocation for adequate evaluation of breeding materials (Vorsa and Zalapa, 2019). Last, the evaluation cycle for new cranberry varieties extends six to eight years after initial planting, creating an exceptionally protracted feedback mechanism between selection decisions and performance verification (Diaz-Garcia et al., 2020). This extended timeline means a cranberry breeder might realistically complete only 3-4 selection cycles during their professional career—a significant limitation compared to breeding programs for annual crops.

These constraints underscore the necessity for technologies that accelerate the breeding process. Advanced genomic tools offer potential for performance prediction prior to phenotypic expression, while modern phenomic approaches may identify early indicators of desirable traits, thus reducing evaluation timelines. The integration of these technologies could transform cranberry improvement from its historically incremental pace to a more efficient breeding system.

Genomic Advances in Cranberry Research

Recent years have witnessed substantial progress in cranberry genomics, facilitated by technological innovations that have democratized access to sophisticated molecular tools. The reduction in DNA sequencing costs, combined with advancements in analytical methods, has enabled cranberry researchers to develop genomic resources comparable to those available for major agricultural crops (Covarrubias-Pazaran et al., 2018).

This genomic advancement has proceeded through several significant achievements. Researchers have assembled complete organellar genomes, including plastid and mitochondrial sequences (Fajardo et al., 2012, 2014; Diaz-Garcia et al., 2019), providing insights into maternal inheritance patterns and evolutionary relationships. More significantly, two high-quality nuclear genome assemblies now serve as reference resources (Diaz-Garcia et al., 2021; Kawash et al., 2022), enabling precise characterization of genetic variations and functional elements.

Transcriptome analyses have elucidated the molecular mechanisms governing important biochemical pathways, particularly those involved in flavonoid synthesis and regulation (Georgi et al., 2013; Sun et al., 2015; Diaz-Garcia et al., 2021). These investigations establish connections between genetic variations and specific biochemical outcomes, providing potential targets for quality improvement.

High-density genetic maps constructed using microsatellite markers and genotyping-by-sequencing approaches have established the foundation for understanding trait inheritance patterns (Georgi et al., 2013; Schlautman et al., 2015, 2017; Covarrubias-Pazaran et al., 2016; Daverdin et al., 2017). These maps enable precise localization of genetic factors controlling agriculturally important characteristics, facilitating marker-assisted selection approaches.

Initial applications in association mapping and genomic selection have begun translating these fundamental resources into practical breeding tools (Covarrubias-Pazaran et al., 2018; Diaz-Garcia et al., 2020; Neyhart et al., 2022). These approaches offer potential to accelerate genetic gain by enabling selection decisions based on DNA markers rather than requiring complete phenotypic expression.

Collectively, these genomic resources have transformed cranberry from a genetically understudied crop to one with sophisticated molecular breeding capabilities. The integration of these tools with advanced phenotyping methods creates opportunities to substantially accelerate the crop improvement process, which is particularly valuable for a perennial crop with extended evaluation cycles.

Quantitative Trait Loci: Elucidating the Genetic Architecture of Cranberry Traits

Quantitative trait loci (QTL) studies in cranberry have progressed from initial exploratory investigations to comprehensive multi-trait analyses across diverse populations. This progression has expanded understanding of the genetic factors controlling economically important characteristics while establishing connections between traditional phenotyping approaches and modern high-throughput methods.

Early QTL mapping studies utilized limited marker sets but nonetheless identified significant genetic factors. Georgi et al. (2013) conducted pioneering work by combining cranberry sequence characterized amplified regions (SCAR), blueberry-derived microsatellites, and targeted markers for functional candidate genes. Their analysis across four related populations identified QTL for critical production characteristics, including field fruit-rot resistance, titratable acidity, fruit weight, and sound fruit yield.

Subsequent investigations employed more sophisticated genetic maps and phenotyping methodologies. Schlautman et al. (2015) constructed a high-density microsatellite linkage map in the elite breeding population CNJ02-1, enabling detection of QTL governing mean fruit weight, total yield, and biennial bearing patterns. This research established the foundation for understanding yield component traits with direct economic implications.

Daverdin et al. (2017) focused on disease resistance, developing detailed linkage maps across four diverse populations selected for segregation in fruit rot susceptibility. Their research identified

15 QTL controlling resistance responses while demonstrating the genetic independence of yield characteristics from disease resistance mechanisms—an important finding for developing cultivars that combine both traits.

The integration of advanced imaging technologies with QTL mapping has further enhanced understanding of fruit quality traits. Diaz-Garcia et al. (2018b, 2018a) developed novel image analysis techniques to efficiently quantify fruit shape, size, and color parameters—characteristics relevant to both fresh market quality and processed product applications. Their approach identified multiple QTLs controlling anthocyanin production, fruit dimensions, and shape factors, which are particularly important for the sweetened dried cranberry market.

A comprehensive meta-QTL analysis by Maule et al. (2024) synthesized findings across multiple years, populations, and studies, identifying genomic regions with consistent effects across diverse environments and measurement methods. This research validated high-throughput phenotyping approaches against traditional evaluation techniques, establishing the reliability of modern methods for future breeding applications.

These QTL studies serve a fundamental purpose beyond academic understanding—they establish the predictive relationships between DNA markers and phenotypic outcomes that enable more efficient selection. By mapping these associations, breeders can potentially reduce evaluation costs and shorten breeding cycles, addressing the key constraints that have historically limited cranberry improvement efficiency. Even modest improvements in selection accuracy translate to significant efficiency gains in perennial crops like cranberry, where each evaluation cycle represents substantial time and resource investment (Bernardo, 2008; Desta and Ortiz, 2014).

The Evolution of Phenotyping Approaches in Cranberry Evaluation

Concurrent with advances in genomics, cranberry phenotyping has undergone significant methodological evolution. Traditional assessment methods, while effective, have been limited by their labor-intensive, subjective, and often destructive nature (Hawkesford and Lorence, 2017; Araus et al., 2018). These constraints have created efficiency limitations in the breeding process that modern technologies now address through automated, high-throughput approaches.

Historically, cranberry evaluation focused on reproductive upright measurements, with breeders meticulously assessing individual fruiting branches to characterize yield potential and fruit quality. This approach provided detailed data but required substantial time and labor investment. The manual nature of these assessments also introduced potential inconsistencies due to evaluator variability, fatigue effects, and subjective judgment. Recognizing these limitations, the field has progressively transitioned toward "whole plot" measurements that characterize overall performance across standardized growing areas (Vorsa and Zalapa, 2019).

The technological transformation of cranberry phenotyping began with digital imaging applications for fruit assessment. GiNA, developed specifically for horticultural crops, established quantitative analysis of external fruit characteristics including precise shape parameters and color values previously assessed through subjective ratings (Diaz-Garcia et al., 2016). This platform has subsequently been succeeded by BerryPortraits, a fruit image analysis platform with enhanced segmentation algorithms and color correction capabilities that standardize evaluations across variable imaging conditions (Loarca et al., 2024).

Remote sensing technologies have expanded phenotyping capabilities from individual fruit to whole-plant and field-scale assessments. Hyperspectral and satellite imaging have demonstrated utility for mapping nutritional status across cranberry beds (Liu et al., 2023; Huang et al., 2024), though their application in genetic studies remains constrained by resolution limitations that cannot adequately distinguish the small plots typical in breeding trials.

Uncrewed aerial vehicles (UAVs) equipped with specialized imaging systems have emerged as a particularly promising solution for breeding applications. Recent innovations combine high-resolution UAV imagery with neural network analysis to assess plant health, development stage, and productivity potential (Akiva et al., 2020, 2022; Johnson et al., 2023). These systems overcome the resolution limitations of satellite platforms while providing the field-scale coverage impossible with laboratory-based imaging systems.

The transition from traditional to high-throughput phenotyping approaches has not eliminated the requirement for ground-truth validation, but it has significantly expanded the scale and precision of data collection while reducing subjective elements in the evaluation process. As these technologies mature and integrate with genomic selection approaches, they offer potential to address the phenotyping limitations that have traditionally constrained breeding progress.

Technological Integration in Modern Crop Phenomics

The emergence of phenomics represents more than the application of individual technologies—it reflects the systematic integration of multiple innovations that collectively transform plant evaluation capabilities. This integration encompasses advancements in sensing platforms, data management systems, analytical algorithms, and integration methodologies that enable comprehensive insights into plant performance (Zhao et al., 2019; Song et al., 2021).

Multiple technological advances have converged to facilitate this phenomic revolution. The reduction in sensor costs has expanded access to sophisticated imaging equipment, placing multispectral, thermal, and high-resolution optical capabilities within reach of more research programs. Concurrently, robotics and UAVs have evolved from experimental prototypes to reliable platforms for deploying these sensors at scale (Yang et al., 2017; Araus et al., 2018).

The data management infrastructure has transformed through cloud computing solutions capable of processing the extensive datasets generated by repeated, high-resolution imaging across large field trials. Open-source software tools have standardized analysis workflows, ensuring methodological consistency while reducing implementation barriers (Gehan et al., 2017; Berry et al., 2018; Zhao et al., 2019).

Machine learning algorithms have revolutionized the interpretation of complex, multi-dimensional phenotypic data. These approaches now extract meaningful patterns from imagery to predict diverse plant characteristics, including developmental timing, stress responses, morphological traits, and yield components (Zhao et al., 2019; Weiss et al., 2020).

Within this broader phenomic context, conventional RGB photography has emerged as an effective and accessible tool. Despite their apparent simplicity, RGB cameras capture wavelengths that correspond with key plant pigment absorption and reflection patterns, particularly for chlorophyll and anthocyanin compounds (Carter and Knapp, 2001; Mahlein, 2016; Del Valle et al., 2018). This spectral correspondence enables extraction of biologically meaningful information from standard photographs through vegetation indices—mathematical combinations of color channel values that enhance specific plant characteristics.

Several RGB-derived indices have demonstrated utility for assessing pigment composition and physiological status. The red-to-green ratio effectively quantifies leaf anthocyanin levels (Gamon and Surfus, 1999), while normalized difference calculations between red and green channels enable vegetation segmentation and biomass estimation (Woebbecke et al., 1993; Pérez et al., 2000). More sophisticated indices like the excess green minus excess red index provide effective separation between plant tissue and soil backgrounds (Meyer and Neto, 2008), while specialized formulations like anthocyanin content-chroma ratio (AC-CR) achieve accurate pigment quantification across diverse plant tissues (Del Valle et al., 2018).

These vegetation indices utilize the characteristic absorption patterns of plant pigments across the visible spectrum. Chlorophylls exhibit peak absorption in blue (435-460nm) and red (660-680nm) wavelengths, while anthocyanins absorb primarily in green wavelengths around 550nm. By monitoring these spectral signatures, researchers can track physiological changes associated with development, stress responses, and seasonal transitions (Sims and Gamon, 2002).

The integration of these diverse technologies—from sensors and platforms to algorithms and indices—has transformed phenotyping from a limiting constraint to a facilitator of breeding progress. Systems that previously required specialized facilities and expertise have evolved into accessible tools applicable across diverse research contexts, including traditionally under-resourced crop improvement programs.

Monitoring Spring Development to Address Frost Vulnerability

Spring development represents a critical vulnerability period for cranberry production, when emerging floral buds face substantial risk from late frost events that can significantly reduce yield potential. Understanding and potentially modifying the timing of this development phase offers an approach to enhancing cranberry resilience under variable climate conditions.

The evergreen nature of cranberry provides an opportunity to monitor spring development through leaf color transitions. During winter dormancy, cranberry foliage accumulates anthocyanin pigments that produce characteristic red coloration. As temperatures increase in spring, these pigments diminish while chlorophyll levels increase, creating a visible shift from red to green that corresponds with physiological changes in the plant (Workmaster, 2001).

This pigmentation pattern allows assessment of developmental progression using imaging approaches. High-definition RGB cameras mounted on UAV can effectively capture these color transitions across entire breeding populations, enabling quantitative comparison of development rates among different genotypes. The timing and rate of these transitions may indicate underlying differences in cold adaptation strategies that influence frost susceptibility.

Recent advancements in UAV monitoring have combined high-resolution imagery with neural network analyses to evaluate cranberry development patterns (Akiva et al., 2020, 2022; Johnson et al., 2023). While initially focused on modeling heat stress risks and ripening patterns, these approaches demonstrate potential for characterizing spring development and frost vulnerability.

By tracking anthocyanin reduction and chlorophyll accumulation through vegetation indices derived from RGB imagery, researchers can quantify development rates without destructive sampling. This non-invasive monitoring enables repeated assessment of the same plants throughout the critical spring period, generating developmental profiles for each genotype that may reveal patterns associated with enhanced frost tolerance.

The capacity to monitor these transitions efficiently across large breeding populations presents opportunities for selecting genotypes with optimized development patterns, potentially identifying those that maintain dormancy during high-risk frost periods but transition rapidly to active growth once danger has passed. This approach could facilitate development of varieties with inherent resistance to spring frost damage, providing genetic solutions to challenges traditionally addressed through resource-intensive management interventions.

Genomic Selection Applications for Cranberry Improvement

The integration of genomic technologies with comprehensive phenotypic data represents a significant opportunity for cranberry breeding. By combining dense molecular marker information with precisely measured traits, researchers can develop predictive models that potentially circumvent the extended evaluation cycles that have historically limited cranberry improvement.

The foundation for this genomic approach lies in the extensive molecular resources developed over the past decade (Covarrubias-Pazaran et al., 2016, 2018; Diaz-Garcia et al., 2018b, 2019, 2021; Kawash et al., 2022; Maule et al., 2024). These investments in sequencing, genotyping, and genetic mapping provide the data infrastructure for implementing advanced selection methodologies.

Genomic selection represents a fundamental shift in the breeding process by employing statistical models trained on a reference population with phenotypic and genotypic data, then applying these models to predict performance in individuals characterized solely by genetic markers (Meuwissen et al., 2001). This methodology captures major and minor genetic effects across the entire genome, offering advantages over traditional marker-assisted selection that typically focuses on a limited number of large-effect loci.

For cranberry improvement, the potential benefits of genomic selection are particularly significant. By predicting performance prior to reproductive maturity, breeders could potentially reduce evaluation time by several years, substantially accelerating selection intensity and genetic gain per unit time. This acceleration effect becomes particularly valuable in perennial crops with extended generation intervals (Bernardo, 2008; Covarrubias-Pazaran et al., 2018; Ferrão et al., 2021).

Genomic prediction models also offer potential for simultaneous improvement of multiple traits through selection indices that balance competing characteristics. By weighting traits according to their economic importance and genetic correlations, breeders can address complex trade-offs that frequently complicate selection decisions.

Initial applications of genomic selection in cranberry have yielded promising results (Covarrubias-Pazaran et al., 2018), suggesting that accurate prediction of complex traits like yield is achievable using current marker technologies and modeling approaches. As phenotyping capabilities expand to include novel traits such as spring development patterns and frost tolerance, these models can incorporate additional selection targets to develop increasingly resilient and productive cultivars.

The continued refinement of phenomic and genomic technologies, combined with advances in statistical modeling approaches, positions cranberry breeding for significant acceleration in improvement rates. By reducing dependence on extended field evaluations while enhancing selection accuracy, these approaches address the fundamental constraints that have historically limited cranberry domestication and improvement.

Research Gaps and Future Directions

Despite substantial advances in cranberry genomics and phenomics, several significant challenges and research opportunities remain. These gaps represent both limitations of current approaches and potential avenues for future investigation that could further accelerate crop improvement.

The development and validation of phenomic models continues to require substantial investment in ground-truth data collection—a process that remains labor-intensive despite technological advances. While imaging systems efficiently capture plant characteristics, the interpretation of this visual information requires calibration against direct measurements to ensure accuracy. Establishing efficient protocols for collecting validation data requires balancing methodological rigor with practical resource limitations.

Technical challenges persist in standardizing data across variable environmental conditions, particularly for field-based imaging approaches. Multi-spectral and hyperspectral sensors require precise calibration to account for atmospheric conditions, illumination variations, and differences in canopy architecture (Zhao et al., 2019). While color correction algorithms have improved, achieving consistent measurements across diverse sampling conditions remains challenging.

Current phenomic models frequently demonstrate limited transferability between environments or populations—a constraint that necessitates local calibration before application in new settings. Developing more generalizable approaches that maintain accuracy across diverse conditions would enhance the utility of these tools for breeding programs operating across multiple locations.

Human factors in phenotypic data collection introduce additional variables that technological solutions have yet to fully address. Traditional scoring approaches remain susceptible to inconsistency, subjective judgment, and evaluator fatigue (Andersson and Prager, 2006; Araus et al., 2018). While automated systems mitigate these concerns for some traits, many complex characteristics still rely on human assessment, creating potential evaluation bottlenecks.

Despite these challenges, several research directions could substantially advance cranberry improvement. Validating modern phenotyping methodologies against traditional approaches would establish confidence in newer technologies while potentially identifying complementary strengths of each method. This validation process could facilitate effective transitions between evaluation systems while maintaining historical continuity.

Developing integrated phenomic-genomic approaches for specific high-value traits represents another significant opportunity. By focusing advanced technologies on characteristics with substantial economic impact, particularly those difficult to assess through conventional means, researchers could maximize return on investment for these methodological innovations.

Identifying the genetic control mechanisms governing spring development timing and frost tolerance could provide valuable selection targets for enhancing cranberry resilience to climate variability. Understanding these adaptive traits at both physiological and molecular levels would enable more precise breeding interventions to address the significant production risks associated with spring frost events.

Standardizing phenomic data collection and analysis protocols across research institutions would facilitate more effective knowledge sharing and collaborative improvement efforts. Establishing common frameworks for sensor calibration, data processing, and model validation would enable more direct comparison between studies while accelerating the adoption of effective methodologies.

The integration of genomics, phenomics, and traditional breeding expertise offers potential to develop climate-resilient cranberry cultivars. By combining frost tolerance with other adaptive traits, breeders could create varieties capable of maintaining productive stability across increasingly variable environmental conditions, addressing one of the most significant challenges facing modern agriculture.

These research directions collectively indicate potential for cranberry improvement to advance beyond its historically incremental pace to achieve substantial progress in both productivity and sustainability. By addressing current limitations while leveraging emerging technologies, cranberry breeding is positioned to enter a period of accelerated genetic progress and agricultural innovation.

Thesis Objectives

This research aims to build a throughline from traditional trait evaluation modalities to newer and more advanced phenotyping methods in cranberries. It endeavors to bridge traditional 'upright' traits with 'whole plot' traits, both to validate the integrity of newer selection criteria and to enable utilization of newer, more robust, and higher throughput technologies in cranberry breeding. These technological applications empower and accelerate cranberry breeding, conventionally a slow process, in a world with fluctuating market dynamics and emerging climate instability.

In addition to connecting and validation old techniques with new, this research looks to demonstrate the utility of a low-cost UAV with a RGB camera to phenotype spring developmental patterns in three distinct cranberry breeding populations. By monitoring spring leaf color shifts, building predictive models to infer plot pigmentation levels, and generating a selection index tied to favorable pigmentation development patterns, we enable mapping to elucidate associated genetic markers.

This thesis consists of the following objectives:

- Map QTL from traditional phenotyping methods and link findings with QTL found in previous studies from 'whole plot' traits.
- Develop models to efficiently and effectively infer spring developmental patterns from UAV imagery in three cranberry breeding populations.
- Map QTL from the model-predicted phenotypes found in objective II.

This research demonstrates an efficient UAV-phenotyping and genetic mapping pipeline to accelerate cranberry selection and breeding. The application of this pipeline to additional UAV-monitored traits collected as a part of this research will help not just to build robust selection indices for single traits but will also provide a means to develop more holistic, multi-trait indices that drive breeding decisions. Moreover, the knowledge acquired here can be applied using other image-associated tools, such as ground-based rovers, with the goal of filling in the temporal or spatial data resolution gaps that limit the current research's statistical power. Although UAVs are more flexible than ground-based imaging tools, they are often constrained by intrinsic and extrinsic factors. Intrinsic factors include their high-power requirements, hard weight limits, and utility to navigate high winds. Extrinsic factors include ever-changing purchase restrictions driven by national security policy decisions and onerous flight rules determined by weather conditions, dynamic airspace restrictions, and increasingly complicated device identification, registration, and licensing rules. Ground rovers are not as limited by these factors, allowing them to attach more sensors, more flexible power systems, fewer bureaucratic restrictions, and lower ground sample distance (higher resolution) owing to their ground proximity.

References

Akiva, P., Dana, K., Oudemans, P., and Mars, M. (2020). Finding Berries: Segmentation and Counting of Cranberries using Point Supervision and Shape Priors., in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, (Seattle, WA, USA: IEEE), 219–228. doi: 10.1109/CVPRW50498.2020.00033

Akiva, P., Planche, B., Roy, A., Oudemans, P., and Dana, K. (2022). Vision on the bog: Cranberry crop risk evaluation with deep learning. *Comput. Electron. Agric.* 203, 107444. doi: 10.1016/j.compag.2022.107444

Alston, Julian M., Medellín-Azuara, Josué, and Saitone, Tina L. (2014). Economic Impact of the North American Cranberry Industry. Cranberry Marketing Committee. Available at: https://www.bccranberries.com/pdfs/Economic_Impact_of_the_NA_Cranberry_Industry_August2014.pdf

Andersson, S., and Prager, M. (2006). "Quantifying Colors," in *Bird Coloration, Volume 1*, eds. G. E. Hill and K. J. McGraw (Harvard University Press), 41–89. doi: 10.2307/j.ctv22jnscm.5

Araus, J. L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S., and Cairns, J. E. (2018). Translating High-Throughput Phenotyping into Genetic Gain. *Trends Plant Sci.* 23, 451–466. doi: 10.1016/j.tplants.2018.02.001

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: Learning from the last 20 years. *Crop Sci.* 48, 1649–1664. doi: 10.2135/cropsci2008.03.0131

Bolivar-Medina, J. L., Zalapa, J. E., Atucha, A., and Patterson, S. E. (2019). Relationship between alternate bearing and apical bud development in cranberry (*Vaccinium macrocarpon*). *Botany* 97, 101–111. doi: 10.1139/cjb-2018-0058

Carter, G. A., and Knapp, A. K. (2001). Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. *Am. J. Bot.* 88, 677–684. doi: 10.2307/2657068

Chandler, F., Wilcox, R., Bain, H., Bergman, H., and Dermen, H. (1947). Cranberry breeding investigation of the US Dept. of Agriculture. *Cranberries* 12, 6–9.

Covarrubias-Pazaran, G., Diaz-Garcia, L., Schlautman, B., Deutsch, J., Salazar, W., Hernandez-Ochoa, M., 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. doi: 10.1186/s12864-016-2802-3

Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., et al. (2018). Multivariate GBLUP Improves Accuracy of Genomic Selection for Yield and Fruit Weight in Biparental Populations of *Vaccinium macrocarpon* Ait. *Front. Plant Sci.* 9, 1–13. doi: 10.3389/fpls.2018.01310

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

Del Valle, J. C., Gallardo‐López, A., Buide, M. L., Whittall, J. B., and Narbona, E. (2018). Digital photography provides a fast, reliable, and noninvasive method to estimate anthocyanin pigment concentration in reproductive and vegetative plant tissues. *Ecol. Evol.* 8, 3064–3076. doi: 10.1002/ece3.3804

Desta, Z. A., and Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19, 592–601. doi: 10.1016/j.tplants.2014.05.006

Diaz-Garcia, L., Covarrubias-Pazaran, G., Johnson-Cicalese, J., Vorsa, N., and Zalapa, J. (2020). Genotyping-by-Sequencing Identifies Historical Breeding Stages of the Recently Domesticated American Cranberry. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.607770

Diaz-Garcia, L., Covarrubias-Pazaran, G., Schlautman, B., Grygleski, E., and Zalapa, J. (2018a). Image-based phenotyping for identification of QTL determining fruit shape and size in American cranberry (*Vaccinium macrocarpon* L.). *PeerJ* 6, e5461. doi: 10.7717/peerj.5461

Diaz-Garcia, L., Covarrubias-Pazaran, G., Schlautman, B., and Zalapa, J. (2016). GiNA, an Efficient and High-Throughput Software for Horticultural Phenotyping. *PLOS ONE* 11, e0160439. doi: 10.1371/journal.pone.0160439

Diaz-Garcia, L., Garcia-Ortega, L. F., González-Rodríguez, M., Delaye, L., Iorizzo, M., and Zalapa, J. (2021). Chromosome-Level Genome Assembly of the American Cranberry (*Vaccinium macrocarpon* Ait.) and Its Wild Relative *Vaccinium microcarpum*. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.633310

Diaz-Garcia, L., Rodriguez-Bonilla, L., Smith, T., and Zalapa, J. (2019). Pacbio Sequencing Reveals Identical Organelle Genomes Between American Cranberry (*Vaccinium macrocarpon* Ait.) and a Wild Relative. *Genes* 10, 1–15. doi: 10.3390/genes10040291

Diaz-Garcia, L., Schlautman, B., Covarrubias-Pazaran, G., Maule, A., Johnson-Cicalese, J., Grygleski, E., et al. (2018b). Massive phenotyping of multiple cranberry populations reveals novel QTLs for fruit anthocyanin content and other important chemical traits. *Mol. Genet. Genomics* 293, 1379–1392. doi: 10.1007/s00438-018-1464-z

Eck, P. (1990). *The American cranberry*. New Brunswick : Rutgers University Press, [1990] ©1990. Available at: https://search.library.wisc.edu/catalog/999621210702121

Fajardo, D., Schlautman, B., Steffan, S., Polashock, J., Vorsa, N., and Zalapa, J. (2014). The American cranberry mitochondrial genome reveals the presence of selenocysteine (tRNA-Sec and SECIS) insertion machinery in land plants. *Gene* 536, 336–43. doi: 10.1016/j.gene.2013.11.104

Fajardo, D., Senalik, D., Ames, M., Zhu, H., Steffan, S. a., Harbut, R., et al. (2012). Complete plastid genome sequence of Vaccinium macrocarpon: structure, gene content, and rearrangements revealed by next generation sequencing. *Tree Genet. Genomes* 9, 489–498. doi: 10.1007/s11295-012-0573-9

Ferrão, L. F. V., Amadeu, R. R., Benevenuto, J., De Bem Oliveira, I., and Munoz, P. R. (2021). Genomic Selection in an Outcrossing Autotetraploid Fruit Crop: Lessons From Blueberry Breeding. *Front. Plant Sci.* 12, 676326. doi: 10.3389/fpls.2021.676326

Georgi, L., Johnson-Cicalese, J., Honig, J., Das, S. P., Rajah, V. D., Bhattacharya, D., 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. doi: 10.1007/s00122-012-2010-8

Hall, C. J. (1941). Cape Cod Winds Gave Idea of Cultivation of the Cranberry. *Cranberries -- Natl. Cranberry Mag.* 6, 4.

Hall, C. J. (1952). I Put by My Chart and Glass, Took to Raising Cranberry Sass. *Cranberries -- Natl. Cranberry Mag.* 17, 10–16.

Hawkesford, M. J., and Lorence, A. (2017). Plant phenotyping: increasing throughput and precision at multiple scales. *Funct. Plant Biol.* 44, v. doi: 10.1071/FPv44n1_FO

Huang, Y., Liu, N., Wagner Hokanson, E., Hansen, N., and Townsend, P. A. (2024). Exploring the potential of multi-source satellite remote sensing in monitoring crop nutrient status: A multi-year case study of cranberries in Wisconsin, USA. *Int. J. Appl. Earth Obs. Geoinformation* 132, 104063. doi: 10.1016/j.jag.2024.104063

Johnson, F., Lowry, J., Dana, K., and Oudemans, P. (2023). Vision-Based Cranberry Crop Ripening Assessment. doi: 10.48550/arXiv.2309.00028

Kawash, J., Colt, K., Hartwick, N. T., Abramson, B. W., Vorsa, N., Polashock, J. J., et al. (2022). Contrasting a reference cranberry genome to a crop wild relative provides insights into adaptation, domestication, and breeding. *PLOS ONE* 17, e0264966. doi: 10.1371/journal.pone.0264966

Liu, N., Wagner Hokanson, E., Hansen, N., and Townsend, P. A. (2023). Multi-year hyperspectral remote sensing of a comprehensive set of crop foliar nutrients in cranberries. *ISPRS J. Photogramm. Remote Sens.* 205, 135–146. doi: 10.1016/j.isprsjprs.2023.10.003

Loarca, J., Wiesner-Hanks, T., Lopez-Moreno, H., Maule, A. F., Liou, M., Torres-Meraz, M. A., et al. (2024). BerryPortraits: Phenotyping Of Ripening Traits cranberry (*Vaccinium macrocarpon* Ait.) with YOLOv8. *Plant Methods* 20, 172. doi: 10.1186/s13007-024-01285-1

Mahlein, A.-K. (2016). Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. *Plant Dis.* 100, 241–251. doi: 10.1094/PDIS-03-15-0340-FE

Maule, A. F., Loarca, J., Diaz-Garcia, L., Lopez-Moreno, H., Johnson-Cicalese, J., Vorsa, N., et al. (2024). Of buds and bits: a meta-QTL study identifies stable QTL for berry quality and yield traits in cranberry mapping populations (*Vaccinium macrocarpon* Ait.). *Front. Plant Sci.* 15, 1294570. doi: 10.3389/fpls.2024.1294570

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157, 1819–1829. doi: 10.1093/genetics/157.4.1819

Meyer, G. E., and Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* 63, 282–293. doi: 10.1016/j.compag.2008.03.009

Neyhart, J. L., Kantar, M. B., Zalapa, J., and Vorsa, N. (2022). Genomic-environmental associations in wild cranberry (*Vaccinium macrocarpon* Ait.). *G3 GenesGenomesGenetics* 12, jkac203. doi: 10.1093/g3journal/jkac203

Norwood, J. W. (1936). “Cranberry Eater”-an American King. *Cranberries -- Natl. Cranberry Mag.* 1, 7.

Pérez, A. J., López, F., Benlloch, J. V., and Christensen, S. (2000). Colour and shape analysis techniques for weed detection in cereal fields. *Comput. Electron. Agric.* 25, 197–212. doi: 10.1016/S0168-1699(99)00068-X

Schlautman, B., Covarrubias-Pazaran, G., Diaz-Garcia, L. A., Johnson-Cicalese, J., Iorrizo, M., Rodriguez-Bonilla, L., et al. (2015). Development of a high-density cranberry SSR linkage map for comparative genetic analysis and trait detection. *Mol. Breed.* 35, 177. doi: 10.1007/s11032-015-0367-5

Schlautman, B., Covarrubias-Pazaran, G., Diaz-Garcia, L., Iorizzo, M., Polashock, J., Grygleski, E., 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 GenesGenomesGenetics* 7, 1177–1189. doi: 10.1534/g3.116.037556

Sims, D. A., and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81, 337–354. doi: 10.1016/S0034-4257(02)00010-X

Song, P., Wang, J., Guo, X., Yang, W., and Zhao, C. (2021). High-throughput phenotyping: Breaking through the bottleneck in future crop breeding. *Crop J.* 9, 633–645. doi: 10.1016/j.cj.2021.03.015

Sun, H., Liu, Y., Gai, Y., Geng, J., Chen, L., Liu, H., et al. (2015). De novo sequencing and analysis of the cranberry fruit transcriptome to identify putative genes involved in flavonoid biosynthesis, transport and regulation. *BMC Genomics* 16, 652. doi: 10.1186/s12864-015-1842-4

United Nations FAO (2023). FAO Statistical Databases. Available at: https://www.fao.org/faostat/en/#home (Accessed March 27, 2025).

USDA National Agricultural Statistics Service (2023). Survey of Agriculture. Available at: https://quickstats.nass.usda.gov/ (Accessed April 2, 2025).

Vander Kloet, S. P. (1988). *The genus Vaccinium in North America*. [Ottawa] : Research Branch, Agriculture Canada, 1988. Available at: https://search.library.wisc.edu/catalog/999600224502121

Vander Kloet, S. P., and Avery, T. S. (2010). *Vaccinium* on the Edge. *Edinb. J. Bot.* 67, 7. doi: 10.1017/S0960428609990199

Vorsa, N., and Zalapa, J. (2019). "Domestication, Genetics, and Genomics of the American Cranberry," in *Plant Breeding Reviews*, (John Wiley & Sons, Ltd), 279–315. doi: https://doi.org/10.1002/9781119616801.ch8

Wagner, W. L., Herbst, D. R., and Sohmer, S. H. (1990). *Manual of the flowering plants of Hawaii.* Honolulu: University of Hawaii Press : Bishop Museum Press.

Weiss, M., Jacob, F., and Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* 236, 111402. doi: 10.1016/j.rse.2019.111402

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., and Mortensen, D. A. (1993). Plant species identification, size, and enumeration using machine vision techniques on near-binary images., eds. J. A. DeShazer and G. E. Meyer (Boston, MA), 208–219. doi: 10.1117/12.144030

Workmaster, B. A. A. (2001). Cold hardiness, ice nucleation, and growth modeling in the cranberry plant. 2001. Available at: https://search.library.wisc.edu/catalog/999926188702121

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., et al. (2017). Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Front. Plant Sci.* 8, 1111. doi: 10.3389/fpls.2017.01111

Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., et al. (2019). Crop Phenomics: Current Status and Perspectives. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00714

frontiers | Frontiers in Plant Science

TYPE Original
Research PUBLISHED
17 September 2024
DOI
10.3389/fpls.2024.129
4570

OPEN ACCESS

EDITED BY

Patrick Vincourt,

Institut National de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE), France
REVIEWED BY

Kerstin Neumann,

Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany
Javaid Akhter Bhat,

Nanjing Agricultural University, China Igor Pacheco,
University of Chile, Chile

*CORRESPONDENCE

Juan E. Zalapa ✉ jezalapa@wisc.edu; ✉ juan.zalapa@usda.gov Andrew F. Maule ✉ maule2@wisc.edu

RECEIVED 14 September 2023

ACCEPTED 31 July 2024

PUBLISHED 17 September 2024

CITATION

Maule AF, Loarca J, Diaz-Garcia L, Lopez- Moreno H, Johnson-Cicalese J, Vorsa N, Iorizzo M, Neyhart JL and Zalapa JE (2024) Of buds and bits: a meta-QTL study identifies stable QTL for berry quality and yield traits in cranberry mapping populations (*Vaccinium macrocarpon* Ait.).
Front. Plant Sci. 15:1294570. doi: 10.3389/fpls.2024.1294570

Chapter II Of buds and bits: a meta-QTL study identifies stable QTL for berry quality and yield traits in cranberry mapping populations (*Vaccinium macrocarpon* Ait.)

Andrew F. Maule [1,2]*, Jenyne Loarca [1,2], Luis Diaz-Garcia [3], Hector Lopez-Moreno[1,2], Jennifer Johnson-Cicalese [4,5], Nicholi Vorsa [4,5], Massimo Iorizzo[6,7], Jeffrey L. Neyhart[8] and Juan E. Zalapa [2]*

[1]Department of Plant and Agroecosystem Sciences, University of Wisconsin-Madison, Madison, WI, United States, [2]Vegetable Crops Research Unit, United States Department of Agriculture - Agricultural Research Service, Madison, WI, United States, [3]Department of Viticulture and Enology, University of California, Davis, Davis, CA, United States, [4]P.E. Marucci Center for Blueberry and Cranberry Research and Extension Center, Rutgers University, Chatsworth, NJ, United States, [5]Department of Plant Biology, School of Environmental and Biological Sciences, Rutgers University, New Brunswick, NJ, United States, [6]Department of Horticultural Science, North Carolina State University, Raleigh, NC, United States, [7]Plants for Human Health Institute, North Carolina State University, Raleigh, NC, United States, [8]Genetic Improvement for Fruits & Vegetables Laboratory, United States Department of Agriculture-Agricultural Research Service, Chatsworth, NJ, United States

Introduction: For nearly two centuries, cranberry (*Vaccinium macrocarpon* Ait.) breeders have improved fruit quality and yield by selecting traits on fruiting stems, termed "reproductive uprights." Crop improvement is accelerating rapidly in contemporary breeding programs due to modern genetic tools and high- throughput phenotyping methods, improving selection efficiency and accuracy.

Methods: We conducted genotypic evaluation on 29 primary traits encompassing fruit quality, yield, and chemical composition in two full-sib cranberry breeding populations—*CNJ02* (*n* = 168) and *CNJ04* (*n* = 67)—over 3 years. Genetic characterization was further performed on 11 secondary traits derived from these primary traits.

Results: For *CNJ02*, 170 major quantitative trait loci (QTL; $R^2 \geq 0.10$) were found with interval mapping, 150 major QTL were found with model mapping, and 9 QTL were found to be stable across multiple years. In *CNJ04*, 69 major QTL were found with interval mapping, 81 major QTL were found with model mapping, and 4 QTL were found to be stable across multiple years. Meta-QTL represent stable genomic regions consistent across multiple years, populations, studies, or traits. Seven multi-trait meta-QTL were found in *CNJ02*, one in *CNJ04*, and one in the combined analysis of both populations. A total of 22 meta-QTL were identified in cross-study, cross-population analysis using digital traits for berry shape and size (8 meta-QTL), digital images for berry color (2 meta-QTL), and three-study cross-analysis (12 meta-QTL).

Discussion: Together, these meta-QTL anchor high-throughput fruit quality phenotyping techniques to traditional phenotyping methods, validating state- of-the-art methods in cranberry phenotyping that will improve breeding accuracy, efficiency, and genetic gain in this globally significant fruit crop.

KEYWORDS

American cranberry, QTL, meta-QTL, BLUP, phenotyping, perennial crops, fruit breeding

Introduction

The "American" cranberry (*Vaccinium macrocarpon* Ait.) is part of a rich and diverse genus with over 500 species adapted to live in marginal habitats—thriving in acidic, peaty bogs, on the rims of sulfur-belching volcanoes, and as epiphytes in the upper story of forests (Vander Kloet, 1988; Vander Kloet and Avery, 2010). Prior to its breeding in the past two centuries, cranberry had been (and continues to be) a cultural, economic, and culinary facet of some indigenous peoples (especially those from modern-day "North America") for several millennia. The development and advancement of *Vaccinium* germplasm for wide commercial use offers a chance to expand agriculture beyond fast-dwindling arable land, offering opportunities to expand nutritional diversity to the human diet.

Cranberry is one of the most important commercial species in the *Vaccinium* genus, along with blueberries, bilberries, and lingonberries (Chandler et al., 1947; Vorsa and Zalapa, 2019). In 2022, global yield of cranberries was around 600,000 metric tonnes, with the United States (USA) producing just over 60% of global yield, followed by Canada at around 35% and Chile (extrapolated) at just under 4% (FAOSTAT, 2022). The estimated raw production value of the USA 2022 cranberry harvest was around $304 million USD (USDA NASS, 2022), not considering value-added products (Alston et al., 2014). Despite its economic importance and nearly a century of breeding, most planted cultivars are limited to only one or two generations beyond wild germplasm, with only recent introduction of third-generation cultivars in managed marshes (Diaz-Garcia et al., 2020). Limited breeding progress stems from large planting space requirements, long establishment times (3–4 years), and extended evaluation times of 6–8 years before commercial release (Vorsa and Zalapa, 2019).

Cranberries exhibit unique plant architecture among woody fruit crops, with a growth habit that is characterized by low-growing vines producing short vertical lateral branches known as “uprights.” These uprights are classified into two types: vegetative uprights and reproductive uprights. Vegetative uprights develop apical buds with only vegetative meristems (vegetative buds) and thus only produce leaves. Reproductive uprights develop both vegetative and floral meristems (reproductive buds) and thus produce leaves while also sustaining the development of flowers and fruit. Apical bud induction and differentiation is determined by both management practices and by genetics (Bolivar-Medina et al., 2019). Cranberry yield traits, such as fruit size, quality, shape, and number, were traditionally measured on these flowering vines on a per-upright basis. As such, “reproductive upright traits” have been the traditional target of selection to phenotypically improve berry quality and yield. However, given the substantial cost and time required to accurately phenotype these traits, convention is transitioning toward measuring yield and chemistry on a plot-level (or per-unit-area) basis (Vorsa and Zalapa, 2019).

The past 7 years have shown incredible progress in cranberry molecular resource development and utilization (Covarrubias-Pazaran et al., 2018; Vorsa and Zalapa, 2019). In cranberry, advancement and cost-reduction in high-throughput genome sequencing technologies have enabled assembly of plastid and mitochondrial genomes (Fajardo et al., 2013; Diaz-Garcia et al., 2019 , 2014), construction of two high-quality genome (Diaz-Garcia et al., 2021; Kawash et al., 2022), *de-novo* sequencing of cranberry transcriptomes (Georgi et al., 2013; Sun et al., 2015; Diaz-Garcia et al., 2021), linkage map development for agronomic traits (Georgi et al., 2013; Schlautman et al., 2015, 2017a; Covarrubias-Pazaran et al., 2016; Daverdin et al., 2017), and pilot applications in association mapping and genomic selection (Covarrubias-Pazaran et al., 2018; Diaz-Garcia et al., 2020; Neyhart et al., 2022).

These technologies have enabled several marker-trait association studies on an array of commercially important cranberry traits. For example, Georgi et al. (2013) found evidence of several quantitative trait loci (QTL) in four related populations for *field fruit-rot resistance*, *titratable acidity* (*TA*), f*ruit weight*, and *sound fruit yield* (*SFY*). Schlautman et al. (2015) constructed a high- density microsatellite linkage map in a breeding population and discovered QTL for mean fruit weight (*MFW*), total yield (*TY*), and biennial bearing index (*BBI*). Daverdin et al. (2017) generated high-density linkage maps in four diverse populations selected to demonstrate high levels of segregation for *field fruit-rot resistance* and found 15 QTL across all populations while demonstrating that yield traits segregate independently of *field fruit-rot resistance*. Image analysis and wet chemistry techniques have demonstrated the power to rapidly generate fruit shape, size, and color descriptors, which are important proxies for fruit quality and yield, and have identified QTL for anthocyanin production, *MFW*, and shape descriptors that are important for sweetened dried cranberry (SDC) production (Covarrubias-Pazaran et al., 2018; Diaz-Garcia et al., 2018a, 2018b). Finally, pilot applications in association mapping and genomic selection have been recently conducted in cranberry (Covarrubias-Pazaran et al., 2018; Diaz-Garcia et al., 2020; Neyhart et al., 2022).

Despite these significant advancements in molecular methods and genetic tools, no study has yet compared or validated trait phenotypes derived from traditional versus contemporary methods. We conducted QTL analysis on trait phenotypes obtained from contemporary phenotyping methods developed in the last three decades and on traditional phenotyping techniques that have been implemented by cranberry breeders in the last two centuries (Franklin et al., 1958; Eck, 1990; Vorsa and Zalapa, 2019). Many of the traits used in the current study are relevant to modern high-throughput phenomic tools, and others are traits traditionally collected per upright to make selection decisions (Diaz-Garcia et al., 2016). Leveraging advanced genomic and phenomic tools will accelerate genetic gain in cranberry breeding programs and improve understanding of which traditional traits are still useful and relevant in the modern context. This paper is the first to report correlations, heritabilities, and QTL based on both traditional and modern phenotyping methods. We propose that modern phenotyping methods are at least as accurate as traditional phenotyping methods, if not more so, for detecting heritable variation in reproductive upright traits, leading to identification of multi-year QTL for fruit yield and quality. We further expect that QTL identified across years and across studies (meta-QTL) will offer robust opportunities to perform marker-assisted selection for cranberry fruit quality and yield traits. Finally, validation of modern methods against traditional methods will open the doors to more efficient selection for these critically important traits.

Materials and methods

Plant material and traits collected

The two full-sib populations used in this study, *CNJ02* (*n* = 168) and *CNJ04* (*n* = 67), represent the most highly studied populations in the cranberry fruit yield and quality scientific literature (Schlautman et al., 2015, 2017b; Covarrubias-Pazaran et al., 2018; Diaz-Garcia et al., 2018b; Vorsa and Zalapa, 2019; Diaz-Garcia et al., 2021). These populations were established in 2007 in 2.3 m^2 field plots at Rutgers'

P.E. Marucci Center for Blueberry and Cranberry Research and Extension, Chatsworth, NJ (Vorsa and Johnson-Cicalese, 2012), with experimental design described in previous studies (Schlautman et al., 2015; Covarrubias-Pazaran et al., 2016; Schlautman et al., 2017b; Diaz-Garcia et al., 2018b). *CNJ02* was derived from a cross between seed-bearing parent, *CNJ97_105_3* (cv. *Mullica Queen*®) and pollen-donating parent, *NJS98_23* (cv. *Crimson Queen*®); *CNJ04* was derived from reciprocal crosses between *CNJ97_105_3* (cv. *Mullica Queen*®) and cv. *Stevens*. Naming conventions of individual genotypes for populations *CNJ02* and *CNJ04* are as follows: *CNJ<YY>_<CN>_<GID>*, where *<YY>* is a two-digit designation indicating the year of the cross (2002 or 2004, respectively), *<CN>* is a number indicating the cross number (CN) for that year, *<GID>* is a genotype identifier (GID), and the underscore separates identifiers in the string. For example, *CNJ02_1_38* is a progeny from *CNJ02*, derived from the first cross (CN = 1) made in 2002, with genotype individual number 38 of that cross (GID = 38).

From 2011 to 2014, reproductive uprights were collected from field plots, 10 uprights per genotype. Traits from *CNJ02* and *CNJ04* each were recorded over 3 years, with *CNJ02* sampled for the years 2011–2013 and *CNJ04* sampled for 2011, 2012, and 2014. The traits measured largely comprise those traditionally considered to be commercially important in cranberries (Vorsa and Zalapa, *personal communication*), along with new fruit quality traits. Traits were loosely categorized into attributes measured directly, and traits features derived from other attributes. Traits measured directly include the categories *Upright Traits*, *Largest Berry Traits*, and *Plot Traits*; derived traits fall into the category *Berry Shape Chimera Parameters* calculated from berry shape composite representations (Supplementary Table S1). Traits will frequently be referred to by both their full names and acronyms to facilitate figure interpretation and cross-referencing with the manuscript. Trait acronyms will be defined once again in each new section or paragraph.

Plot Traits were sampled the same years as the uprights and are based on 0.09 m^2 plot samples of fruit. These include *TY*, *SFY*, *mean fruit mass* (*MFM*), *percent fruit rot* (*PFR*), *total anthocyanins* (*Tacy*), *soluble solids* (*Brix*), *TA*, and *proanthocyanins* (*PACs*) (Supplementary Table S1) (Diaz-Garcia et al., 2018b).

Upright Traits, which include *Largest Berry Traits*, were recorded per upright, and their abbreviations are always prefixed with the letter "*U*" (Supplementary Table S1). *Upright Traits* include *total berry mass* (*UTBM*), *length* (*UL*), *secondary length* (*USL*), *dry leaf mass* (*UDM*), *rebud* (*URB*), *mean fruit mass* (*UMFM*), *number of pedicels* (*UNPs*), *number of pedicels without berries* (*UN0*), *number of pedicels with mature berries* (*UNBs*), *number of pedicels with aborted flowers* (*UNAFs*), and *number of pedicels with aborted berries* (*UNABs*).

Largest Berry Traits were measured on the largest berry per upright (10 berries per genotype). *Largest Berry* yield traits include *berry length* (UBL), *berry width* (*UBW), berry length:width ratio* (*ULvW*), and *berry mass* (*UBM*). *Largest Berry* quality traits include *berry shape* (*UBS)*, *number of seeds* (*UNSs*), *calyx diameter* (*UCD*), *calyx lobe fold pattern* (*UCLP)*, *calyx lobe size* (*UCLS)*, *calyx end shape* (*UCES)*, *berry pedicel end shape* (UBES), and *berry bloom level* (*UBBL*)—a measure of berry epicuticular wax levels.

Berry Shape Chimera Properties are derived from the composite representations, or chimeras, of five berry shape categories (Figure II-1). The berry chimera is rendered from the average shape of the 10 largest berries (Figure II-2). These parameters were extracted and mapped to provide a data integrity benchmark of the subjective, categorical berry shape traits against their per-upright berry trait analogs (Supplementary Tables S1 and the section *Trait evaluation and transformation*).

Trait evaluation and transformation

All categorical traits were transformed to numeric, discrete values (except for largest berry shape). Higher values represent more favorable characteristics. Supplementary Table S3 displays the conversions of categorical traits to numeric values.

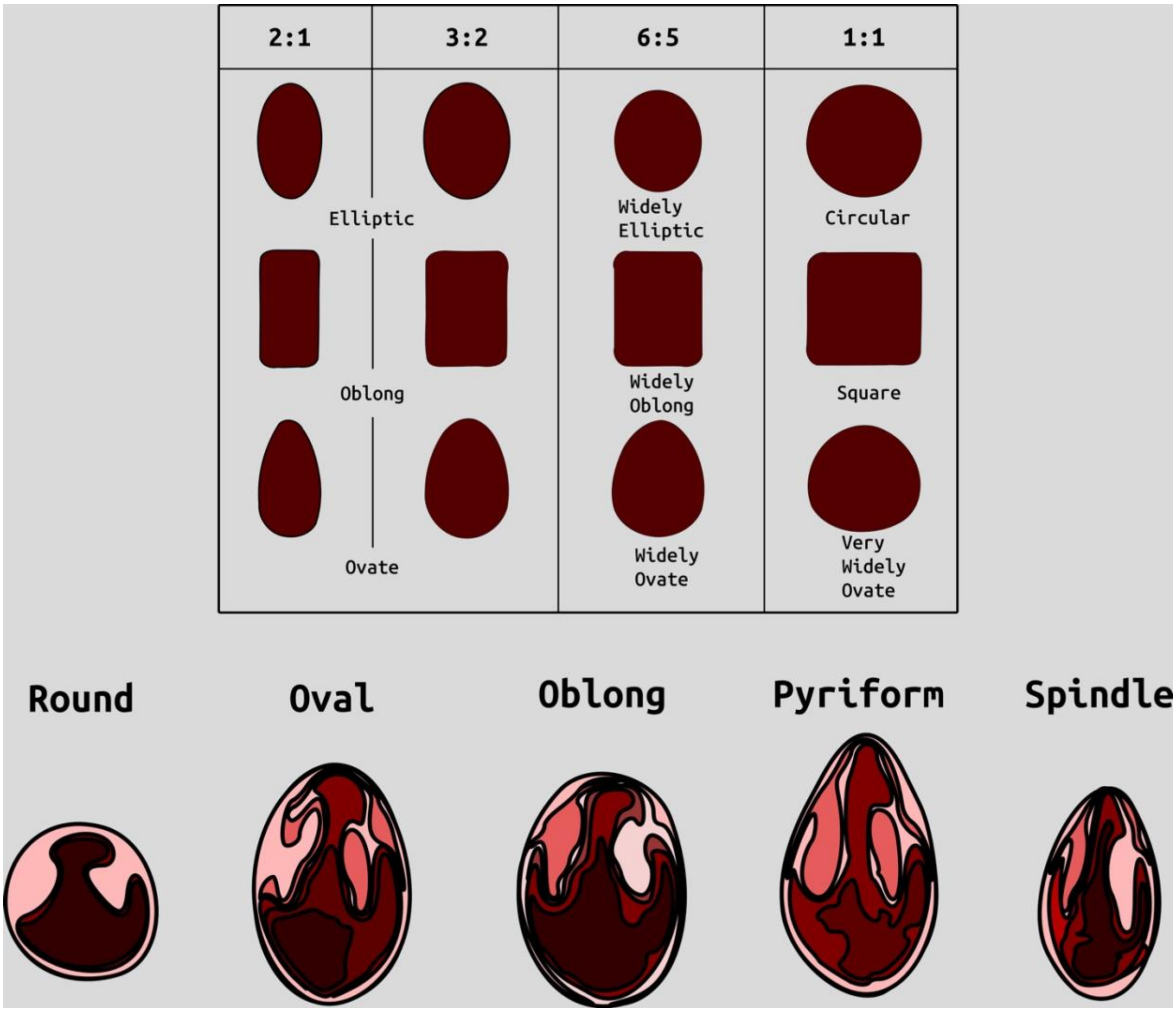


Figure II-1. Examples of categories of berry shape parameters, used to help classify cranberry shape (Franklin *et al.*, 1958).

Largest berry shape was quantified by replacing scored shape categories with their canonical digital image shapes using the template images provided to trait evaluators (Figure II-1). Digital image shape descriptors derived from the representative berry chimera are listed in Supplementary Table S1 under category *Berry Shape Chimera Parameters*. These images were initially converted to binary images, normalized by centering all representations, and scaling so that all shape templates have the same area. Thereafter, a chimeric berry representation was derived for each genotype year by combining and thresholding the differential distance transformed representations of each corresponding upright's normalized berry shape (Figure II-2). A distance transform representation of a binary image is the Euclidean distance of any foreground pixel to its nearest background pixel. The differential distance transformation is the difference between the distance transformation of the normalized binary berry template image and the same image inverted. Positive values indicate foreground pixels, with higher values indicating the centroid of binary images. By combining the differential distance transforms and thresholding on positive values, a chimeric binary berry image representation per genotype year can be generated. The berry chimera therefore represents a synthesized berry image derived from the composition of multiple categorical berry shapes (10 berry shape classes per genotype). Shape descriptors are calculated from the synthesized berry chimera, allowing for quantitative analysis of shape traits for mapping. Traits *upright chimera unsigned manhattan chain code–X-axis* (*UKUX)* and *upright chimera unsigned manhattan chain code–Y-axis* (*UKUY)* are *log* 10 derivations of unsigned Manhattan chain codes (UMCC) (Žalik et al., 2016). A chain code is a numerical representation describing the contour path of an object. UMCCs are one of many chain codes used to describe contour shapes and were chosen here for their highly compressible representation. Other chimera shape descriptors outlined in Supplementary Table S1 include *chimera shape eccentricity* (*UKEC*), *chimera length:width ratio* (*UKLvW*), *chimera tortuosity* (*UKTO*), and *chimera solidity* (UKSO). *Eccentricity* (*EC*) of a closed contour is a mathematical descriptor for the curvature of an ellipse, with zero indicating a perfect circle and values closer to one signifying a higher length:width ratio. Tortuosity describes the "*waviness*" of an object's contour and, in this case, a slope chain code (SCC) method was applied to generate this tortuosity value (Bribiesca, 2013). Solidity describes the shape density relative to its convex hull. Lower solidity values indicate more *waviness* in the berry contour, while higher values have smoother contour curvature.

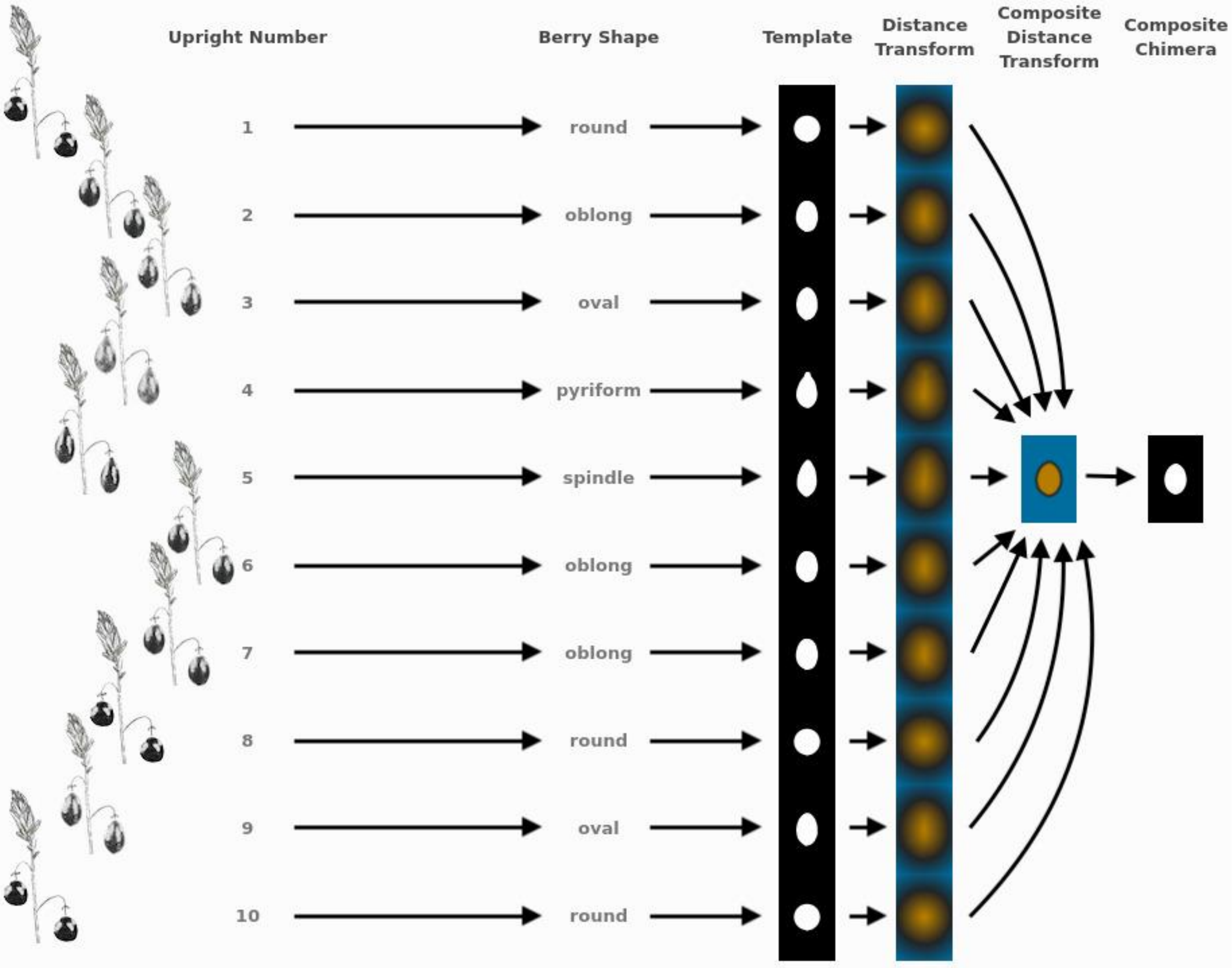


Figure II-2. Example showing the methodology for generating representative genotype shape, or berry chimera, from ten-upright samples. This berry chimera is subsequently used as an image to generate digital image processing shape descriptors used in creating quantifying features for QTL mapping.

All traits were curated by removing entries marked as rotten, and outliers were detected and trimmed with the outlierTest() function of the car package (RRID: SCR_022137) with a default cutoff of $p < 0.05$ using a linear model of the trait regressed on population, genotype, and year (Fox and Weisberg, 2019). Additional observations were culled if they exceeded three standard deviation units from the mean under a Gaussian standard distribution. Subsequent analysis and trait mapping were applied to all traits based on the mean trait values across the ten sampled uprights, for each genotype year. From these upright means, Pearson correlation coefficients were calculated for all traits and averaged across months and all sampled years. Correlation heatmap plots were generated using the corrplot package (RRID: SCR_023081), with statistical *p* values calculated using the ggcorrplot package (Kassambara, 2019; Wei and Simko, 2021). Trait correlations were partitioned using hierarchical agglomerative clustering cut at an absolute value correlation tree height equal to 0.6. These partitions form clusters, or cliques, of traits, delineated by heavy black lines in the generated heatmaps. Only non-singleton clusters are reported in the results.

Linkage maps

Linkage maps used in this study were previously created from a combination of robust single sequence repeat (SSR) markers and genotype-by-sequencing (GBS) single-nucleotide polymorphic (SNP) markers (Schlautman et al., 2017b). SSR marker data generated for *CNJ02* (541 SSRs) and *GRYG* (189 SSRs) were discovered and curated as described in Schlautman et al. (2015) and Covarrubias-Pazaran et al. (2016). For GBS markers, genomic DNA was extracted from flash frozen leaf tissue and EcoT221-digested DNA fragments were uniquely barcoded for all progeny and parents of *CNJ02* and *CNJ04* using the approach described by Elshire et al. (2011). These fragments were sequenced (single-end) at the Cornell University Biotechnology Resource Center Genomics Facility on an Illumina HiSeq 2000 platform. A reference Tassel GBS analysis pipeline (Bradbury et al., 2007) was used to filter and process the resulting sequence reads and call SNPs in the resulting datasets for the *CNJ02* and *CNJ04* populations using the parameters described in Schlautman et al. (2017b) and aligned to the cranberry reference genome produced by Polashock et al. (2014). SNPs with >20% missing data, minor allele frequency (MAF) of <10%, or severely distorted segregation ratios were removed. Linkage analysis on the filtered SSR and

SNP markers was performed with JoinMap v4.1, using a pseudo-testcross method, and biparental consensus linkage maps were separately generated for *CNJ02* (n = 3925) and *CNJ04* (n = 3081) (Schlautman et al., 2017b). Additionally, a composite linkage map (n = 1560) derived from three cranberry populations, *CNJ02*, *CNJ04*, and *GRYG*, was generated using a linear programming approach as described by Schlautman et al. (2017), and is available on the Genome Database for Vaccinium site under the map identifier Cranberry- Composite_map-F1 (https://www.vaccinium.org/bio_data/1070). QTL generated herein used the composite linkage map in order to facilitate cross-population comparison. Population-specific biparental consensus linkage maps were used to produce the additive genomic relationship matrices described in Equation II-1.

Estimating breeding values and heritability

Equation II-1 shows the mixed model used in estimating best linear unbiased predictors (BLUPs) within years (Henderson, 1975). The equation variables are defined as follows: y = phenotype value, μ = intercept (global mean of trait), Z_g = genotype random-effect incidence matrix, g = genotypic effects (BLUPs), Z_r = row random effect incidence matrix, r = row effect, Z_c = column random effect incidence matrix, c = column effect, Z_s = 2D-spline random effect incidence matrix, s = spline effect, ε = residuals, G = genotype variance-covariance matrix (De Los Campos et al., 2015), A = additive genomic relationship matrix (Endelman, 2011), σ_a^2 = additive genomic variance, I = identity matrix, σ_r^2 = row variance, σ_c^2 = column variance, σ_s^2 = 2D-spline variance, σ_ε^2 = residual error variance.

Equation II-1. Mixed model for estimating BLUPs within years.

$$y = \mu + Z_g g + Z_r r + Z_c c + Z_s s +,$$

$$where\ g \sim \mathcal{N}(0, G), G = A\sigma_a^2, r \sim \mathcal{N}(0, I\sigma_r^2), c \sim \mathcal{N}(0, I\sigma_c^2), s \sim \mathcal{N}(0, I\sigma_s^2), and\ \varepsilon \sim \mathcal{N}(0, I\sigma_\varepsilon^2).$$

Equation II-2 displays the across year mixed model used in estimating BLUPs. Year is modeled as a fixed effect. All symbols are the same as in Equation II-1, but with the additional term X_e = fixed-effect year incidence matrix, e = year effect, Z_{ge} = genotype-by-year random-effect incidence matrix, ge = genotype-by-year effect, and σ_{ge} = genotype-by-year variance.

Equation II-2. Mixed model for estimating BLUPs across years.

$$y = \mu + X_e e + Z_g g + Z_{ge} ge + Z_r r + Z_c c + Z_s s +,$$
$$where\ levels\ of X_e \in \{2011{,}2012{,}2013{,}2014\}, and\ g_e \sim \mathcal{N}\left(0, \sigma_{ge}^2\right).$$

The mixed models detailed Equation II-1 and Equation II-2 were fit using utilities defined in the R package (R Core Team, 2024) sommer (Covarrubias-Pazaran, 2016). The additive genomic relationship matrix *A* (also known as additive relationship matrix) is a variance-covariance matrix that was constructed using the sommer function A.mat() on biallelic markers (SNP) from each population’s respective biparental consensus linkage map. Modeling the genetic relatedness between genotypes with the *A* matrix adjusts for differences absent a structured experimental design. Spatial effects were also compensated for by modeling this variation as 2D-spline random effects and as row and column random effects. The 2D-spline effects are continuous random variables that model spatial variation that does not track first-order polynomial (straight-line) field effects or does not trace along the row and column effects. The row and column random effects models spatial variation by blocking according to the predefined row and column indices of the plots. The best linear unbiased predictors (BLUPs) (also known as genomic estimated breeding values (GEBV)), were estimated using the mmer() function of the sommer package.

Model selection was performed from a full model search on the random terms of models defined in Equation II-1 or Equation II-2 using the *Akaike Information Criteria* (AIC) (Akaike, 1974). The model with the lowest AIC was subsequently used to estimate BLUPs, balancing model complexity with parsimony. For across-year mixed model BLUP estimates, the genotype-by-year interaction effect was excluded from the model fit if it did not display significance ($p < 0.05$) using a likelihood ratio test. Random term variance estimates were used to calculate additive genomic heritability (h^2) of each trait (De Los Campos et al., 2015).

Equation II-3 displays the formula for calculating within-year narrow-sense genomic heritability, using variances estimated from fitting the mixed model (Equation II-1). Across year genomic heritability is calculated as shown in Equation II-4, using variances estimated from the mixed model (Equation II-2), with n representing the number of distinct years fit per trait.

Equation II-3. Narrow sense heritability calculation for within-year mixed model.

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\varepsilon^2}$$

Equation II-4. Narrow sense heritability calculation for across-years mixed model.

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \frac{\sigma_{ge}^2}{n} + \frac{\sigma_\varepsilon^2}{n}}$$

QTL mapping

QTL mapping was performed by using the CRAN (RRID: SCR_003005) package r/QTL (RRID: SCR_009085), a software toolkit for mapping experimental crosses (Broman et al., 2003). To infer QTLs, previously modeled genotype BLUPs were substituted in lieu of raw phenotypes in the R/qtl cross table. QTL were detected using two methods: a single-QTL interval mapping method and a model selection approach. Both methods used Haley-Knott regression to model QTL between genetic map markers (Haley and Knott, 1992). The single-QTL method uses the scanone() function, with significant QTL determined using scanone () run against 1,000 permutations of the phenotypes in order to simulate the log of odds (LOD) distribution of the NULL model. The model selection approach uses the stepwiseqtl() function, which runs a forward/backward model search algorithm by which additive and interacting terms are successively added to the model, followed by "backward" pruning of other model terms that optimizes a penalized LOD score. The penalized LOD score uses n^{th} percentile thresholds derived from running scantwo() against 1,000 permutations of the phenotypes in order to control the false positive rate at *n* percent (Broman et al., 2009). More complex models are penalized higher to reduce model overfitting. For single-QTL interval mapping, the significance threshold was set for QTL with LOD scores above the 80^{th} permutation percentile, while for model selection, penalized LOD scores were derived from thresholds determined from the 95^{th} permutation percentile. QTL inferred through statistical association mapping approaches are subsequently referred to as principal or primary QTL.

Meta-QTL

Co-located QTL, or meta-QTL, represent collections of principal marker associations that are stable across corresponding traits, years (multiple models), populations, QTL inference methods, and/or other studies. Major QTL are defined as those QTL with a mean percent marker variance explained (marker $\overline{R^2}$) greater than or equal to 10%. Within-trait major QTL found for populations *CNJ02* and CNJ04 were assembled from principal QTL that were stable in at least three of the four models (single-year models on three separate years as defined in Equation II-1 and the multi-year model defined in Equation II-2). To construct multi-trait composite meta-QTL, traits were grouped into correlated sets based on observed trait Pearson correlation blocks (Figure II-3) for this study, or similar groups of traits across studies. Multi-trait composite meta-QTL were grouped according to the categories detailed in (Table II-1, Table II-1, Supplementary Tables 1A, 1B. Within-trait major meta-QTL were assigned a separate group per trait. Using the lower and upper LOD 1.5 interval extents, all pairwise principal QTL within the same group were represented in an undirected graph structure, where nodes constitute the primary QTL and edges indicate QTL–QTL overlap. A graph was assembled for each set and trait group using the Python library NetworkX, and composite meta-QTL were synthesized from a maximal clique approximation algorithm, where each maximal clique represents a new meta-QTL (Van Rossum, 2007; Hagberg et al., 2008). For each maximal clique, the new left and right extents of the meta-QTL were taken from the maximal LOD 1.5 left extent and the minimal LOD 1.5 right extent. The meta-QTL's synthetic position was calculated as the center, or mean, of its extents. Additionally, information was recorded for each meta-QTL by compiling summary statistics such as population count, study count, year count, mapping method count, trait count, and mean marker variance explained by QTL. Other meta-QTL information compiled included the composite list of traits, populations, models (temporal consistency), and studies. Genomic plots of composite meta-QTL were generated using a customized version of the CRAN package LinkageMapView

(Ouellette et al., 2018). For each multi-trait set annotated in Table II-1, a representative linkage map of selected meta-QTL was generated. Line segments represent the maximal LOD position of the principal trait QTL, linkage map fills represent the original LOD 1.5 intervals, labels identify the respective traits of the composite, and interval bars demarcate the intersection of principal QTL 1.5 LOD intervals. Traits and intervals are colored as defined in each figure's respective legend (Figure II-4-Figure II-6). Within-trait meta-QTL for *CNJ02* and *CNJ04* are graphed similarly to multi-trait meta-QTL, but with each linkage group subdivided into four columns delineating the year models (Figure II-4).

Results

Correlations

Prominent trait correlations illuminate phenotypic relationships for population *CNJ02* (A and B) and *CNJ04* (C and D) ($p< 0.05$) (Figure II-3). The bar plots in Figure II-3B,D show the 15 top positive and 15 bottom negative inter-trait correlations per respective population. Only significant correlations ($p < 0.05$) are displayed in Figure II-3B,D. Correlated trait clusters in *CNJ02* include {*PACs*, *TY*, *SFY*}, {*MFM*, *UMFM*, *UBM*, *UTBM*}, and {UCLP, UCLS}, {UKTO, UKUY}, {biennial bearing index-total yield (BBITY), biennial bearing index-sound fruit yield (BBISFY)}, {UNAFs}, {UBL, UDM}, and {UBL, UKLvW, UKEC, ULvW, UKUX} (Figure II-3A). Pairwise positive associations evident in the *CNJ02* dataset include {*SFY* × *TY*}, {*MFM* × *UBM*}, {*MFM* × *UMFM*}, {*UBM* × *UTBM*}, and {*MFM* × *UBW*}. Negative pairwise correlations found in the *CNJ02* dataset include {*TY* × *PAC*}, {UCLP × UCLS}, {SFY × PAC}, {SFY × Tacy}, {PAC × UNABs}, {MFM × PAC}, {MFM × UNAFs}, and {MFM × UBBL}. For CNJ04, correlation clusters include the {*TY, SFY*, *UTBM*, *UNBs*} cluster and the {*MFM*, *UMFM*, *UBM*} cluster (Figure II-3C).

Heritabilities

Generally, *all-year* model trait BLUPs (Equation II-2) have larger heritabilities relative to their respective fitted model trait BLUPs for individual years. Traits with consistently high heritability in both populations are displayed in Figure II-7. Heritability values are calculated from the inter-population mean of the *all-year* model heritability (Figure II-7A), and separately, the within-year model heritability averages between *CNJ02* and *CNJ04* (Figure II-7B). *ULvW* has remarkably high heritability in both populations, followed by *UBL* and *UNSs*. *Tacy*, a desirable trait associated with fruit color, also shows evidence of high heritability. High heritabilities in *UBM*, *UMFM*, and *TY* indicate strong potential for selecting for higher yields in both populations. The fruit rot *SFY* and *PFR* exhibited low to modest heritabilities ($0.3 \leq h^2 \leq 0.6$) in both populations. Fruit quality traits such as *TA* and *Brix* exhibit very low *all-year* heritabilities in population *CNJ02* but higher heritability in *CNJ04* (Figure II-8).

Breeding value estimates

Population *CNJ02* and *CNJ04* both displayed evidence of transgressive segregation for many of the traits. Table II-2, Table II-3 show summary statistics for traits and their associated BLUPs for the *all-year* model. Rows are sorted by decreasing heritability, and only display the top 50% of models based on heritability ($h^2 \geq 0.47$) and with significant genotypic effects ($p < 0.05$). Raw trait associated statistics are subscripted with the italicized letter *r*, and BLUP associated statistics are subscripted with the italicized letter *b*. The top and bottom five representative genotypes are also presented both for recorded traits and for genotype BLUPs.

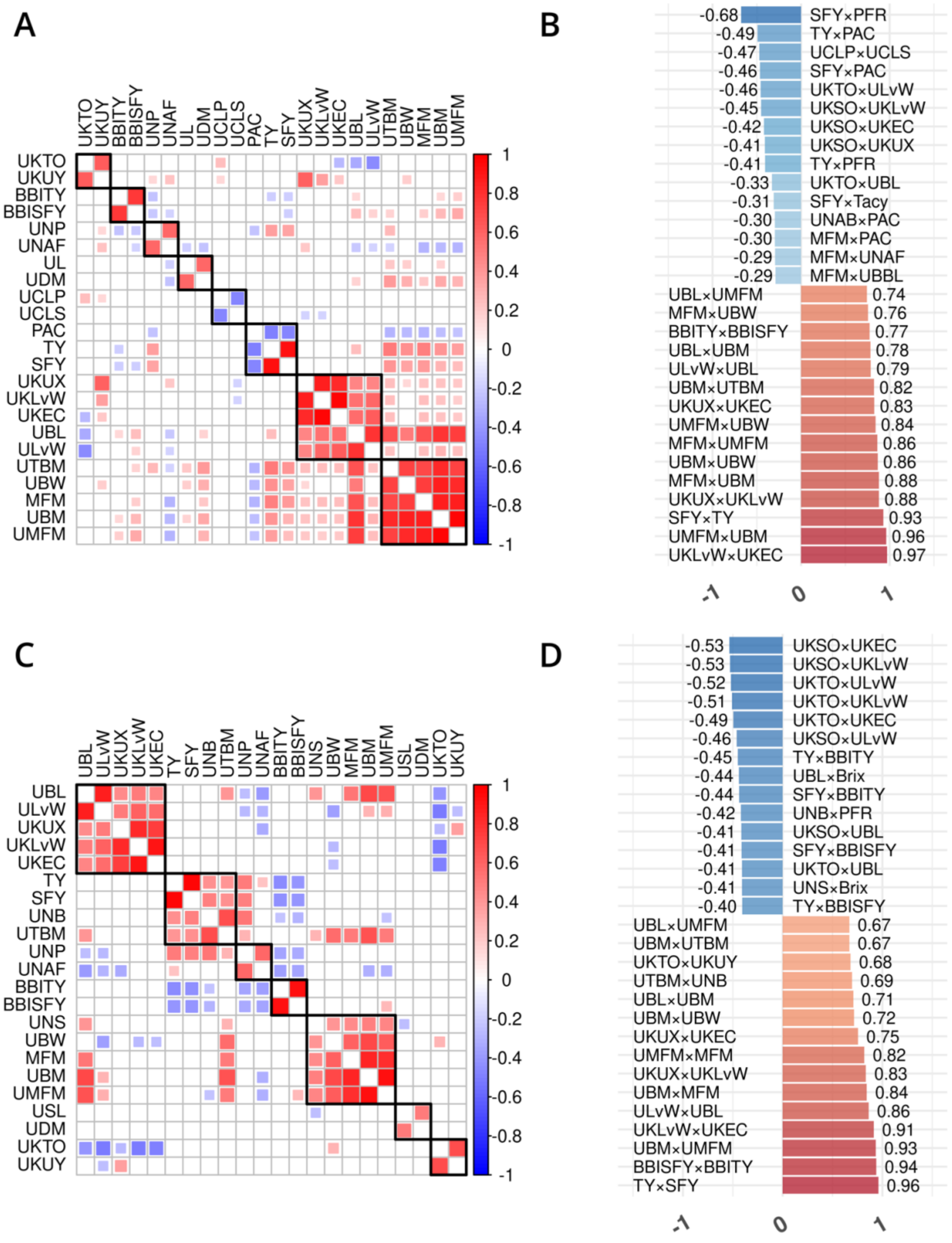


Figure II-3. Trait phenotype correlation heatmaps for traits in cranberry populations *CNJ02* (A and B) and *CNJ04* (C and D). Only traits found within non-singleton hierarchical clusters are shown in (A) and (C), where black squares delineate clusters. The 15 largest positive and 15 smallest negative correlations are shown in (B) and (D). Only significant pairwise trait correlations ($p < 0.05$) are displayed.

QTL analysis

CNJ02 population

For the *CNJ02* population, around six QTL per *trait* × *model* were found for the scanone method and 10 QTL per *trait* × *model* were found for the stepwiseqtl method. Of the 385 additive scanone QTL found in models with significant genotype effects ($p < 0.05$), 170 were major QTL (*marker PVE* ≥ *0.1*). Of the stepwiseqtl additive QTL found in models with significant genotype effects, 150 of 702 were major. The average scanone percent variance explained (*PVE*) per QTL was 9.8%, with an average model *PVE* of 71.5%. The mean stepwiseqtl QTL *PVE* was 7.2%, with an average model *PVE* of 89.9%. A full list of QTL and their effect sizes can be found in Supplementary Tables S4, S5.

A summary of stable multi-model, within-trait meta-QTL for *CNJ02* are displayed in REF _ch2_figure4 \h **Figure 4**A, Table II-4. The *Tacy* trait displays a prominent meta-QTL on linkage group 3 (position 55.8 cM) with an average paternal effect of −11 mg units per 100 mg fruit. *UBM*, *UTBM*, and *PFR* all have proximal QTL on linkage group 11 (around 30 cM) with *UBM* and *UTBM* having similar average paternal effect sizes of −0.34 g and −0.29 g. A meta-QTL for MFM (per upright) is found very close to the berry mass and upright berry mass on linkage group 11 (position 38.1 cM) with a similar paternal effect size of −0.3 g/fruit. *PFR* has around 10% decrease for both the maternal and paternal effects. A meta-QTL is found very close to the *UBM* and *UTBM* on linkage group 11 (position 38.1 cM), with a similar paternal effect size of −0.3 g/fruit. *SFY* has a meta-QTL on linkage group 12 at 40.4 cM, with a maternal effect size of 10 g/0.09 m^2 and a paternal effect size of 46 g/0.09 m^2. *UKLvW* has a modest meta-QTL on linkage groups 3 (position 30.7 cM) and 9 (position 82.7 cM). *UBL* has a meta-QTL on linkage group 10 (position 53.0 cM) with a maternal effect size of −1.8 mm and paternal effect size of 1.1 mm.

Table II-1A. Trait groupings for each set of co-located QTL in present study.

Trait Groups[a]	Filter[b]
CNJ02	$\overline{R}^2 \geq 10\%$ Year count ≥ 4
mean fruit mass (MFM, UMFM); berry width (UBW); berry mass (UBM); total berry mass (UTBM)	
total yield (TY); sound fruit yield (SFY); proanthocyanins (PAC)	
berry length (UBL); chimera eccentricity (UKEC); berry length:width (ULvW), chimera length:width UKLvW);	
upright length (UL); upright dry leaf mass (UDM)	
number of pedicels (UNP); number of pedicels with aborted flowers (UNAF)	
CNJ04	$\overline{R}^2 \geq 10\%$ Year count ≥ 4
UNP; UNAF	
upright secondary length (USL); UDM	
number of seeds (UNS); UBW, MFM, UMFM; UBM	
number of pedicels with mature berries (UNB); TY; SFY; UTBM	
UBL; ULvW, UKLvW; UKEC	
CNJ0x	$\overline{R}^2 \geq 10\%$ Year count ≥ 4 Populations ≥ 2
UNP; UNAF	
MFM, UMFM; UBM; UBW;	
TY; SFY	
UBL; ULvW, UKLvW; UKEC	

[a] Bold entries designate the separate sets used in co-location QTL analysis.

[b] Multiple filter rules are combined with a logical AND operation. Trait groupings for each set of co-located QTL in present study. Trait groups were selected based on correlated clusters or similar categories of traits. The filter column details the constraints applied to the synthetic meta-QTL that are displayed in the linkage maps. $\overline{R^2}$ is the mean percent marker variance explained across all composite QTL. Trait acronyms are provided upon first use in table and are abbreviated thereafter.

Table II-1B. Trait groupings for each set of co-located quantitative trait loci (QTL) across cranberry studies (present study included).

Trait groups[a]	Populations	Filter[b]
CNJ0x & Diaz-Garcia 2018a	*CNJ02*, *CNJ04*, *GRYG*	$\overline{R}^2 \geq 10$ % *Trait count* ≥ 3 *Study count* ≥ 2
MFM, UMFM, UBM, UTBM, BW, UBW, UBL, BL, berry area (BA)		
TY, SFY, PAC		
EC, UKEC, LvW, ULvW		
CNJ0x & Diaz-Garcia 2018b	*CNJ02*, *CNJ04*, & *GRYG*	$\overline{R}^2 \geq 10$ % *Populations* ≥ 2 *Study count* ≥ 2
MFM, UMFM, UTBM, UBW, UBM, UBL		
TY, SFY, PAC		
Berry color (BCOLOR), berry color variance (BCOLORVAR) total anthocyanin (Tacy), *total anthocyanin in September (TACY_SEP), total anthocyanin in October (TACY_OCT),* *difference in Tacy between Sept. and Oct. (TACY_DIFF)*		
CNJ0x & Schlautman 2015	*CNJ02*, *CNJ04*	$\overline{R}^2 \geq 10\%$ *Study count* ≥ 2
MFM, UMFM, UTBM, UBM, UBL, UBW		
TY, SFY, PAC		
CNJ0x & Schlautman 2015 & Diaz-Garcia 2018a	*CNJ02*, *CNJ04*, *GRYG*	$\overline{R}^2 \geq 10$ % *Trait count* ≥ 3 *Study count* ≥ 2
MFM, UMFM, UTBM, UBM, BL, UBL, BW,UBW, BA		
TY, SFY, PAC		
EC, UKEC, LvW, ULvW		

[a] Bold entries designate the separate sets used in co-location QTL analysis.

[b] Multiple filter rules are combined with a logical AND operation. Trait groupings for each set of co-located quantitative trait loci (QTL) across cranberry studies (present study included). Trait groups were selected based on correlated clusters or similar categories of traits. The filter column details the constraints applied to the synthetic meta-QTL that are displayed in the linkage maps. $\overline{R^2}$ is the mean percent marker variance explained across all composite QTL. Trait acronyms are provided upon first use in table and are abbreviated thereafter.

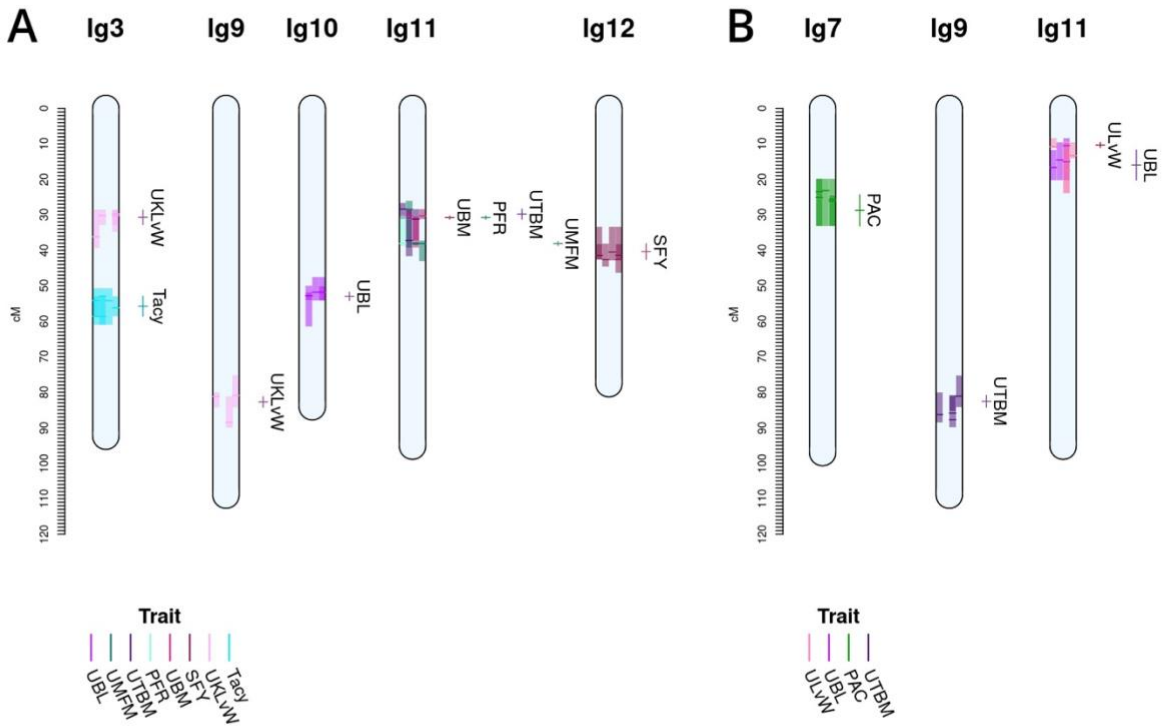


Figure II-4. Co-located, stable QTL in cranberry populations *CNJ02* (A) and *CNJ04* (B) found in at least three of four BLUP models. Only QTL with $\overline{\mathrm{R}^2} \geq 10\%$ are shown. Shaded bars represent the 1.5LOD confidence interval for the original QTL, with the darker line segments indicating the location of the QTL. The linkage groups are divided into four columns, one for each fitted BLUP model — from left to right: 2011, 2012, 2013, and *all-years* (A); and 2011, 2012, 2014, and *all-years* (B). Each QTL interval is color-coded by its corresponding trait. The labels and interval lines to the right of linkage groups show the stable metaQTL associated with their respective genomic regions.

CNJ04 population

Despite having a lower population size (*n*= 67), and thus lower power to find statistically significant QTL, prominent QTL were found in *CNJ04*. An average of three QTL per *trait* × *model* were found using scanone and nine QTL per *trait* × *model* were mapped with stepwiseqtl. Of the 90 scanone QTL found in models with significant genotype effects, 69 were major QTL. Of the stepwiseqtl QTL found in models with significant genotype effects, 81 of 365 were major. The scanone average total BLUP variance explained per QTL was 16.9%, with an average BLUP variance explained by all QTL per *trait* × *model* of 67%. The stepwiseqtl average total BLUP variance explained per QTL was 7%, with an average BLUP variance explained by all

QTL per *trait* × *model* of 96.6%. A full list of QTL and their effect sizes can be found in Supplementary Tables S6, S7. Table II-5, Figure II-4B show the inventory of durable multi-model, within-trait meta-QTL for *CNJ04*. *PAC* content has a meta-QTL on linkage group 7 (position 28.7 cM) with a mean maternal effect size of 0.14 mg/g fruit and a mean paternal effect size of 0.19 mg/g fruit. A meta-QTL associated with the trait *UKLvW* can be found on linkage group 11 (position 10.3 cM), with similar maternal and paternal effect sizes of 0.1. A meta-QTL associated with trait *UBL* is located nearby the trait *ULvW* meta-QTL on linkage group 11 (position 16.0 cM) and has maternal and paternal effect sizes of 1.1 and 1.3, respectively. A meta-QTL for *UTBM* is found on linkage group 9 (position 82.6 cM) with a maternal effect size of 0.44 g fruit/ upright and paternal effect size of 0.19 g fruit/upright.

Multi-trait Cross-population meta-QTL

Co-located trait-marker associations representing stable, cross- population QTL are shown in Figure II-5. Together, they represent stable meta-QTL derived from the primary QTL found in the current study for *CNJ02*, *CNJ04*, and *CNJ0x*. The traits that pass the filtering constraints outlined in Table II-1 are all yield or yield- adjacent traits. Composite multi-trait QTL are found on linkage groups 2, 3, 10, 11, and 12 for set *CNJ02* (Table II-1, Figure II-5A); linkage group 11 for set *CNJ04* (Table II-1, Figure II-5B); and linkage groups 3, 11, and 12 for set *CNJ0x* (Table II-1, Figure II-5C).

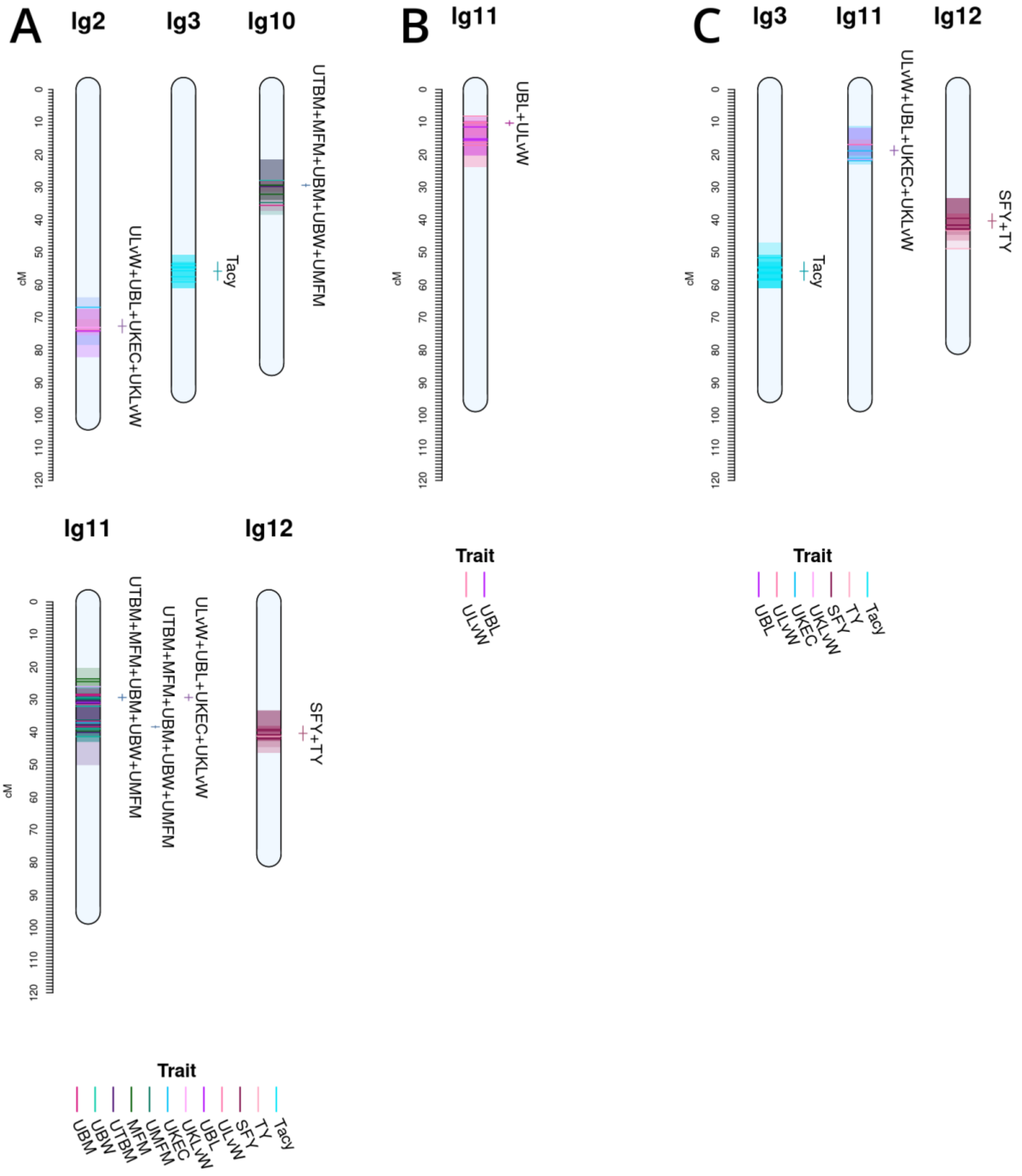


Figure II-5. Noteworthy multi-trait, co-located QTL for cranberry populations CNJ02 (A), CNJ04 (B), and both CNJ02 and CNJ04 (C), based on results from this study.

For set *CNJ02* (Table II-1, Figure II-5A), linkage group 2 has a composite QTL for {*UBL*, *ULvW*, *UKLvW*, *UKEC*} (position 72.6 cM); linkage group 3 has a year-stable meta-QTL for *Tacy (*position 55.8 cM)*;* linkage group 10 has a compound QTL for {*UBM*, *UTBM*, *UBW*, *MFM* and *UMFM*} (position 29.4 cM); linkage group 11 has three meta-QTL at two positions for {*MFM*, *UMFM*, *UTBM*, *MFM*,*BW*} and {*UBL, ULvW, UKEC, UKLvW*} (position 29.4 cM) and{*UBM*, *UTBM*, *UMFM*, *MFM*, *UBW*} (position 38.4 cM); and, last, linkage group 12 has a co-located QTL for {*SFY*, *TY*} (position 40.4 cM). Set *CNJ04* (Table II-1, Figure II-5B), linkage group 11 has a co-located QTL for {*UBL, ULvW*} (position 10.3 cM). For set *CNJ0x* (Table II-1, Figure II-5C), linkage group 3 has a meta-QTL for *Tacy* (position 55.8 cM); linkage 11 has compound QTL for upright traits{*UBL*, *ULvW*, *UKEC*, *UKLvW*} (position 18.7 cM); and linkage 12 has a composite QTL for {*SFY*, *TY*} (position 40.4 cM).

Figure II-6 shows meta-QTL found through comparison of this study's QTL to QTL previously found in Diaz-Garcia et al. (2018a); Diaz-Garcia et al. (2018b), and Schlautman et al. (2015), collectively containing meta-QTL from populations *CNJ02*, *CNJ04*, and the additional linkage mapping population *GRYG*. Cross-study meta-QTL co-located with populations *CNJ02* or *CNJ04* include those associated with sets *CNJ0x &* Diaz-Garcia, 2018b (Table II-1; Figure II-6)A; *CNJ0x &* Diaz-Garcia, 2018a (Table II-1; Figure II-6)B; *CNJ0x &* Schlautman 2015 (Table II-1; Figure II-6C), and *CNJ0x &* Schlautman, 2015 *&* Diaz-Garcia, 2018a (Table II-1; Figure II-6D; Supplementary Table S2). The shared set *CNJ0x &* Diaz-Garcia, 2018b (Figure II-6A) contains meta-QTL associated with chemistry traits, yield traits, and color traits derived from digital images of berries. Color traits are tightly associated with *Tacy* content (Diaz- Garcia, 2018b). Set *CNJ0x &* Diaz-Garcia, 2018b has two composite QTL on linkage group 3 that are stable across this study and the Diaz-

Garcia et al. (2018b) QTL study for composite traits {BCOLOR, Tacy, BCOLORVAR} (34.0 cM) and {Tacy, BCOLOR, TACY_SEP, BCOLORVAR, TACY_OCT} (56.1 cM). Set CNJ0x & Diaz-Garcia, 2018a (Figure II-6B) contains meta-QTL for traits affiliated with yield and berry quality, including digital traits measuring berry shape and size parameters (Diaz-Garcia et al., 2018a). For set *CNJ0x &* Diaz-Garcia, 2018a, linkage group 1 has a composite QTL for {*BA*, *BL*, *UBW*, *BW*} (102 cM); linkage group 2 has a meta-QTL for {*upright berry length* (*UBL*), *BL, UMFM*} (position 23.3 cM); linkage group 9 has a composite QTL for {*UTBM*, *UBM*, *BW*} (82.7 cM); linkage group 11 has three meta-QTL for {*BA*, *BL*, *BW, UBL, UMFM*}, {*LvW*, *ULvW*, *UKEC*}, and {*LvW*, *ULvW*, *EC*} clustered at positions 13.8, 21.1, and 77.8 cM, respectively; and linkage group 12 has two separate but close clusters of meta-QTL for {*LvW*, *UKEC*, *EC*} at positions 44.9 and 50.3cM, respectively. Set C*NJ0x & Schlautman* (Figure II-6C), linkage group 6 has a compound QTL for {*UTBM*, *MFM*} (position ~77.0 cM*)*; linkage group 10 has two meta-QTL for {*UBM, MFM, UMFM*} (positions 33.3 and 42.0 cM), linkage group 11 has four composite QTL for {*UBL, MFM, UMFM)*} (12.6 cM), {*PACs*, *SFY*, *TY*} (18.7 cM), multi-year {*BBITY*} (19.6 cM), and for {*UBM, UTBM*, *MFM*, *UMFM*, *UBW*} (38.4 cM). For set C*NJ0x &* Schlautman, 2015 *&* Diaz-Garcia, 2018a (Figure II-6D), linkage group 1 has a compound QTL for {*berry area* (*BA*), *BW*, *UBW*, *BL*} (position 101.6 cM); linkage group 2 has a meta-QTL for {*BL*, *UBL*, *UMFM*} (position 23.2 cM); linkage groups 9 and 10 each have a single meta-QTL— linkage group 9 for {*UBM*, *UTBM*, *BW*} (position 82.7 cM) and linkage group 10 for {*UBM*, *MFM*, *UMFM*} (position 33.3 cM); linkage group 11 has the highest number and density of meta-QTL—a cluster of three composite QTL (13–21 cM), one meta-QTL

(38.4 cM), and two co-located QTL (77.8 cM to 81.2 cM). The first cluster has meta-QTL at 13.1 cM and 16.0 cM for {*UBL*, *MFM*, *UMFM*, *BW*, *BL*} and {*EC*, *ULvW*} respectively, and meta-QTL at 18.7 cM and 21.1 cM for {PAC, SFY, TY} and {UKEC, LvW, ULvW}. Linkage group 11 also has compound QTL for {*UBM*, *UTBM*, MFM, UMFM, UBW} (38.4 cM), {EC, LvW, ULvW} (77.8 cM), and {BA, BL, MFM} (81.2 cM); and linkage group 12 has two meta-QTL for {LvW, EC, UKEC}—one at 45.0cM and one at 50.3 cM.

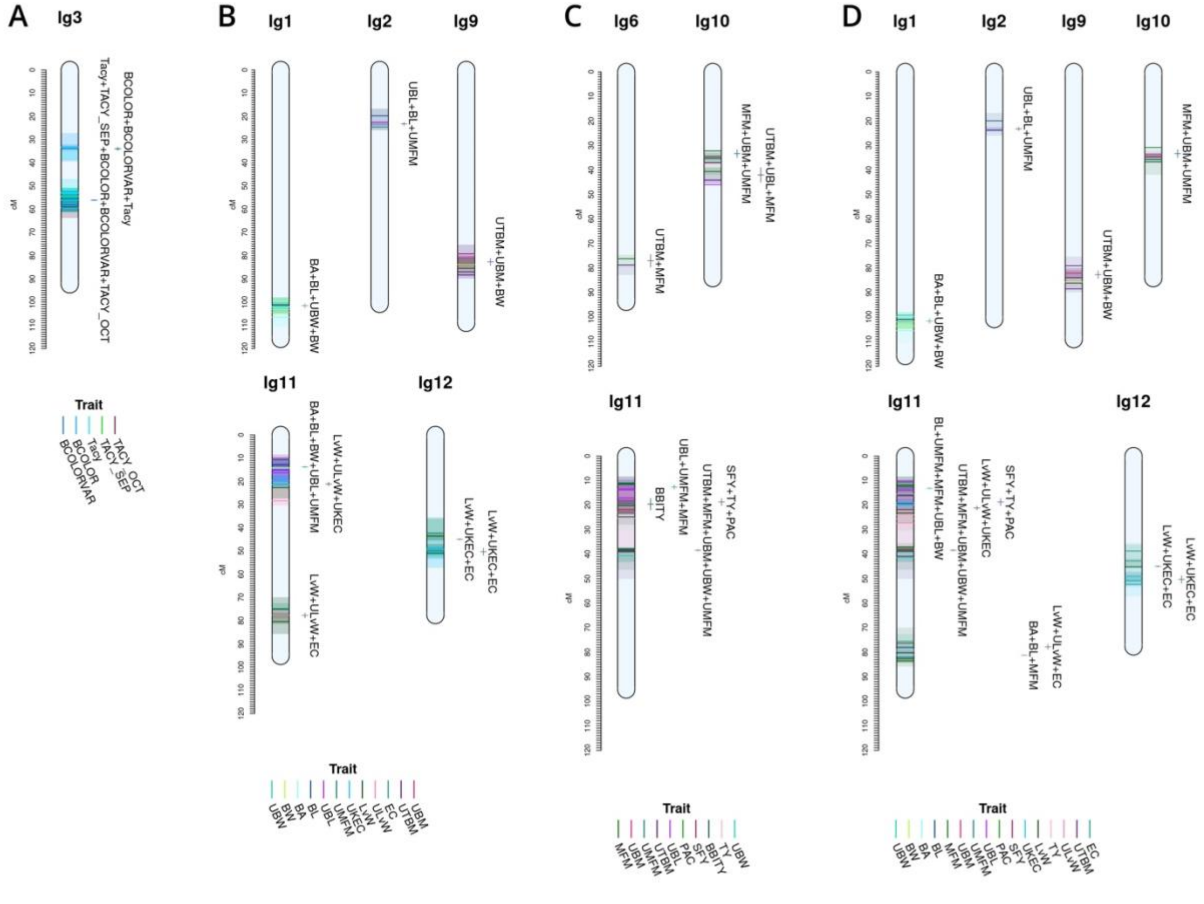


Figure II-6. Cross study co-located meta-QTL found comparing QTL from this study with QTL from Diaz-Garcia et al. (2018b) (A), Diaz-Garcia et al. (2018a) (B), Schlautman et al. (2015) (C), and a combined synthesis of Diaz-Garcia et al. (2018a) and Schlautman et al. (2015) (D) in cranberry populations CNJ02, CNJ04, and GRYG.

Table II-2. Cranberry population CNJ02 trait and BLUP summary for all-year model with heritability above 0.47. Only models with significant ($p < 0.05$) genotypic effects are displayed.

		F$_1$ progeny					Parents				
Trait	h^2*	Min_r†	Min_b‡	$m_r \pm$ SE§	$m_b \pm$ SE#	Max_r⚭	Max_b⚔	♠ P_{1r}	♥ P_{1b}	P_{2r}♦	♣ P_{2b}
ULvW	0.97	*1.10*	1.10	*1.30 ± 0.10*	1.30 ± 0.09	*1.50*	1.50	1.20	1.30	1.30	1.30
UBL	0.91	*17.0*	15.0	*22.0 ± 1.90*	20.0 ± 1.40	*26.0*	23.0	22.0	20.0	23.0	19.0
UNS	0.88	*8.70*	9.60	*20.0 ± 3.60*	16.0 ± 2.50	*30.0*	22.0	20.0	17.0	18.0	16.0
Tacy	0.88	*16.0*	23.0	*30.0 ± 7.20*	34.0 ± 5.30	*48.0*	46.0	-	32.0	-	35.0
UBM	0.77	*1.40*	1.50	*2.40 ± 0.37*	2.0 ± 0.18	*3.30*	2.50	2.20	1.90	2.80	2.0
UMFM	0.74	*1.20*	1.40	*2.20 ± 0.36*	1.80 ± 0.18	*3.00*	2.30	2.00	1.80	2.50	1.80
TY	0.74	*79.0*	−110	*310 ± 100*	41.0 ± 44.0	*570*	160.0	-	21.0	-	63.0
UCLP	0.78	*0.23*	1.00	*1.20 ± 0.36*	1.60 ± 0.23	*1.80*	2.10	0.45	1.40	1.40	1.90
UCLS	0.65	*0.40*	0.73	*1.10 ± 0.25*	1.10 ± 0.15	*1.90*	1.60	1.60	1.20	1.20	1.00
SFY	0.77	*24.0*	−140	*230 ± 91.0*	34.0 ± 46.0	*460.0*	150	-	8.10	-	60.0
UKLvW	0.69	*1.10*	1.20	*1.50 ± 0.09*	1.40 ± 0.05	*1.70*	1.50	1.40	1.40	1.50	1.40
PFR	0.73	*9.70*	17.0	*30.0 ± 13.0*	30 ± 7.40	*84.0*	60.0	-	32.0	-	29.0
UKEC	0.65	*0.35*	0.60	*0.73 ± 0.06*	0.70 ± 0.03	*0.81*	0.78	0.71	0.71	0.73	0.69
UNPs	0.51	*2.80*	2.80	*4.10 ± 0.57*	3.40 ± 0.22	*5.60*	3.90	5.20	3.20	4.40	3.60
UBW	0.60	*14.0*	14.0	*17 ± 0.93*	15.0 ± 0.32	*21.0*	16.0	17.0	15.0	18.0	15.0
UTBM	0.49	*1.90*	2.40	*4.60 ± 0.96*	3.30 ± 0.32	*7.70*	4.20	4.40	3.0	5.70	3.60
UKTO	0.55	*84.0*	85.0	*100 ± 9.40*	92.0 ± 3.50	*130*	100.0	95.0	87.0	100.0	98.0

*Narrow-sense genomic heritability for trait/model. †Minimum raw trait value. ‡Minimum BLUP trait value. §Mean raw trait value ± standard error. #Mean BLUP trait value ± standard error.

⚭Minimum raw trait value. ⚔Minimum BLUP trait value. ♠Maternal raw trait value. ♥Maternal BLUP trait value. ♦Paternal raw trait value. ♣Paternal BLUP trait value

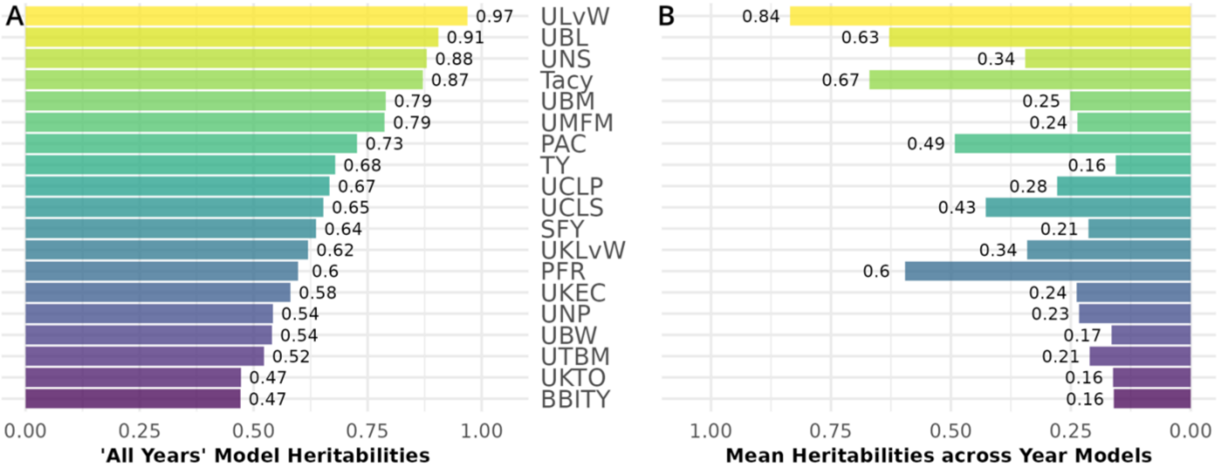


Figure II-7. Mean heritability between cranberry populations CNJ02 and CNJ04 of all-year model (A) and across within-year models (B). Only the top 50% of traits based on mean within-year model heritability are displayed.

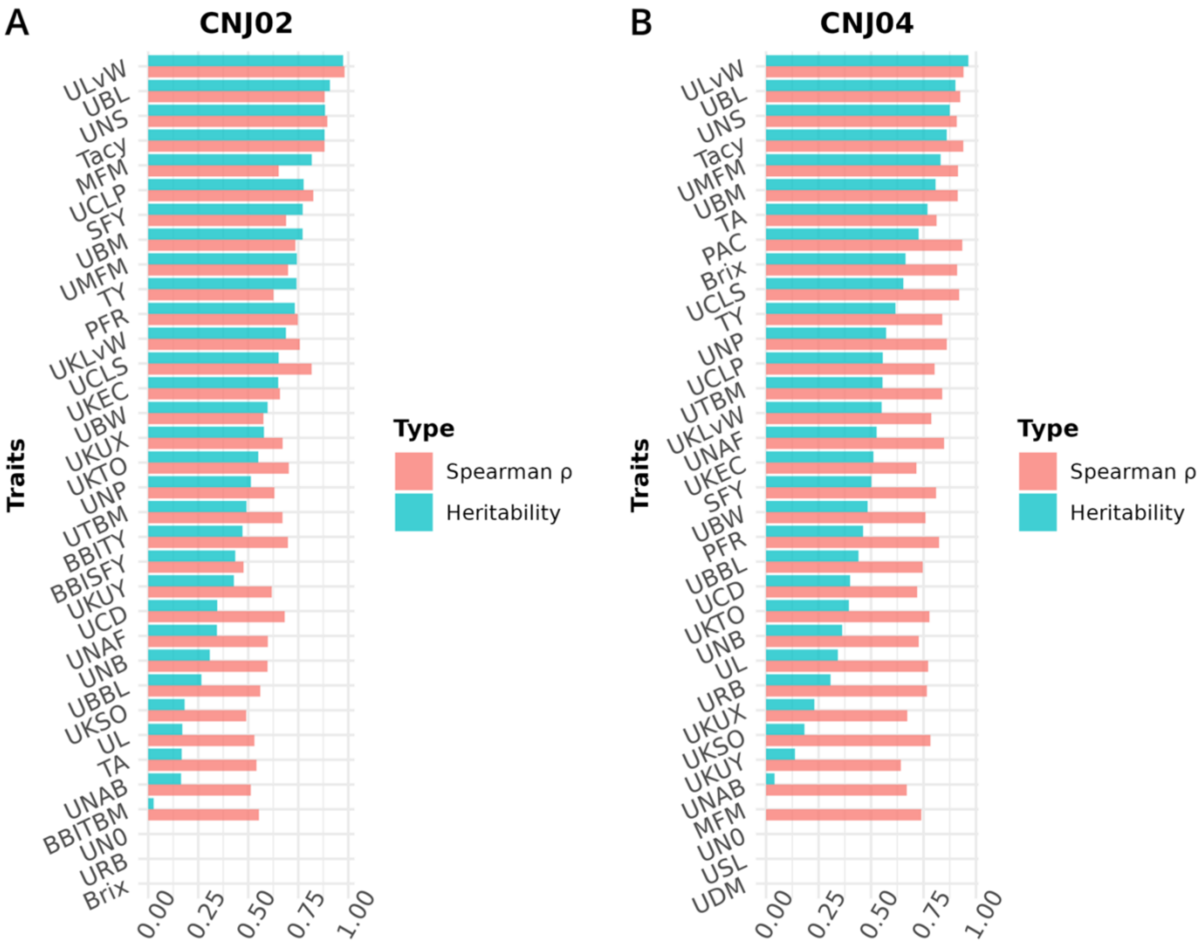


Figure II-8. Heritabilities and Spearman's rank correlation coefficients (ρ) for cranberry traits in populations (A) *CNJ02* and (B) *CNJ04*. Correlation coefficients represent the associations between genotype BLUPs and their associated phenotypes for the all-years mixed model. Cross study co-located metaQTL found comparing QTL from this study with QTL from Diaz-Garcia *et al.* (2018b) (A), Diaz-Garcia *et al.* (2018a) (B), Schlautman *et al.* (2015) (C), and a combined synthesis of Diaz-Garcia *et al.* (2018a) and Schlautman *et al.* (2015) (D) in cranberry populations *CNJ02*, *CNJ04*, and *GRYG*.

Discussion

A recent acceleration of the quantity and quality of cranberry molecular resources has propelled advances in cranberry breeding (Vorsa and Zalapa, 2019; Diaz-Garcia et al., 2020). Advances in the genetic capital of this important fruit crop include development of high-density linkage maps, construction of mitochondrial, and nuclear genome assemblies, a plethora of QTL mapping studies in a variety of important traits (Vorsa and Zalapa, 2019), and a feasibility study in genomic prediction and genomic selection (Covarrubias-Pazaran et al., 2018).

Shifts and innovations in cranberry phenotyping methods have paralleled advances in genetic resource development. Up until 60 years ago, breeders would select cranberry breeding material based on traits measured from the fundamental unit of cranberry productivity: the "reproductive upright." Since then, modernization of farming technology and management methods, improved understanding of cranberry physiology, emerging cranberry products and markets, expansion of high-yielding cultivars, and climate change have transformed how breeders prioritize and assess traits (Vorsa and Zalapa, 2019). The focus has thus shifted from assessing cranberry uprights to assessing plot- level measurements, increasingly combined with high-throughput digital imaging (Diaz-Garcia et al., 2018a, 2018b).

Despite a widespread acceptance and application of these modern trait collection methods in cranberry breeding, a translation gap exists in how upright traits relate to these newer phenotype scoring methods. With dense marker maps and parallel collection of both reproductive upright attributes and plot-level traits, the current study offers a unique opportunity to use both phenotypic paradigms for genetic mapping. This paper is the first to comprehensively report correlations, heritabilities, and QTL based on both traditional and modern phenotyping methods.

Trait correlations

In the *CNJ02* population, *Tacy* and *PAC* are negatively correlated with *SFY* but positively correlated with *PFR*. These correlations indicate that, as pigments and other flavonoids develop, cranberry fruit develop more decay, which is likely a consequence of biochemical patterns that coordinate with the timing of cranberry ripening, over-ripening, and subsequent rot (Supplementary Table S8) (Georgi et al., 2013; Daverdin et al., 2017). We also found that, in both studied populations, *MFM*—a plot yield trait—was strongly correlated with many of the per upright yield-related traits (e.g., *UBL*, *UBM*, and *UBW*). These correlations are the first reported link between a plot-level yield trait and per-upright yield traits in both *CNJ02* (Figure II-3A,B; Supplementary Table S8) and *CNJ04* (Figure II-3C,D; Supplementary Table S9). In the *CNJ02* population, modest but significant positive correlations between plot trait *TY* and upright traits {*UTBM, UBW, UBM, UMFM, UNP*s, and *UNBs*} further links plot traits with upright traits. These plot yield × upright trait correlations (*TY* × {*UNPs*, *UNB*, *UTBM*}) were also observed in the *CNJ04* population, though to a lesser degree (Figure II-3C,D; Supplementary Table S9). These correlations not only establish important relationships between plot level and upright trait types but also may indicate the existence of co-located trait blocks within the genome (Diaz-Garcia et al., 2018a, 2018b). For both the *CNJ02* and *CNJ04* populations, we found that *UKLvW*, a trait derived from berry shape categorical data, had a moderate positive correlation with quantitative traits such *ULvW*. This result validates that berry chimeras composited from the berry shape scores track well with numerically precise, accurately measured traits such as *ULvW*, despite being a subjective trait prone to imprecision.

Table II-3. Cranberry population CNJ04 trait and BLUP summary for all-year model with heritability above 0.47. Only models with significant ($p < 0.05$) genotypic effects are displayed.

		F$_1$ progeny						Parents			
Trait	h^2*	Min$_r$†	Min$_b$‡	m_r ± SE§	m_b ± SE#	Max$_r$⚭	Max$_b$⚔	P_{1r}♠	P_{1b}♥	P_{2r}♦	P_{2b}♣
ULvW	0.96	*1.10*	1.10	*1.2 ± 0.09*	1.2 ± 0.07	*1.5*	1.40	1.40	1.20	-	-
UBL	0.90	*19.0*	18.0	*22 ± 1.40*	20 ± 1.10	*27*	24.0	24.0	20.0	-	-
UNS	0.88	*8.70*	9.40	*19 ± 3.40*	17 ± 2.40	*26*	22.0	17.0	17.0	-	-
Tacy	0.86	*10.0*	25.0	*23 ± 5.80*	34 ± 4.60	*35*	42.0	-	29.0	-	-
UBM	0.81	*2.10*	2.00	*2.7 ± 0.27*	2.3 ± 0.17	*3.3*	2.80	2.80	2.30	-	-
UMFM	0.83	*1.80*	1.70	*2.4 ± 0.28*	2.0 ± 0.17	*3.3*	2.50	2.30	1.90	-	-
PAC	0.73	*1.00*	1.20	*1.4 ± 0.16*	1.4 ± 0.10	*1.8*	1.70	-	1.40	-	-
TY	0.62	*63.0*	48.0	*370 ± 89*	180 ± 41.0	*540*	260	-	160	-	-
UCLP	0.56	*0.29*	1.10	*1.2 ± 0.35*	1.5 ± 0.14	*1.9*	1.70	1.50	1.40	-	-
UCLS	0.65	*0.47*	1.10	*1.3 ± 0.32*	1.5 ± 0.18	*1.9*	1.80	1.30	1.60	-	-
SFY	0.50	*58.0*	52.0	*320 ± 82.0*	140 ± 30.0	*500*	210	-	130	-	-
UKLvW	0.55	*1.20*	1.20	*1.4 ± 0.07*	1.3 ± 0.03	*1.6*	1.40	1.40	1.30	-	-
UKEC	0.51	*0.59*	0.64	*0.72 ± 0.03*	0.68 ± 0.01	*0.79*	0.70	0.72	0.68	-	-
UNPs	0.57	*3.10*	3.80	*4.2 ± 0.48*	4.3 ± 0.23	*5.2*	4.70	3.60	4.20	-	-
UTBM	0.55	*3.10*	3.90	*5.1 ± 0.81*	4.6 ± 0.38	*7.6*	5.60	4.80	4.30	-	-

*Narrow-sense genomic heritability for trait/model. †Minimum raw trait value. ‡Minimum BLUP trait value. §Mean raw trait value ± standard error. #Mean BLUP trait value ± standard error.

⚭Minimum raw trait value. ⚔Minimum BLUP trait value. ♠Maternal raw trait value. ♥Maternal BLUP trait value. ♦Paternal raw trait value. ♣Paternal BLUP trait value.

Heritabilities

Estimating heritabilities offers breeders a chance to find confidence that their selections will be fruitful in subsequent generations. Although this study calculates heritabilities from mixed model estimates of environmental, additive genetic, and residual variances, the heritabilities derived offer breeding insights when selecting upon model derived BLUPs. Many of the calculated heritabilities in this study are comparable to the results of other studies. For example, the high mean heritability of the *ULvW* (*ULvW ratio*, $h^2 \approx 0.97$) parallels plot data findings found in *GRYG*, a population genetically distinct from *CNJ02* and *CNJ04* (Diaz-Garcia et al., 2018a). Consistently high heritability across these three populations suggests a strong genetic persistence of *ULvW ratio*. High heritabilities in both *ULvW ratio* and in related plot traits indicates that selecting rounder berries (low *ULvW ratio)* with either method is effective. Moderate heritability estimates of *TY* ($h^2 \approx 0.74$ for *CNJ02*, $h^2 \approx 0.62$ for *CNJ04*) are lower relative to other berry size and weight parameters – these estimates are consistent with the highly polygenic nature of this trait in other crops. Schlautman et al. (2015) estimated *CNJ02* heritabilities of $h^2 \approx 0.70$ and $h^2 \approx 0.64$ for mean fruit weight (*MFM*) and *TY*, respectively. Trait heritabilities for *Tacy*, *Brix*, and *titratable acidity (TA)* were consistent with the present study, while their lower heritability estimates for *TY* ($0.29 < h^2 < 0.47$) are likely due to the use of small populations (Vorsa and Johnson-Cicalese, 2012). Furthermore, a study by Johnson-Cicalese et al. (2015) using midparent-progeny mean regression estimates of heritability for fruit rot resistance ($h^2 \approx 0.81$) found consistently higher heritability with *fruit rot (%) (PFR)* in the *CNJ02* population of this study ($h^2 \approx 0.73$). Lower heritability estimates found in *CNJ04* ($h^2 \approx 0.46$) could be due to reduced statistical power from lower population size. Moderate heritability of traits like *SFY* and *PFR* demonstrate limited but possible potential for selecting rot-resistant varieties in regions where berry rot is a problem. The traits *TA* and *Brix*

Table II-4. Co-located, stable QTL in cranberry population CNJ02 found in at least three of four BLUP models. Only QTL with $R^2 \geq 10$ % are shown. Table is arranged in descending order by mean marker variance.

Trait	Models	LG	Position (cM)[*]	R^2[†]	-1.5 LOD (cM)[§]	+1.5 LOD (cM)[#]	Methods[‡]	$\overline{AvB}$[⚭]	$\overline{CvD}$[⤫]	$\overline{Int}$[♠]
Tacy	2011+ 2012+ 2013+ all years	3	55.8	27.9	52.9	58.6	scanone+ stepwiseqtl	−1.70	−11.0	0.48
UBM	2012+ 2013+ all years	11	30.8	20.9	30.3	31.2	stepwiseqtl	−0.06	−0.34	0.06
UBL	2012+ 2013+ all years	10	53.0	12.9	51.8	54.2	scanone+ stepwiseqtl	−1.80	1.10	−0.74

PFR	2011+ 2012+ all years	11	30.8	20.4	30.3	31.2	scanone+ stepwiseqtl	−10.0	−9.90	2.50
UTBM	2011+ 2012+ 2013	11	29.8	18.0	28.4	31.2	scanone+ stepwiseqtl	0.08	−0.29	0.05
UMFM	2012+ 2013+ all years	11	38.1	17.2	37.5	38.7	scanone+ stepwiseqtl	−0.11	−0.30	0.08
SFY	2011+ 2012+ 2013+ all years	12	40.4	16.0	38.1	42.6	scanone+ stepwiseqtl	10.0	46.0	−1.10
UKLvW	2011+ 2012+ all years	3	30.7	15.2	28.6	32.8	scanone+ stepwiseqtl	0.04	0.06	0.02
UKLvW	2011+ 2013+ all years	9	82.7	13.5	81.2	84.2	scanone	3e-02	0.07	4.2e-03

*Mean position of combined QTL. †Mean of variance explained by combined QTL. ‡QTL mapping method applied. §1.5 LOD left interval. #1.5 LOD right interval. ⚭Mean maternal effect size: (AC+AD)-(BC+BD). ⚔Mean paternal effect size: (AC+BC)–(AD+BD). ♠Mean interaction effect size: (AC+BD)-(AD+BC).

Table II-5. Co-located, stable QTL in cranberry population CNJ04 found in at least three of four BLUP models. Only QTL with $R^2 \geq 10$ % are shown. Table is arranged in descending order by mean marker variance. All QTL were reported below were identified with both scanone and stepwiseqtl methods.

Trait	Models	LG	Position (cM)*	R^2 †	-1.5 LOD (cM)§	+1.5 LOD (cM)#	$\overline{AvB}$⚭	$\overline{\text{CvD}}$⚔	$\overline{Int}$♠
PAC	2012+ 2014+ all years	7	28.7	26.6	24.3	33.2	0.14	0.19	−0.03
ULvW	2011+ 2014+ all years	11	10.3	25.6	9.6	11.1	0.10	0.11	0.020
UBL	2011+ 2012+ 2014	11	16.0	20.9	11.7	20.2	1.10	1.30	−2.8e-03
UTBM	2011+ 2014+ all years	9	82.6	14.0	80.9	84.2	0.44	−0.19	0.04

*Mean position of combined QTL. †Mean of variance explained by combined QTL. ‡QTL mapping method applied. §1.5 LOD left interval. #1.5 LOD right interval. ⚭Mean maternal effect size: (AC+AD)-(BC+BD). ⚔Mean paternal effect size: (AC+BC)–(AD+BD). ♠Mean interaction effect size: (AC+BD)-(AD+BC).

exhibited notable differences in heritability between *CNJ02* and *CNJ04*. In *CNJ04*, the heritabilities of *TA* ($h^2 \approx 0.77$) and *Brix* ($h^2 \approx 0.66$) were rather high, while in *CNJ02*, these two traits exhibited low heritability (*TA*, $h^2 \approx 0.17$ and *Brix*, $h^2 \approx 0$).

This likely has to do with differences in the standing genetic variation between the parents for each population. For *CNJ02*, the parents (cv. *Mullica Queen*® and cv. *Crimson Queen*®) are highly elite third-generation hybrids with little distinct genetic variation for *Brix* and *TA*, thus exhibiting low heritability in their progeny. In contrast, *CNJ04* has one highly elite parent (cv. *Mullica Queen*®), while the other parent (cv. *Stevens*) is a first-generation hybrid. Since *CNJ04's* two parents are genetically and phenotypically distinct and consequently have more standing genetic variation between the parents, these traits manifest higher heritability for *Brix* and *TA*.

The heritabilities presented here may have been affected by population sizes, reduced recombination history in F_1 mapping populations, and higher degrees of freedom when working with heterozygous, four-way crosses, and there is lower statistical power to segregate genetic from phenotypic variances in mixed models in traits such as *TY*. However, the heritability estimates were overall consistent in rank relative to other traits measured in recent cranberry studies.

QTL summary

This analysis is one of the most comprehensive QTL mapping studies in cranberry, given the number of traits assessed, the number of factors modeled, and cross-study comparisons. Figure II-5 highlights salient, multi-trait co-located QTL in population sets *CNJ02*, *CNJ04*, and *CNJ0x* (panels A, B, and C, respectively) of the current study. Multi-trait clusters among non- synonymous traits likely represent tight linkage or pleiotropy. Multiple linkage associations found on chromosomes 2, 3, 10, 11, and 12 together constitute traits important to both berry quality and yield. A lower fruit length versus width ratio and a higher MFM translates to larger, spherical

berries—quality traits important in SDC production. Quality traits relevant to SDC production— *ULvW*, *UKLvW*, *UKEC*, and *UMFM* and *MFM*)—have stable and co-located QTL on linkage groups 2, 11, and 12. *CNJ02* also displays a modest composite QTL {*SFY*, *TY*} on linkage group 12, where *SFY* is an important measure of rot. Two co-located meta-QTL on linkage group 11 ({*UBM* and *UTBM*)*, MFM, UMFM, UBW*} and {*UBL, ULvW, UKLvW, UKEC*} are likely identical QTL or the result of pleiotropy (position 29.4 cM). *CNJ04* only displays one robust meta-QTL on linkage group 11 for {*UBL*, *ULvW*} at position 10.32 cM. Despite a lack of a multi-year stable QTL for *Tacy* in *CNJ04*, the stable meta-QTL on linkage group 3@55.8 cM in *CNJ02* is also shared with a *CNJ04* QTL found in the *all-year* model. This shared cross-population *Tacy* QTL on linkage group 3 indicates the importance of this region to *Tacy* production, likely from the shared parent *Crimson Queen*® common to both *CNJ02* and *CNJ04*. The meta-QTL {*UBL*, *ULvW*, *UKEC*, *UKLvW*} common to both *CNJ02* and *CNJ04* (Figure II-5C) on linkage group 11@18.7 cM lacks the position stability relative to the same multi-trait meta-QTL found in population *CNJ02* on linkage group 11@29.4 cM
(Figure II-5A). This position shift from 11@29.4 cM to 11@18.7 cM is consistent with the larger LOD 1.5 interval for the elemental trait QTL, but the relative cross-population stability still
highlights the importance of this region to the composite {*UBL*, *ULvW*, *UKEC*, *UKLvW*}, with most of the region's genetic variation likely driven by *UBL*. As with Tacy on linkage group 3, a meta-QTL for {*SFY*, *TY*} on linkage group 12@40.4cM is found both in *CNJ02* (Figure II-5A) alone and in the combined analysis of *CNJ02* and *CNJ04* (Figure II-5C), indicating the cross-population stability of this QTL.

Figure II-6A features prominent, multi-study, multi-year, and multi-trait QTL from the set *CNJ0x &* Diaz-Garcia et al., 2018b. Tacy and other color-relevant meta-QTL on linkage group 3@56.1cM in populations *CNJ02* and *CNJ04* is consistent with the results found in Figure II-5C. This meta-QTL also demonstrates consistency across two separate studies, underpinning the importance of this genomic region to *Tacy* and color development in *CNJ02* and *CNJ04*. Evidence of an additional meta-QTL on linkage group 3@34.0 cM in population *GRYG* and *CNJ04* indicates an alternative genomic region responsible for regulating fruit color in a distinct population set. This is consistent with previously observed genomic regions that encompass QTL for many related fruit quality traits in cranberry for fruit quality traits.

Comparing the QTL found in the current study against Diaz-Garcia et al. (2018a) offers a unique perspective to compare berry size parameter QTL found using newer digital imaging techniques against parameters assayed using the manually measured traits pertinent to this study (Figure II-6B). The composite meta-QTL on linkage group 1@101.7cM establishes a correspondence of *UBW* (current study) with *BW* and *berry length* (*BL*) (*digital traits;* Diaz-Garcia et al., 2018a). *UBL* (current study) coincides with *BL* (digital trait; Diaz-Garcia et al., 2018a) at meta-QTL found in linkage groups 2@23.2 cM and 11@13.8 cM for populations *CNJ04* and *GRYG*. Meta-QTL associated with round, spherical berries that also demonstrate coincidence between manual and digital traits are found on linkage groups 11 and 12. A meta-QTL in linkage group 11@21.1cM connects the current study traits *UKEC* and *ULvW* with the image-derived digital trait *berry length:width ratio* (LvW; Diaz-Garcia et al., 2018a). The composite QTL at 11@77.8cM links the digital traits *LvW* and *berry* *shape eccentricity* (*EC*; Diaz-Garcia et al., 2018a) to *ULvW* (current study) across populations *GRYG* and *CNJ02*. Two meta-QTL for composite trait {*LvW*, *UKEC*, *EC*} on linkage group 12, positions 45.0 cM and 50.3 cM, relate upright trait (*UKEC*) with digitally measured traits {*LvW*, *EC*} (Diaz-Garcia et al., 2018a) across populations *GRYG* and *CNJ04*. These QTL complexes on linkage groups 11 and 12 together highlight the interchangeability of two distinct phenotyping (upright vs. plot) methods across distinct populations. QTL reported by Schlautman et al. (2015) are associated with plot yield traits and meta-QTL shown in Figure II-6C demonstrate the coincidence of these QTL with QTL found in the current study. The meta-QTL on linkage group 6@77.0cM ties UTBM (current study) to *MFM* (Schlautman et al., 2015), consistent with the high correlation (ρ = 0.71) between these two traits (Supplementary Table S8). Linkage group 10 has two meta-QTL that demonstrate a correspondence of *MFM* (Schlautman et al., 2015) with {*UBM*, *UMFM*} (current study; position 33.3 cM) and {*UTBM*, *UBL*} (current study; position 42.0). Linkage group 11 contains several meta-QTL that associate the yield related plot traits *biennial- bearing index - total yield* (*BBITY*), *TY*, and *MFM* (Schlautman et al., 2015) with upright traits *UBL*, *UMFM*, *UTBM*, *UBM*, and *UBW*, further establishing the congruity between plot traits and upright traits.

A comprehensive comparison of cross-study meta-QTL is shown in Figure II-4D. The meta-QTL displayed include compound, stable QTL found across up to three independent studies (current study; Schlautman et al., 2015; Diaz-Garcia et al., 2018a). QTL found in Figure II-5D are duplicated in Figure II-5B, apart from a three- study meta-QTL on linkage group 11@13.1 cM for the compound trait {*UBL*, *MFM*, *UMFM*, *BW*, *berry length* (*BL*)}, which differs inappreciably from the meta-QTL found for composite trait {*berry area* (*BA*), *UBL*, *UMFM*, *BW*, *BL*} on linkage group 11@13.8cM (Figure II-4B). This meta-QTL for linkage group 11, position 13.1cM includes a *MFM* QTL discovered using a different dataset and different methodology by Schlautman et al. (2015). Consequently, this genomic region shows evidence of a multi-study, multi-trait, multi-year QTL for mean berry size. Other multi-study meta-QTL that include QTL from Schlautman et al. (2015) include upright traits {*UBM*, *UMFM*} and plot trait *MFM* on 10@33.3cM, additional evidence of an association of *PAC* with *SFY* and *TY* on 11@18.7cm, upright traits {*UBM*, *UTBM*, *UMFM*, *UBW*} (current study) with plot trait *MFM* (Schlautman et al., 2015) on 11@ 38.4cM, and digital traits {*BA*, *BL*} (Diaz-Garcia et al., 2018a) with plot trait *MFM* (Schlautman et al., 2015) on 11@81.2cM.

Conclusion

Data availability statement

All raw data can be found on the Genome Database for Vaccinium at https://www.vaccinium.org/bio_data/7547154. Software to generate BLUPs, QTL, and meta-QTL are available at https://github.com/bliptrip/CNJ0x-Trait-Mapping.

Author contributions

AM: Conceptualization, Writing – original draft, Writing – review & editing, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization. JL: Formal analysis, Writing – review & editing. LD-G: Conceptualization, Formal analysis, Writing – review & editing. HL-M: Formal analysis, Writing – review & editing. JJ-C: Writing – review & editing, Conceptualization, Data curation, Resources. NV: Conceptualization, Resources, Writing – review & editing, Funding acquisition. MI: Conceptualization, Funding acquisition, Resources, Writing – review & editing. JN: Resources, Writing – review & editing. JZ: Resources, Writing – review & editing, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project was supported by USDA-ARS (project no. 5090-21220-007-00-D provided to JZ); USDA AFRI NIFA (grant no. 2022-67012-37202 provided to JL); NIH (grant no. 5 T32 GM135066AM; UW- Madison Biotechnology Training Program; provided to AM); USDA NIFA (project no. 2019-51181-30015; VacciniumCAP); Ocean Spray Cranberries, Inc.; Wisconsin Cranberry Growers Association; and Cranberry Institute. All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The funders had no role in the study design, data collection, and analysis, interpretation of the data, writing and preparation of this article, or decision to publish.

Acknowledgments

The authors thank Emily Gustin and Eric Wiesman for their invaluable help with data collection management. AM would also like to thank Allen Centennial Gardens for providing an oasis to recharge his mind. The authors also thank the reviewers who helped enhance the quality of this paper. JZ would like to express his gratitude through PS 136:1.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary Material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1294570/full#supplementary-material

References

Alston, J. M., Medellıń-Azuara, J., and Saitone, T. L. (2014). *Economic impact of the North American cranberry industry* (Cranberry Marketing Committee). Available at: https://www.bccranberries.com/pdfs/Economic_Impact_of_the_NA_Cranberry_Industry_August2014.pdf.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automatic Control* 19, 716–723. doi: 10.1109/TAC.1974.1100705

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Bolivar-Medina, J. L., Zalapa, J. E., Atucha, A., and Patterson, S. E. (2019). Relationship between alternate bearing and apical bud development in cranberry (*Vaccinium macrocarpon*). *Botany* 97, 101–111. doi: 10.1139/cjb-2018-0058

Bribiesca, E. (2013). A measure of tortuosity based on chain coding. *Pattern Recognition* 46, 716–724. doi: 10.1016/j.patcog.2012.09.017

Broman, K. W., Sen, S., and New York, N. Y. (2009). *A guide to QTL mapping with* R/ *qtl* (New York, NY: Springer New York). doi: 10.1007/978-0-387-92125-9

Broman, K. W., Wu, H., Sen, S., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19, 889–890. doi: 10.1093/bioinformatics/btg112

Chandler, F., Wilcox, R., and Bain, H. (1947). Cranberry breeding investigation of the US Dept. of Agriculture. *Cranberries*. 12, 6–9.

Covarrubias-Pazaran, G. (2016). Genome-assisted prediction of quantitative traits using the R package sommer. *PloS One*. 11, 1–15. doi: 10.1371/journal.pone.0156744

Covarrubias-Pazaran, G., Diaz-Garcia, L., Schlautman, B., Deutsch, J., Salazar, W., Hernandez-Ochoa, M., 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. doi: 10.1186/s12864-016-2802-3

Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., et al. (2018). Multivariate GBLUP improves accuracy of genomic selection for yield and fruit weight in biparental populations of *vaccinium macrocarpon* ait. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01310

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

De Los Campos, G., Sorensen, D., and Gianola, D. (2015). Genomic heritability: what is it? *PloS Genet.* 11, 1–21. doi: 10.1371/journal.pgen.1005048

Diaz-Garcia, L., Covarrubias-Pazaran, G., Johnson-Cicalese, J., Vorsa, N., and Zalapa, J. (2020). Genotyping-by-sequencing identifies historical breeding stages of the recently domesticated american cranberry. *Front. Plant Sci.* 11. doi: 10.3389/ fpls.2020.607770

Diaz-Garcia, L., Covarrubias-Pazaran, G., Schlautman, B., Grygleski, E., and Zalapa,
J. (2018a). Image-based phenotyping for identification of QTL determining fruit shape
and size in American cranberry (*Vaccinium macrocarpon* L.). *PeerJ* 6, e5461. doi: 10.7717/peerj.5461

Diaz-Garcia, L., Covarrubias-Pazaran, G., Schlautman, B., and Zalapa, J. (2016). GiNA, an efficient and high-throughput software for horticultural phenotyping. *PloS One* 11, e0160439. doi: 10.1371/journal.pone.0160439

Diaz-Garcia, L., Garcia-Ortega, L. F., González-Rodríguez, M., Delaye, L., Iorizzo, M., and Zalapa, J. (2021). Chromosome-Level Genome Assembly of the American Cranberry (*Vaccinium macrocarpon* Ait.) and Its Wild Relative *Vaccinium microcarpum*. Front. *Plant Sci.* 12. doi: 10.3389/fpls.2021.633310

Diaz-Garcia, L., Rodriguez-Bonilla, L., Smith, T., and Zalapa, J. (2019). Pacbio sequencing reveals identical organelle genomes between american cranberry (*Vaccinium macrocarpon* ait.) and a wild relative. *Genes* 10, 1–15. doi: 10.3390/ genes10040291

Diaz-Garcia, L., Schlautman, B., Covarrubias-Pazaran, G., Maule, A., Johnson- Cicalese, J., Grygleski, E., et al. (2018b). Massive phenotyping of multiple cranberry populations reveals novel QTLs for fruit anthocyanin content and other important chemical traits. *Mol. Genet. Genomics* 293, 1379–1392. doi: 10.1007/s00438-018-1464-z

Eck, P. (1990). *The american cranberry* (New Brunswick : Rutgers University Press).

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One* 6, e19379. doi: 10.1371/journal.pone.0019379

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024

Fajardo, D., Schlautman, B., Steffan, S., Polashock, J., Vorsa, N., and Zalapa, J. (2014). The American cranberry mitochondrial genome reveals the presence of selenocysteine (tRNA-Sec and SECIS) insertion machinery in land plants. *Gene* 536, 336–343. doi: 10.1016/j.gene.2013.11.104

Fajardo, D., Senalik, D., Ames, M., Zhu, H., Steffan, S. A., Harbut, R., et al. (2013). Complete plastid genome sequence of *Vaccinium macrocarpon*: structure, gene content, and rearrangements revealed by next generation sequencing. *Tree Genet. {&} Genomes* 9, 489–498. doi: 10.1007/s11295-012-0573-9

Food and Agriculture Organization of the United Nations (2022). *FAOSTAT statistical database* (Rome: FAO). Available at: https://www.fao.org/faostat/en/ #home. 1997. 08 Jan. 2024.

Fox, J., and Weisberg, S. (2019). *An R companion to applied regression. 3rd ed.* (Thousand Oaks CA: Sage). Available at: https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Franklin, H. J., Kelley, J.Jr., and T.W., E. (1958). Cranberry varieties of north america, bulletin 513. Experiment station, college of agriculture (University of Massachusetts).

Georgi, L., Johnson-Cicalese, J., Honig, J., Das, S. P., Rajah, V. D., Bhattacharya, D., 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. doi: 10.1007/ s00122-012-2010-8

Hagberg, A., Swart, P. J., and Schult, D. A. (2008). “Exploring network structure, dynamics, and function using NetworkX,” in *Proceedings of the 7th python in science conference (SciPy 2008).* (Los Alamos, NM, United States: Los Alamos National Laboratory). Available at: https://www.osti.gov/biblio/960616.

Haley, C. S., and Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69, 315–324. doi: 10.1038/hdy.1992.131

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31, 423–447. doi: 10.2307/2529430

Johnson-Cicalese, J., Polashock, J. J., Honig, J. A., Vaiciunas, J., and Ward, D. L. (2015). Heritability of fruit rot resistance in american cranberry. *J. Am. Hortic. Soc.* 140, 233–242. doi: 10.21273/JASHS.140.3.233

Kassambara, A. (2019). R package “ggcorrplot”: Visualization of a correlation matrix using “ggplot2.”

Kawash, J., Colt, K., Hartwick, N. T., Abramson, B. W., Vorsa, N., Polashock, J. J., et al. (2022). Contrasting a reference cranberry genome to a crop wild relative provides insights into adaptation, domestication, and breeding. *PloS One* 17, e0264966. doi: 10.1371/journal.pone.0264966

National Agricultural Statistics Service (2022). (Washington, D.C: National Agricultural Statistics Service, U.S. Dept. of Agriculture). Available online at: https:// www.nass.usda.gov/. (Accessed August 30, 2023)

Neyhart, J. L., Kantar, M. B., Zalapa, J., and Vorsa, N. (2022). Genomic- environmental associations in wild cranberry (*Vaccinium macrocarpon* Ait.). *G3 Genes|Genomes|Genetics* 12, jkac203. doi: 10.1093/g3journal/jkac203

Polashock, J., Zelzion, E., Fajardo, D., Zalapa, J., Georgi, L., and Bhattacharya, D. (2014). The American cranberry : first insights into the whole genome of a species adapted to bog habitat The American cranberry : first insights into the whole genome of a species adapted to bog habitat. *BMC Plant Biol.* 14.

Ouellette, L. A., Reid, R. W., Blanchard, S. G., and Brouwer, C. R. (2018). LinkageMapView—rendering high-resolution linkage and QTL maps. *Bioinformatics* 34, 306–307. doi: 10.1093/bioinformatics/btx576

R Core Team (2024). *R: a language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: https://www.R-project.org/.

Rossum, G. van (2007). *Python programming language., in USENIX annual technical conference.* (Santa Clara, CA (United States), 1–36. Available at: https://api. semanticscholar.org/CorpusID:45594778.

Schlautman, B., Bolivar-Medina, J., Hodapp, S., and Zalapa, J. (2017a). Cranberry SSR multiplexing panels for DNA horticultural fingerprinting and genetic studies. *Scientia Hortic.* 219, 280–286. doi: 10.1016/j.scienta.2017.03.005

Schlautman, B., Covarrubias-Pazaran, G., Diaz-Garcia, L., Iorizzo, M., Polashock, J., Grygleski, E., et al. (2017b). . Construction of a high-density american cranberry (*Vaccinium macrocarpon* ait.) composite map using genotyping-by-sequencing for multi-pedigree linkage mapping. *G3 Genes|Genomes|Genetics* 7, 1177–1189. doi: 10.1534/g3.116.037556

Schlautman, B., Covarrubias-Pazaran, G., Diaz-Garcia, L. A., Johnson-Cicalese, J., Iorrizo, M., Rodriguez-Bonilla, L., et al. (2015). Development of a high-density cranberry SSR linkage map for comparative genetic analysis and trait detection. *Mol. Breed.* 35, 177. doi: 10.1007/s11032-015-0367-5

Sun, H., Liu, Y., Gai, Y., Geng, J., Chen, L., Liu, H., et al. (2015). *De novo* sequencing and analysis of the cranberry fruit transcriptome to identify putative genes involved in flavonoid biosynthesis, transport and regulation. *BMC Genomics* 16, 652. doi: 10.1186/ s12864-015-1842-4

Vander Kloet, S. P. (1988). *The genus* vaccinium *in north america* (Ottawa: Research Branch, Agriculture Canada).

Vander Kloet, S. P., and Avery, T. S. (2010). *Vaccinium* on the edge. *Edinburgh J. Bot.* 67, 7. doi: 10.1007/978-1-4419-0763-9_6

Vorsa, N., and Johnson-Cicalese, J. (2012). American cranberry. *Fruit Breed. Handb. Plant Breed.* 8, 191–223. doi: 10.1007/978-1-4419-0763-9

Vorsa, N., and Zalapa, J. (2019). "Domestication, genetics, and genomics of the american cranberry," in *Plant Breeding Reviews*, (John Wiley & Sons, Ltd), 279–315. doi: 10.1002/9781119616801.ch8

Wei, T., and Simko, V. (2021). *R package "corrplot": visualization of a correlation matrix*. Available at: https://github.com/taiyun/corrplot

Žalik, B., Mongus, D., Liu, Y.-K., and Lukač, N. (2016). Unsigned Manhattan chain code. *J. Visual Communication Image Representation* 38, 186–194. doi: 10.1016/ j.jvcir.2016.03.001

Chapter III Modeling Spring Greening Patterns among Cranberry Genotypes via Aerial Imaging of Leaf Color

Abstract

Cranberry breeding presents unique challenges due to limited domestication, long lifecycle, and extended evaluation cycles. Spring is a particularly vulnerable time for cranberry production, as late frosts can damage floral buds and compromise crop yields. This study demonstrates the potential of uncrewed aerial vehicles (UAV) mounted with red, green, blue (RGB) cameras to precisely assay spring vegetative developmental variation among cranberry genotypes. As an evergreen broadleaf, cranberry leaves exhibit elevated anthocyanin pigment levels during winter dormancy and turn green during the transition to warmer months. We developed a novel phenotype scoring index that quantifies cranberry spring developmental performance by capturing leaf color imagery across frost periods to assess spring development over time. Session imagery was captured across three populations at two distinct locations encompassing eight distinct capture dates in late spring of the years 2018 and 2019, with multiple sessions captured per date to assemble repeated measures. For each capture date, a random subsample of around ten percent of the plots had leaf tissue collected to ground truth anthocyanin pigment levels. Twenty-three image-derived vegetative indices (VIs) were fit to the ground truth values using five different regression methods: ordinary least squares estimator (OLSE) regression, ridge regression (RR), principal components regression (PCR), partial least squares regression (PLSR), and simulated annealing (SA). SA was ultimately chosen to predict all unfit plot image anthocyanin levels, as it balanced model performance ($R^2 = 0.886$) with interpretability and parsimony, reducing the number of VI predictors to eight. Predicted anthocyanin levels were fit to exponential decay curves at the genotype and population level. A selection index was developed by integrating the area between the genotype curves and their respective population curves, condensing the multivariate temporal and developmental genotype state into a single scalar value for comparison. This index allows researchers to score genotypes based on ideal spring

developmental characteristics, providing a late spring frost-avoidance metric to advance cranberry breeding. Furthermore, this selection index enables downstream marker-trait association studies and genomic selection applications.

Introduction

The American cranberry (*Vaccinium macrocarpon* Ait.) belongs to a rich and diverse genus with over 500 species adapted to marginal habitats. American cranberry is one of only a few species in *Vaccinium* of commercial importance, including blueberries and lingonberries (Vander Kloet, 1988; Vander Kloet and Avery, 2010). Despite its economic importance and over a century of breeding, domestication of most planted cultivars is limited to only one or two generations beyond wild germplasm, with only recent introduction of third-generation cultivars into cultivated marshes (Chandler et al., 1947; Vorsa and Zalapa, 2019). As such, advancing genomic and phenomic tools that accelerate selection and reduce phenotyping costs is critical to maximize genetic gain.

Plant pigments play diverse roles in plant physiology, from photosynthesis and photoprotection to reactive oxygen scavenging and plant defense (Hughes et al., 2005). Most of these plant pigments exhibit distinct spectral absorbance (λ_a) and reflectance (λ_r) signatures. Common pigments include carotenoids (λ_a 450-550 nm); chlorophylls (Chl) *a* and *b* in the blue spectrum (λ_a 435-460 nm); anthocyanins (λ_a 520-560 nm); and Chl *a* and *b* in the red spectrum (λ_a 660-680 nm) (Sims and Gamon, 2002; Gitelson et al., 2009; Hashimoto et al., 2016; Del Valle et al., 2018). Digital photography offers an affordable, accessible, efficient, reliable, and scalable method to capture and characterize pigment signatures in plant tissues (Yang et al., 2017b; Tanaka et al., 2024). By combining digital images with other environmental sensor data, models can be developed and utilized to infer plant physiological states, providing real-time crop management guidance or highlighting selectable, segregating patterns in mapping or diverse populations.

Red-green-blue (RGB) cameras were designed to perceptually mimic human vision in the 400-700 nm light wavelength band, which coincides with most plant pigment reflectances (Carter and Knapp, 2001; Lee, 2005; Del Valle et al., 2018). Various features have been extracted from RGB cameras as inputs to models developed for predicting various crop physiological and phenological processes, including spectral reflectances (SR), vegetation indices (VI), and texture features (TF). In the context of RGB digital images, SRs are the digital pixel values recorded for the red (R), green (G), and blue (B) channels, or their alternate and equivalent color space conversions. VIs are mathematical combinations of SRs and often are normalized to make them less sensitive to variation in capture conditions (Huang et al., 2024). VIs can also be constructed by accumulating histogram frequencies over defined SR ranges, like the green area (GA) and greener area (GGA) indices (Casadesús et al., 2007).

RGB-based VIs have proven effective for estimating plant pigments and phenology, with red and green channels having successfully measured leaf anthocyanin content (Gamon and Surfus, 1999); weed segmentation and biomass quantification (Woebbecke et al., 1993; Pérez et al., 2000); and vegetation-soil segmentation (Meyer and Neto, 2008). More recently, specialized indices like the anthocyanin content-chroma ratio (AC-CR), anthocyanin content-chroma basic (AC-CB), and strength of green (SG) indices have shown high accuracy in detecting anthocyanin levels across plant vegetative and reproductive tissues (Del Valle et al., 2018). Lastly, TFs encapsulate localized spatial variation of image SRs and/or VIs and have shown utility in enhancing model inference performance in many remote sensing crop applications (Hlatshwayo et al., 2019; Zheng et al., 2019; Johansen et al., 2020; Liu et al., 2022; Huang et al., 2024).

Modern phenomics has emerged not only through digital imaging but also through the synergy of several technological advances: affordable high throughput sensing platforms, cost-effective cloud computing solutions, innovative data fusion methods, and sophisticated analytical algorithms (Zhao et al., 2019; Song et al., 2021). These advances have substantially reduced phenotyping bottlenecks in breeding programs (Araus et al., 2018) by expanding access to high fidelity trait monitoring tools (Covarrubias-Pazaran et al., 2016; Yang et al., 2017; Zhao et al., 2019). With the development of mature, open source toolkits and inexpensive cloud computing resources, large datasets are more available and standardized (Gehan et al., 2017; Berry et al., 2018; Zhao et al., 2019). Innovative machine learning algorithms have enabled many successful applications in predicting plant phenology, morphology, yield, and physiological stress responses from images and sensor data (Zhao et al., 2019; Weiss et al., 2020).

The adoption of high throughput phenotyping (HTP) methods for cranberry monitoring and breeding has grown steadily over the last decade. GiNA, an early application of HTP to cranberries, is a digital imaging tool that quantifies external fruit appearance traits including shape and color parameters (Diaz-Garcia et al., 2016). Building upon this foundation, BerryPortraits offers an open-source enhancement that incorporates advanced segmentation algorithms with color correction capabilities (Loarca et al., 2024). These imaging tools have been used for cranberry breeding and genetics studies such as QTL mapping for color and shape in cranberries (Diaz-Garcia et al., 2018a, 2018b). In addition, hyperspectral and satellite imagery have demonstrated utility in mapping foliar nutrient status in cranberry beds (Liu et al., 2023; Huang et al., 2024). However, the applicability of hyperspectral and satellite imaging to genetics and breeding is limited by their low spatial resolution, as initial breeding trials use hundreds of genotypes planted in smaller plots.

Recent novel developments in using UAVs to assess and evaluate cranberry health and phenology have combined high spatial resolution images with deep neural networks (Akiva et al., 2020, 2022; Johnson et al., 2023). These studies primarily modeled economic risks associated with high temperatures that damage cranberries. They combined unique deep-learning training methods to efficiently segment and predict cranberry counts in the field, calculate cranberry albedo, integrate weather data through sensors, and model cloud trajectory and coverage to infer crop risk and inform management decisions. These same methods are also employed to assess ripening patterns in cranberry beds, with sufficient spatial resolution to assess smaller breeding plots (Johnson et al., 2023).

As non-deciduous broadleaf 'evergreens', cranberry plants retain their leaves over the winter months and exhibit transient reddening due to anthocyanin accumulation. Various theories exist on the role of anthocyanin pigments in angiosperm 'evergreens': reactive oxygen species scavengers, ultraviolet B sunscreen, antiherbivory compounds, antifungal agents, light-quenching agents that protect inefficient photosystems in cold temperatures, carbohydrate sinks, and elevated warming under high light and cold temperature conditions (Hughes et al., 2005; Hughes, 2011; Davies et al., 2022). A study by Abdallah A. (1989) on the biophysical and biochemical changes in cranberry leaves found that decreases in anthocyanin and increases in chlorophyll in deacclimating leaves in spring were not directly associated with shifts in leaf freezing stress resistance. In contrast, Workmaster B. (2001) later found color transitions in leaves were synchronous with bud morphology changes from tight bud to swelling, which in turn have been associated with shifts in bud freezing stress resistance (Workmaster and Palta, 2006, 2009).

At minimum, this study explores the application of an efficient HTP monitoring tool to predict cranberry leaf anthocyanin levels in spring. Assuming a modest link between leaf color changes and phenological shifts in bud freezing stress resistance (Workmaster and Palta, 2006, 2009), this study extends its scope to enable efficient assessment of frost resistance in cranberry. To develop an efficient HTP monitoring tool, we used an UAV to capture images of cranberry plots on three populations planted at two locations over two years. We extracted image features, or vegetative indices (VIs), and characterized their inter-associations to relevant developmental parameters, including the physiological metric GDD, anthocyanin spectrophotometer proxy A535 absorbance units (AU), and chlorophyll spectrophotometer proxies A649 AUs and A660 AUs. We then fit and evaluated models using the A535 AU as response and VIs as predictors. We performed model selection to simplify the models, prevent model overfitting, ensure coefficient interpretability, and reduce prediction variance and predict plot A535 AU values from all image plots. Finally, we developed a novel selection index to score and rank model genotypes based on spring-greening patterns that are consistent with a joint post-frost exit from dormancy and accelerated post-dormancy growth.

This study highlights the potential of employing UAVs or other robotic technology outfitted with imaging capabilities to moderate crop economic risks through long term cost-effective breeding and genetic improvement programs in addition to short and medium term management practices.

Materials and Methods

Populations

Tissue samples and UAV images (described below) were collected from three populations and two locations in central Wisconsin, U.S.A, per Figure III-1A. The first site was planted in 2013 with two breeding populations at Cranberry Creek Cranberry, Inc. in Necedah. CNJ02 (n = 176) was derived from a cross between seed-bearing parent, CNJ97_105_3 (Mullica Queen®) and pollen-donating parent, NJS98_23 (Crimson Queen®); CNJ04 (n = 69) was derived from reciprocal crosses between CNJ97_105_3 and Stevens. The planting follows an augmented design with regularly spaced parental checks at four locations in the cranberry bed (Federer et al., 1956; Federer and Raghavarao, 1975; Lin and Poushinsky, 1983). CNJ02 and CNJ04 were assessed in 2018 and 2019.

The second site was planted in 2010 at Valley Corporation in Tomah, surveyed in 2018. GRYG (n = 494) was derived from a cross between seed-bearing parent, [BGx(BLxNL)]95, and the pollen-donating parent, GH1x35.

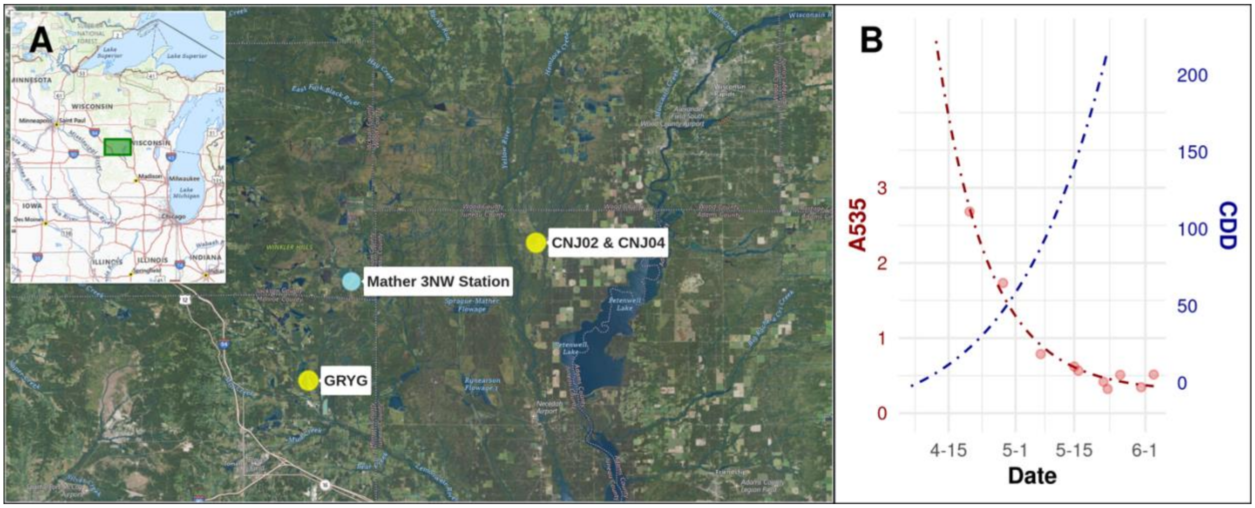


Figure III-1. A. Geographic map showing location of study's three cranberry breeding populations and the proximal weather station with daily temperature data. B. Spring temporal patterns for mean reference A535 (dark red) and CDD (dark blue) for years 2018-19. Dashed lines represent the corresponding best fit curve for both datasets. Points represent the mean A535 values measured on a given month and day.

Leaf Sampling and Processing

Leaf tissue samples were collected for each imaging session as references to quantify relative leaf tissue chlorophyll and anthocyanin levels. The procedure was derived from Workmaster B. (2001) where they modeled spring growth and development in cranberries by analyzing chlorophyll and anthocyanin in upright foliar tissue. In summary, approximately 16 uprights were collected fromaround 30-40 randomly selected plots before performing an UAV imaging session. The top 3 cm of leaves from each upright were stripped, pooled, weighed per genotype sample, and subsequently lyophilized for approximately 48 hours or until fully dry. Dried leaf tissue weights were recorded for each sample, and approximately two random samples of around 0.1 g were allotted into separate 15 mL polystyrene conical tubes. One 6 mm glass bead and three 3 mm glass beads were each added to a conical tube and tissues were homogenized completely using a repurposed paint shaker optimized to process biological samples.

To extract chlorophyll from homogenized leaf tissue, 10 mL of ice-cold 95% v/v ethanol was added to each conical tube, tissues were thoroughly admixed using a vortex instrument and stored out of direct light at 4°C for three days. After three days, tubes were vortexed again to fully mix and centrifuged for four minutes at 3,050 × *g*. Spectrophotometer absorbance readings of 1 mL supernatant were recorded at 649 nm and 660 nm, using a 95% ethanol blank (Wintermans and De Mots, 1965). All samples were processed using a ThermoSpectronic Genesys 5 model 336001 spectrophotometer (Thermo Electron Corporation, Madison, WI U.S.A.). Values were recorded as absorbance units (AU) normalized on a per volume solvent and per mass of lyophilized tissue.

To extract and record total anthocyanins, the general method of Fuleki and Francis (1968) was used. To the same conical tubes used to assay chlorophyll, 1.59 mL of 1.5 N HCl was added. Tubes were vortexed to admix tissue thoroughly and incubated at 4°C in the dark for three days. Following incubation, tubes were again vortexed to mix tissue in the solution, insoluble solids were precipitated in a centrifuge for four minutes at 3,050 × *g*. Spectrophotometer absorbance readings of 1 mL supernatant were recorded at 535 nm, using a 95% ethanol + 1.5N HCl blank. Instrumentation used and units were identical to those for chlorophyll.

UAV Imaging

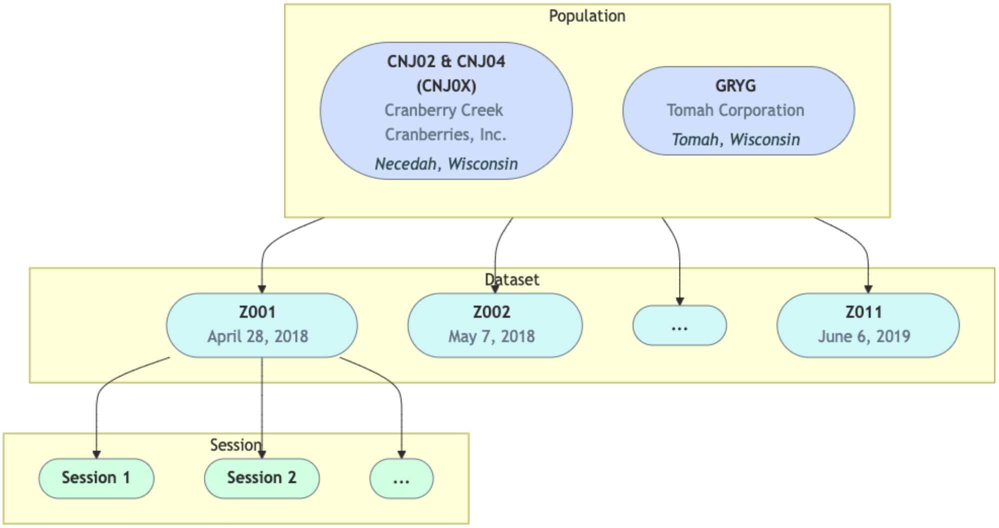


Figure III-2. UAV cranberry imaging process topology. Breeding field plots for population CNJ02 and CNJ04 were imaged at Cranberry Creek Cranberries (site 1) and field plots for population GRYG were assessed at Tomah Corporation (site 2). Datasets represent a given capture date and are numbered incrementally from Z001. Sessions represent distinct, redundant, and comprehensive captures of all field plots at a given location.

Cranberry plots were imaged with a DJI Phantom 4 Pro quadcopter UAV (Da-Jiang Innovations, Ltd., Shenzhen, Guandong, China). The DJI UAV was chosen to balance cost, quality, and ease-of-use to phenotype plots at high throughput. The UAV has a gimbal mounted RGB camera with a 20-megapixel, 2.54 cm CMOS RGB sensor and a lens with a focal length of 8.8 mm. Flight sessions were programmed to fly at 5.5 - 7.5 meters above ground level, using video in 2018 and still image captures in 2019. Image overlap was between 30 and 40 percent, balancing a DJI firmware limitation of 99 photos per session with complete field coverage per flight session at low altitudes. Ground sampling distances were estimated to be between 0.39 cm/pixel at 5.5 meters and 0.53 cm/pixel at 7.5 meters for video captures (4096 × 2160 pixel resolution) and between 0.29 cm/pixel at 5.5 meters and 0.40 cm/pixel at 7.5 meters for image captures (5472 × 3648 pixel resolution).

Session imagery was captured between two populations and across the eight datasets, with each dataset representing a distinct capture date. Within each dataset, multiple sessions were captured to increase data depth and to introduce redundancy in the event of reconstruction errors or unstable capture lighting and altitude. Figure III-2 illustrates the terminology relevant to the UAV image capture process used in this study.

Vegetative Indices

Twenty-three different image features, referred to herein as vegetative indices (VIs), were extracted from UAV images of cranberry plots (Table III-1). For all image pixels corresponding to a cranberry plot, VI values were computed. VI means for each plot in an image were subsequently used to train regression models to spectrophotometer absorbances.

Table III-1. Vegetative indices (VIs) used to study cranberry breeding field plots for cranberry populations CNJ02, CNJ04, and GRYG.

Abbrev.	Name	Formula*	Reference
A	a* axis	Nonlinear transformations of RGB spectral reflectances to CIE Lab colorspace. CIE Lab a*	

		values map to green-red opponent colors, where negative values approach green and positive values tend towards red.	
ac.cb	anthocyanin content - chroma basic	(B+R)/G	Del Valle et al. (2018)
ac.cr	anthocyanin content - chroma ratio	G/((B+R)/2)	Del Valle et al. (2018)
b	blue chromatic coordinate	B/(R+G+B)	Woebbecke et al. (1995)
cive	color index of vegetation extraction	$0.441R_n-0.811G_n+0.385B_n+18.78745$	Kataoka et al. (2003)
ebi	excess blue index	1.4b-g	Guijarro et al. (2011)
egi	excess green index	2g-r-b	Woebbecke et al. (1995)
egri	excess green minus excess red index	egi-eri=3g-2.4r-b	Meyer and Neto (2008)
eri	excess red index	1.4r-g	Meyer et al. (1999)
g	green chromatic coordinate	G/(R+G+B)	Woebbecke et al. (1995)
ga	green area index	sum of histogram frequencies of HSV hue channel, from 60° to 180°	Casadesús et al. (2007)
gga	greener area index†	sum of histogram frequencies of HSV hue channel, from 80° to 180°	Casadesús et al. (2007)
gli	green leaf index	(2G-R-B)/(2G+R+B)	Louhaichi et al. (2001)
H	hue	Hue is the color, often represented in units of degrees on a circle from 0° to 360°. Red primary starts at 0°, green at 120°, and blue at 240°.	
ipca	principal component analysis index	0.994|R-B|+0.961|G-B|+0.914|G-R|	Saberioon et al. (2014)
kai	kawashima index	(R-B)/(R+B)	Kawashima (1998)
LB	b* axis	Nonlinear transformations of RGB spectral reflectances to CIE Lab color space. CIE Lab b* values map to blue-yellow opponent colors, where negative values approach blue and positive values move in the direction of yellow.	
ngrdi	normalized green-red difference index	(G-R)/(G+R)	Tucker (1979)
r	red chromatic coordinate	R/(R+G+B)	Woebbecke et al. (1995)

S	saturation	Saturation represents the color richness, with values towards 0 representing grey and higher values showing more vivid color tone.	
vari	visible atmospherically resistant index	(G-R)/(G+R-B)	Gitelson et al. (2002)
vegi	vegetation index - a is a constant set to 0.667	$G_n/(R_n^{a}B_n^{1-a})$	Hague et al. (2006)
woi	Woebbecke index	(g-b)/(r-g)	Woebbecke et al. (1995)

[*] R_n, G_n, B_n are normalized ([0-1]) relative to their respective channel values. For this study, these values are calculated by dividing the channel value by its maximum, which is 255 for the 8-bit channel images captured in this study.
[†] The gga index avoids yellow-hued greens included by the ga index, which can indicate senescence.

Growing Degree Days

Temperature data was collected from past weather records using the National Weather Service (NWS) online data portal. Maximum and minimum temperatures were downloaded from NWS at the Mather 3 NW meteorological station (id USC00475164). This NWS station was selected for its proximity to both cranberry populations (see Figure III-1A) and for its comprehensive daily minimum and maximum temperature records for the years 2018 to 2020. Growing degree days (GDD) were calculated using an adjusted method from ambient temperature (Equation III-1).

Equation III-1. Calculation for adjusted growing degree day (*GDD*) and cumulative degree days (*CDD*) for American cranberry.

$$GDD_d = \begin{cases} 0, if\ DH < 13 \\ \bar{T}_d - \dot{T} \end{cases}$$

$\mathrm{CDD_N} = \sum_{\mathrm{d}=1}^{\mathrm{N}} \mathrm{GDD_d} = \mathrm{DD_{N-1}} + \mathrm{GDD_N}$Where DH is the total daylight hours (sunrise to sunset), $\bar{\mathrm{T}}_\mathrm{d}$ is the mean daily temperature ($\frac{\mathrm{T_{max}}+\mathrm{T_{min}}}{2}$), $\dot{\mathrm{T}}$ is the baseline temperature for cranberry, d

is the respective day of year for which $\mathrm{GDD_d}$ is calculated, and N is the day of the year for which $\mathrm{CDD_N}$ is calculated (day 1 corresponds to January 1st and resets every new year).

The reference $\dot{\mathrm{T}} = 5°C$ was determined to be an accurate and stable baseline temperature for key cranberry developmental stages when GDD is adjusted for daylength (Workmaster, 2001; Jeranyama and Kennedy, 2021). At the two locations in this study, the 13-hour daylight threshold occurred on April 6th for all years of data collection. Ambient temperatures measured by NWS stations are taken by instruments shaded from direct sunlight, acting as reasonable proxies for under-canopy temperatures. Cumulative growing degree days (CDD) were subsequently calculated as the accumulated GDD for each year.

Correlations

Pairwise Pearson's correlation coefficients were calculated for all VIs, CDDs, spectrophotometer absorbance values for A535, and absorbance ratios A660:A535, A649:A535, and A6XX:A535 (see Table III-2 and Figure III-3).

Model Fit and Prediction

Several linear regression methods and a simulated annealing (SA) model selection method were applied to fit models to each of the reference data types in Table III-2. These include principal components regression (PCR), partial least squares regression (PLSR), ridge regression (RR), and multivariate linear regression (ordinary least squares - OLS). The general regression model is described in Equation III-2. Metrics used to validate model quality were coefficient of determination (R^2), root mean square error (*RMSE*) and mean absolute error (*MAE*). The median value of each of these metrics and their 95% confidence intervals were derived from a distribution of 200 repeats of 10-fold cross-validation on training data.

Table III-2. Reference data types and descriptions for cranberry leaf tissue chlorophyll and anthocyanin levels. This reference data is derived from documented anthocyanin and chlorophyll pigment spectral absorbance curves (Wintermans and De Mots, 1965; Fuleki and Francis, 1968; Workmaster, 2001).

Values are unitless but recorded as absorbance units (AU) at defined spectral wavelength (nm), or are ratios between two absorbance readings.

Reference	Description
A535	Units are recorded as AUs relative to an acidified ethanol blank at spectral wavelength 535 nm (green). Analogous metric to the anthocyanin levels in cranberry leaf tissue. Values are normalized to the leaf tissue dried mass.
A649	Units are recorded as AUs relative to a 95% ethanol blank at spectral wavelength 649 nm (red). Representative measure for chlorophyll a and b levels in cranberry leaf tissue. Values are normalized to the leaf tissue dried mass.
A660	Units are recorded as AUs relative to a 95% ethanol blank.at spectral wavelength 660 nm (red). Representative measure for chlorophyll a and b levels in cranberry leaf tissue. Values are normalized to the leaf tissue dried mass.
A660:A535	As anthocyanin absorption curves also overlap with chlorophyll, the ratio of red:green can be a more useful metric to measure chlorophyll:anthocyanin levels compared to either pigment independently from spectral absorbances (Gamon and Surfus, 1999).
A649:A535	As anthocyanin absorption curves also overlap with chlorophyll, the ratio of red:green can be a more useful metric to measure chlorophyll:anthocyanin levels compared to either pigment independently from spectral absorbances (Gamon and Surfus, 1999).
A6XX:A535	A6XX is the mean of the A660 and A649 spectrophotometer absorbance values. As anthocyanin absorption curves also overlap with chlorophyll, the ratio of red:green can be a more useful metric to measure chlorophyll:anthocyanin levels compared to either pigment independently from spectral absorbances (Gamon and Surfus, 1999).

Equation III-2. General linear regression model form.

$$a_n = V_{n \times m} \cdot \beta_m + \epsilon_n$$

The response variable a_n is a column vector of n reference data observations (spectrophotometer absorbances in Table III-2), each element representing a distinct image capture observation of a reference field plot. $V_{n \times m}$ is a matrix of the computed mean VI values for each of n reference plot images, m VI predictors per observation. β_m is a column vector of regression coefficients associated with each of m VI values. ϵ_n represents the random noise, or error, and is normally distributed stochastic variable with a mean of zero and variance of σ^2.

To avoid unstable regression coefficients and unnecessary model complexity, redundant features were removed from all regression models. Specifically, *egi*, *rgb.G*, *lab.a*, and *ndi* were eliminated due to their high correlation with other features, identified using the CRAN caret package's findCorrelation() function with a 0.99 threshold (Kuhn, Max, 2008).

Compared to OLS, the PCR, PLSR, RR , and SA methods avoid overfitting by either reducing model complexity with ablation of predictors (sparsity) or through regularization. These methods create simpler and more generalizable models that perform better on untrained data.

The PCR and PLSR regression methods serve to generate model parsimony by compressing highly correlated predictors (PCR) and response (PLSR) variables into a condensed model of uncorrelated synthetic variables that ideally segregate model representation into fewer terms. PCR and PLSR regressions were performed from utilities in the R package `pls`, using 8 repetitions of 10-fold cross-validation to tune the number of components and measure model performance (Liland et al., 2023). The number of components for PCR and PLSR regressions were selected using the one standard error method on *RMSE* (Breiman et al., 2017), which selects the sparsest model within one standard error of the highest performing cross-validated model.

RR reduces model overfitting by introducing an L2-norm penalty term to the objective function, effecting better model performance through a bias-variance tradeoff. RR was performed using the `ridge()` function from the `fastmatrix` package, using a generalized cross-validation method to tune the lambda penalty factor when fitting the model (Osorio and Ogueda, 2024).

A simulated annealing (SA) method was performed both to select a subset of model predictors and to fit a reduced parsimonious model. SA searched the feature space iteratively by performing small perturbations on the features of the model, accepting performance degradations with diminishing probability as the iterations increase (Kirkpatrick et al., 1983)., thus avoiding suboptimal solutions within localized maxima. The SA regression was executed using the `safs()` function of the caret package (Kuhn, Max, 2008), using 5 repetitions of 10-fold cross validation to tune and estimate model performance.

Multiple UAV images of the same cranberry plot from different angles afforded repeated estimates of a plot's A535 value. Since these measurements are taken of the same plots under nearly identical conditions, their model predictions on untrained data should be highly consistent. The coefficient of variation (CV) measures this consistency - a low CV indicates that the model produces stable A535 predictions regardless of viewing angle, offering a complementary performance benchmark that approximates model prediction variance.

Selection Index

To develop a phenotype scoring index, predicted A535 values were plotted against cumulative growth degree days (CDD) for all cranberry plot images. Exponential decay functions were subsequently fit on a per population and a per genotype level, combining predicted values from 2018 and 2019. Scores were calculated as the conditional integral difference between the genotype-fit and population-fit curves, as defined below in Equation III-3. The last week of May in central Wisconsin, where most cranberries are grown, is a time when the probability of a frost event drops below 30% (National Centers for Environmental Information, 2025). For this study, the last week of May corresponds to around 200 accumulated heat units, denoted as $\mathrm{x_c}$ in Equation III-3. The accumulated integral difference before the $\mathrm{x_c}$ threshold was combined with the inverted accumulated integral difference after $\mathrm{x_c}$. This selection index, subsequently referred to as *leaf maturity index* (LMI), scores genotypes with delayed spring development higher: greater red pigmentation before the critical period threshold and less pigmentation after, favoring rapid developmental post-frost.

Equation III-3. Leaf maturity index (LMI) calculation, where $\mathrm{s_g(x)}$ is the score at CDD x, $\mathrm{\widehat{r_g}(x)}$ is the genotype-fit curve A535 value at CDD x, $\mathrm{\widehat{r_p}(x)}$ is the population-fit curve A535 value at CDD x, $\mathrm{x_c}$ is the critical frost period threshold CDD, $\mathrm{S_g}$ is the selection index of genotype g, $\mathrm{x_s}$ and $\mathrm{x_f}$ are the starting and ending CDD values that bound integration of $\mathrm{s_g(x)}$.

$$s_g(x) = \begin{cases} \widehat{r_g}(x) - \widehat{r_p}(x), x \leq x_c \\ \widehat{r_p}(x) - \widehat{r_g}(x), x > x_c \end{cases}$$

$$S_g = \int_{x_s}^{x_f} s_g(x) d_x$$

Results

Correlations

Pairwise correlations among Table III-1's SRs and VIs, observed A535 nm spectrophotometer absorbances, and CDD are displayed below in Figure III-3 A-C. All pairwise SR correlations are available in Supplementary Table A-1 and Supplementary Table A-2.

Figure III-3A shows rank-sorted pairwise correlations between VIs and CDD to highlight VIs associated with a traditionally important metric in modeling cranberry physiology and phenology. In addition to VIs, A535 is also included in Figure III-3A to highlight the relationship between leaf anthocyanin content and CDD. Features with strong CDD pairwise correlation ($|r| > 0.75$) are *cive*, *LB*, *ac.cb*, *gli*, *egri*, *vari*, A535, *eri*, and *ebi*.

Ordered pairwise correlations of VI and A535 in Figure III-3B emphasize those features that have high association with cranberry spring leaf color development. CDD is also included in Figure III-3B to demonstrate the connection between an important metric used in modeling cranberry development and A535. Prominent A535 covariates ($|r| > 0.75$) are *ac.cb*, *gli*, *egri*, *cive*, *ebi*, *eri*, *vegi,* and CDD, *LB*, and *vari*.

The strength of association between VIs, calculated as the absolute value of correlation coefficients, are illustrated in the Figure III-3C heatmap. VIs are ordered using the first order principal component of the strength of associations using the `corrplot()` function in the corrplot CRAN package (Wei and Simko, 2021). VI associations are generally higher in the upper left, with a decreasing gradient from top-left to bottom-right. The cluster encompassing *gli*, *ac.cb*, *ebi*, *egri*, *cive*, *eri*, *vari*, and *LB* is a particularly tight grouping of highly correlated VIs.

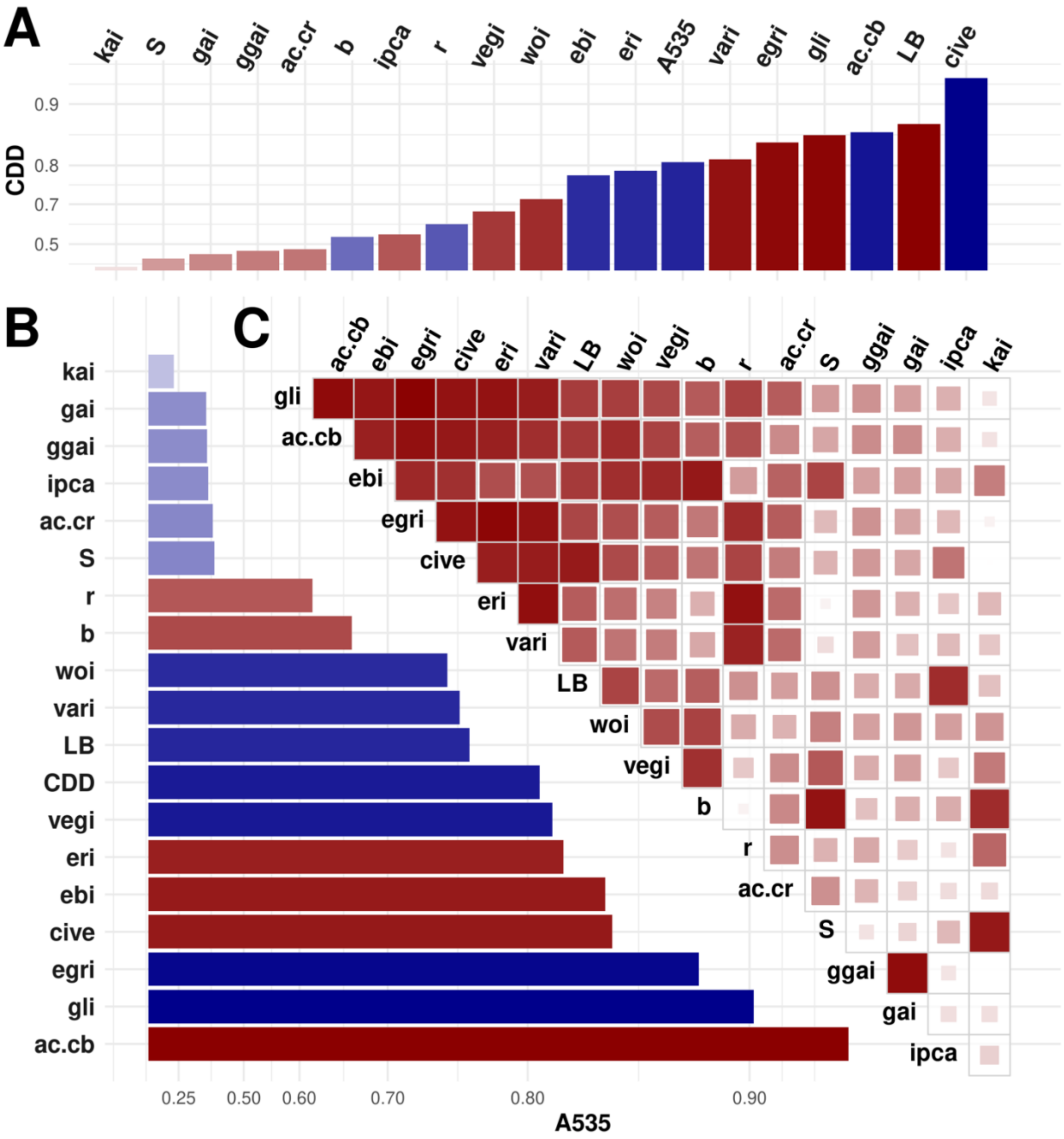


Figure III-3. Pearson's correlations across three cranberry populations. A. Pairwise CDD-by-trait correlations ranked by strength of association. B. Pairwise A535-by-trait correlations ranked by strength of association. A.-B. Non-significant ($p \geq 0.05$) correlations are labeled with "NS". Correlations are displayed and sorted as absolute values and are colored according to their correlation sign – positive correlations are red, while negative correlations are blue. The correlations are scaled exponentially to better differentiate stronger associations. C. Heatmap of absolute pairwise correlations between VIs.

Model Fit

Model fit benchmarks for all five models are displayed below in Figure III-4. The regression models PCR, PLSR, RR, and OLSE all have nearly identical predictive power, with *RMSE* values around 0.2 and R^2 around 0.9. They vary in the number of features used to fit the model, with both PCR and PLSR fitting 16 synthetic variables each, while RR and OLSE fit all specified input predictors. The SA-derived model exhibits a modest performance decline (median R^2, RMSE, and MAE), although its RMSE and MAE prediction error variances overlap substantially with their corresponding metrics in other models. The SA method narrowed the number of predictors from 19 down to 7, producing a more parsimonious model that trades insignificant performance declines for higher interpretability and stable predictions.

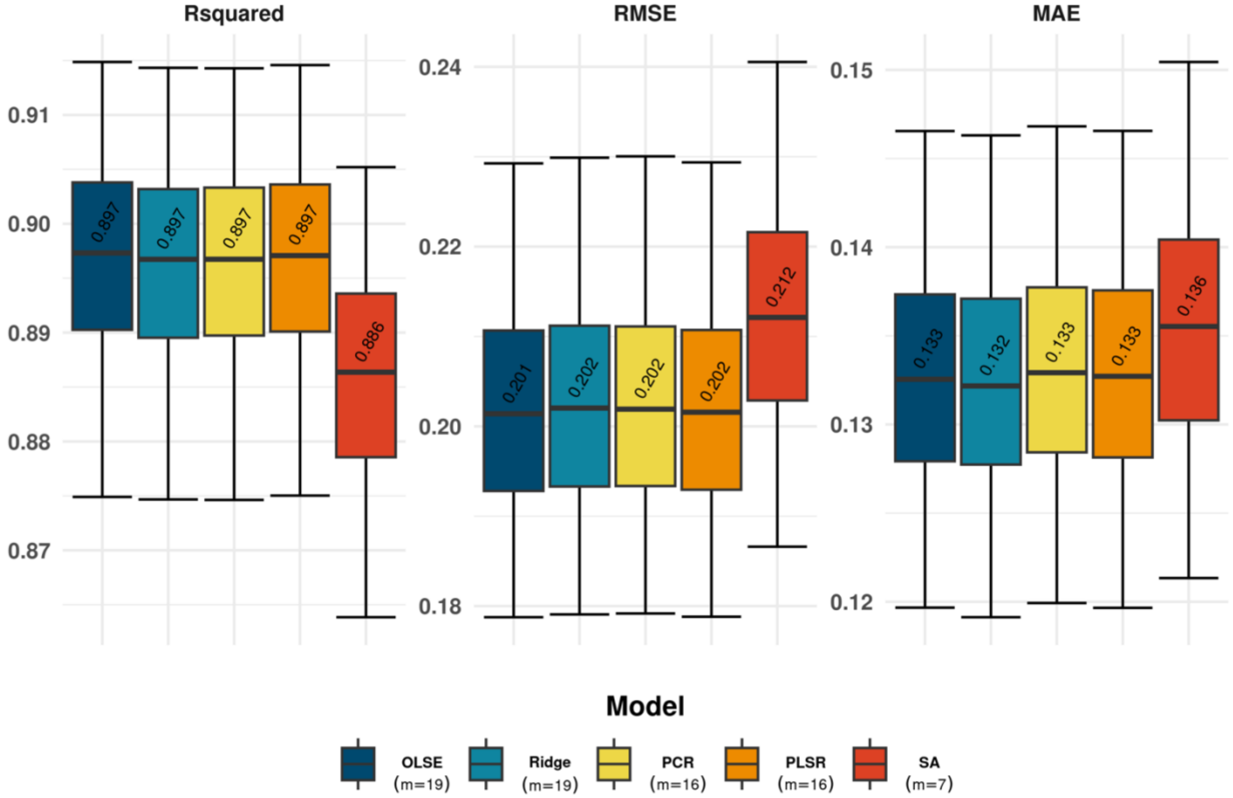


Figure III-4. Fit benchmarks across five regression model types in three cranberry populations. Each column's error bar represents the 95% confidence interval of the respective benchmark after 10-fold cross-validation repeated 200 times. The horizontal black line with corresponding text indicates the median value and the column extents representing the 25% and 75% quartiles. R^2 is the model coefficient of determination, *RMSE* is the root mean square error, and MAE is the mean absolute error. Number of features (*m*), displayed in the legend for each model type, represents the number of predictor variables used to fit the model. Predictors are synthetic principal components for PCR and PLSR or relevant SRs and VIs for RR, OLSE, and SA methods.

Figure III-5 displays the SA model parameter estimates. Model coefficients and their 95% confidence intervals are shown in Figure III-5A, along with their corresponding p-values from the ANOVA. Effect size estimates ($\widehat{\eta_p^2}$) for each predictor, representing their predictive impact on the model, are shown in Figure III-5B. All model coefficients are highly significant ($p < 0.001$). Regarding predictive efficacy, VIs *gli* ($\widehat{\eta_p^2} \approx 0.88$) and *ac.cb* ($\widehat{\eta_p^2} \approx 0.35$) contribute substantially to the model, while *S* ($\widehat{\eta_p^2} \approx 0.04$), *cive* ($\widehat{\eta_p^2} \approx 0.03$), and *ipca* ($\widehat{\eta_p^2} \approx 0.02$) demonstrate moderate influence. Marginal effects are contributed by predictors *gai* and *egri*, but they benefit model performance by keeping repeated measure variance low (Figure III-6).

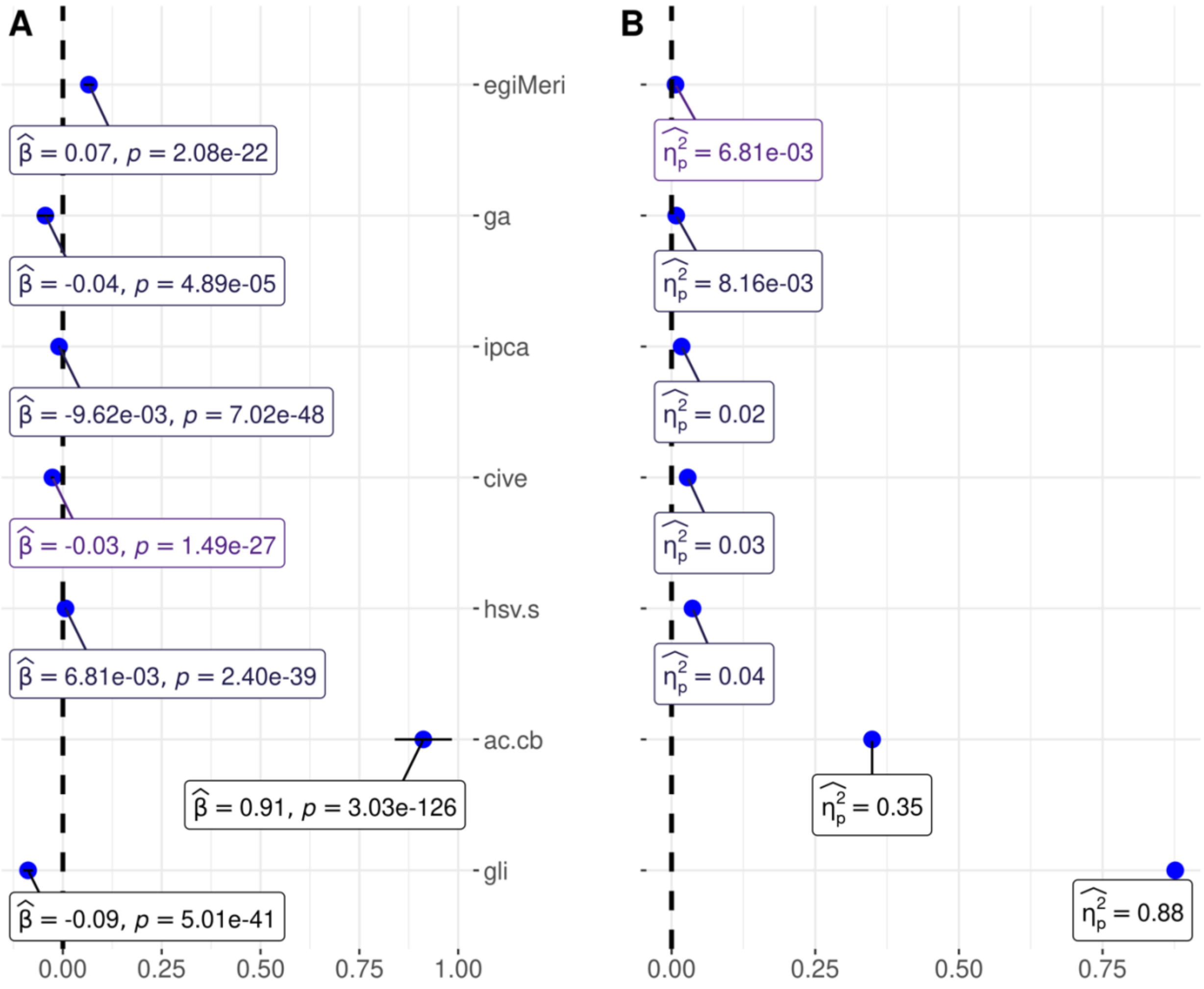


Figure III-5. Model fit parameter estimates for simulated annealing (SA) selected vegetative indices (VIs) across three cranberry breeding populations. A. SA regression coefficient estimates ($\widehat{\beta}$) for the VIs listed in the y-axis along with the coefficient p-values. The horizontal bars represent the 95% confidence interval of estimates for each predictor coefficient. B. ANOVA effect size estimates ($\widehat{\eta_p^2}$) of each VI, where larger relative values indicate higher meaningful effect contributions. Figure was generated using the ggcoefstats() function of the ggstatsplot R package (Patil, 2021).

Model distributions of repeated predicted A535 CV values are shown in Figure III-6, which includes median A535 CVs, statistical hypothesis test calculations, effect size estimates, and statistical groupings. A one-way ANOVA nonparametric Friedman rank-sum test demonstrated strongly significant statistical differences ($p \ll 0.001$) between regression methods for repeated measure predictability (Friedman, 1937, 1939, 1940). Statistical groupings between regression methods were determined using Holm-adjusted pairwise p-values with a differential threshold below

0.05 (Holm, 1979). Kendall's coefficient of concordance (Kendall and Smith, 1939), a non-parametric measurement of how reliably different groups can be distinguished using different scorers, was estimated at 0.02 for CV.

The SA model had the lowest median CV ($\hat{u}_{median}$) of 5.91 and the smallest spread in CV. The RR, OLSE, and PCR models were all tied for the largest $\hat{u}_{median}$ of 6.20, with RR having the largest variance in repeated measure CV. RR and SA each form their own distinct statistical groups, while OLSE, PLSR, and PCR all are within the same statistical group ($p < 0.05$).

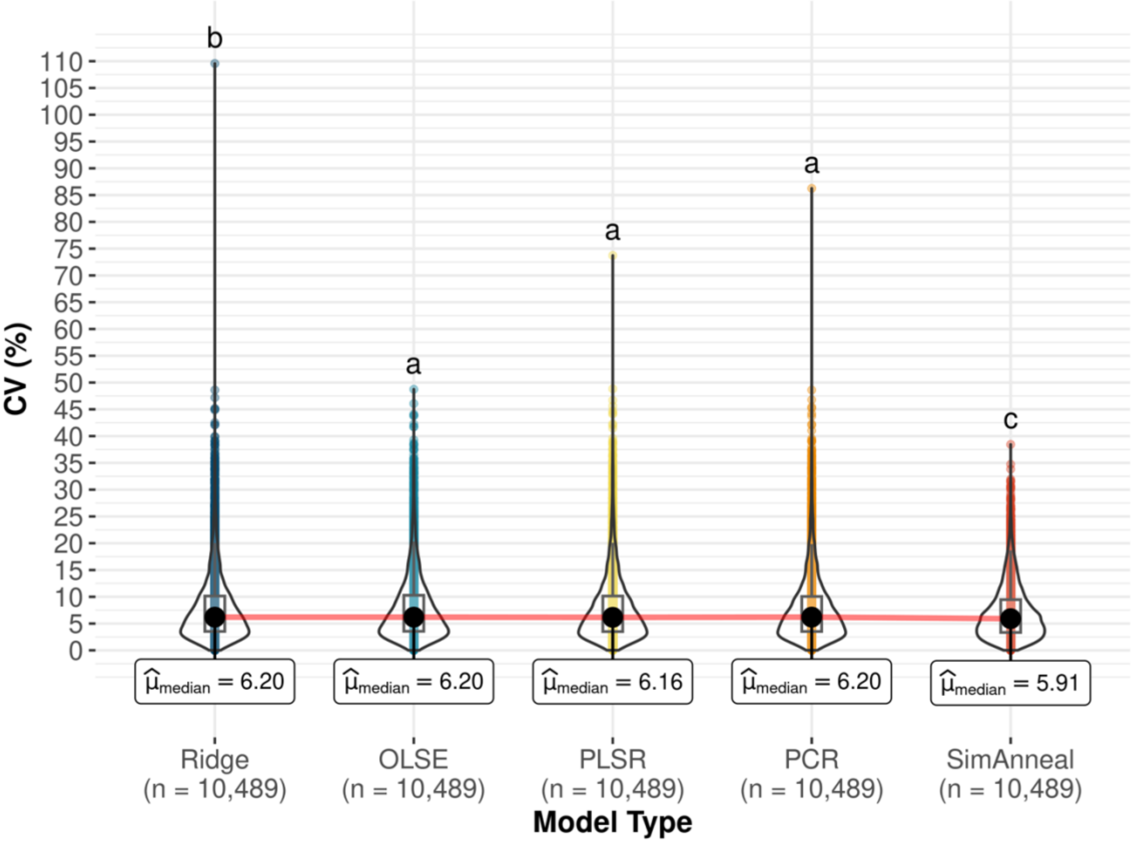


Figure III-6. Coefficient of variation (CV) of regression A535 predictions across repeated photos of the same plot genotype within an imaging session. Plot was generated using the ggwithinstats() function of the ggstatsplot R package (Patil, 2021). CV was calculated across all three cranberry populations of this study. Statistical groupings between regression methods were determined using Holm-adjusted (Holm, 1979) pairwise p-values, with group membership designated using letter characters and distinctions set at $p < 0.05$.

SA regression fit plots are displayed below in Figure III-7 for all training data. The dashed line indicates the ideal fit line, where observed A535 values equal model-predicted A535. Figure III-7A shows the fit plot for populations CNJ02 and CNJ04 at the Cranberry Creek marsh, while panel B displays the fit plot for population GRYG at the Valley Corporation marsh. The SA model demonstrates reduced prediction accuracy when observed A535 absorbances are greater than two.

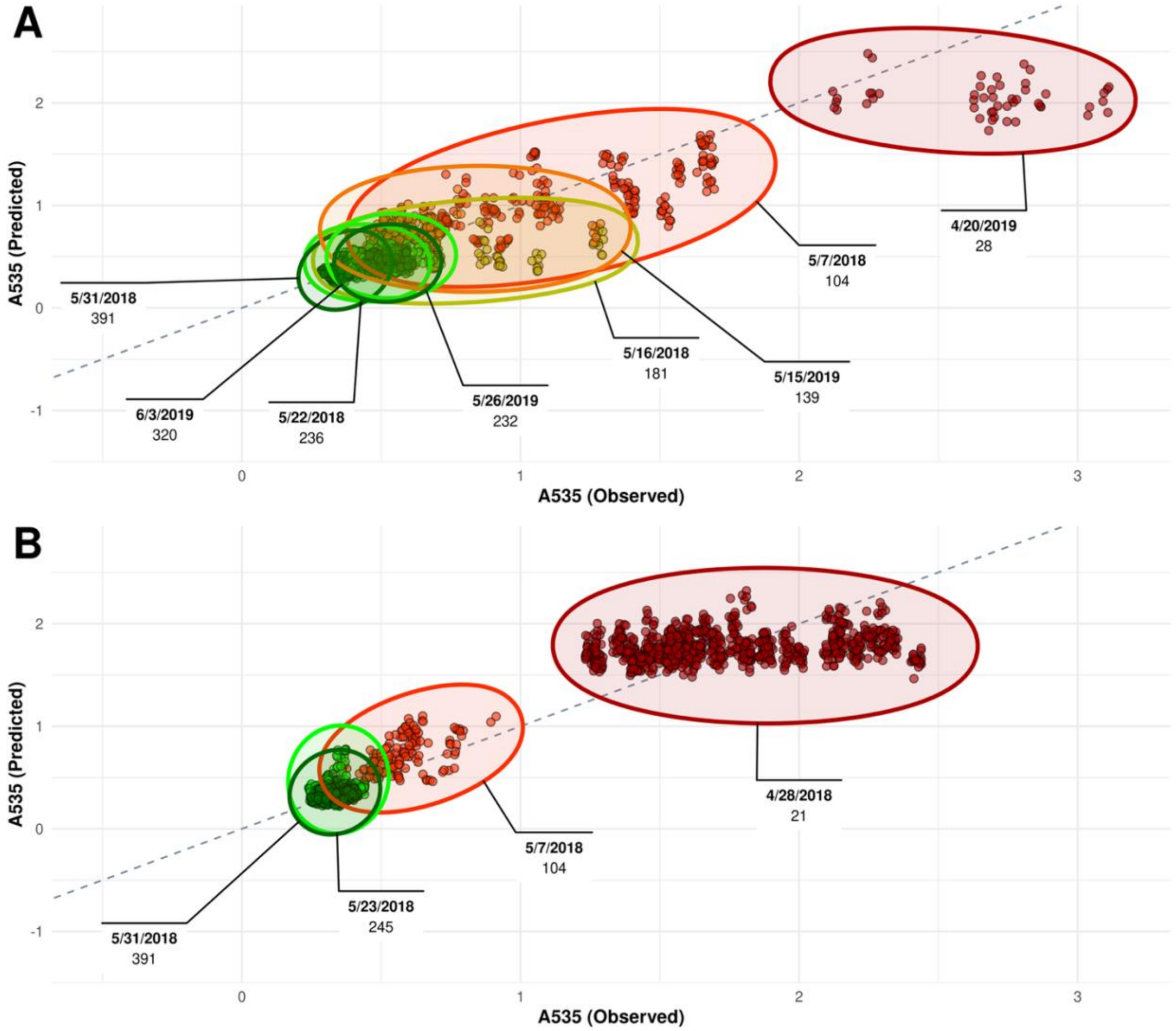


Figure III-7. SA regression of three cranberry populations - fitted vs. observed plots. Panel A represents the regression fit for populations CNJ02 and CNJ04. Panel B illustrates the regression fit for population GRYG. Ellipses circumscribe all datapoints associated with a given UAV flight date. The colors associated with the ellipses and datapoints are graded using a linear function mapping CDD inputs to informative color outputs: low CDD values map to dark red and high CDD values map to dark green. Each ellipse is labeled with its associated date in month-day-year format and its associated CDD.

Developmental Progression

To illustrate that the SA regression model shows visual consistency between predicted A535 values for plots and the qualitative redness of the plots, SA predictions for all plots were sorted from highest to lowest predicted A535. Representative plot images were selected near quantiles 1 (top), 0.75 (upper quartile), 0.5 (median), 0.25 (lower quartile), and 0 (bottom). Figure III-8 below illustrates two representative plot images for each quantile. The developmental progression, from top (left) to bottom (right) predicted A535, is visually consistent with the expected phenotypic expression of cranberry leaf redness.

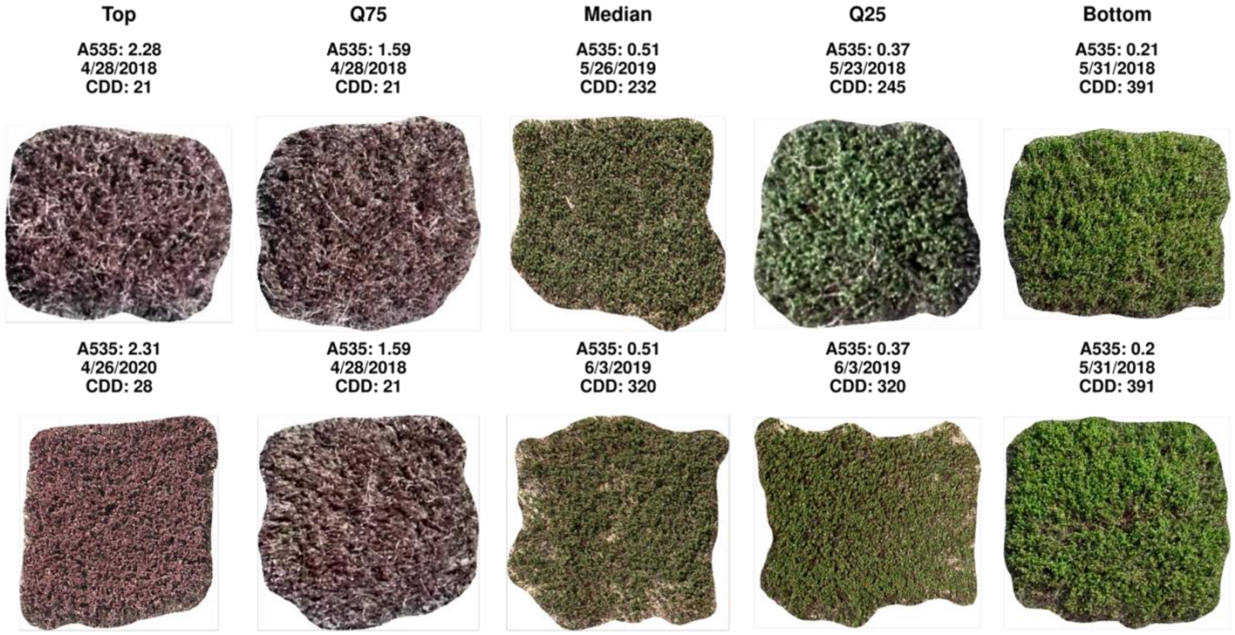


Figure III-8. Sample UAV images from three cranberry populations representing gradient of top quintile to bottom quintile of SA model predicted A535. Each column represents a sample of two plot images within that column's associated quintile. Quintiles are ordered highest to lowest, from left to right. Quintiles are labeled top (quintile 1), Q75 (upper quartile), median (quintile 0.5), Q25 (lower quartile), and bottom (quintile 0). Each plot is labeled with its predicted A535, the date at which the plot image was captured (in month-day-year format), and the calculated CDD on that image date.

Genotype Performance

Top- and bottom-ranking CNJ02 and GRYG genotypes per the selection index defined in Equation III-3 are illustrated below in Figure III-9 to Figure III-11. In each figure, a representative low-scoring genotype is shown under a high-scoring genotype to highlight their developmental differences. Genotypes ranked in either the top 5% or top 10 selections (whichever is lower) of selection index scores are displayed in Table III-3, while the lowest 5% or bottom 10 selections are displayed in Table III-4.

CNJ02-1-70 and CNJ02-1-67 are representative top- and bottom-scoring genotypes displayed in Figure III-9 (2018) and Figure III-10 (2019). The LMI scores for CNJ02-1-70 and CNJ02-1-67 are 66.8 and -57.6, respectively. Genotype fit curves in Figure III-9 and Figure III-10, columns *A*, rows *2-3*, show the predicted A535 developmental differences and visually differentiate the LMI scores between the genotypes. Figure III-9, column *B* (*CDD = 104*) and Figure I-10, column *C* (*CDD = 139*) demonstrate snapshots of the development timeline where there is a larger range of segregation in predicted A535. Other developmental timepoints (and their associated CDD values) exhibit tighter predicted A535 distributions, exhibiting less segregation among genotypes.

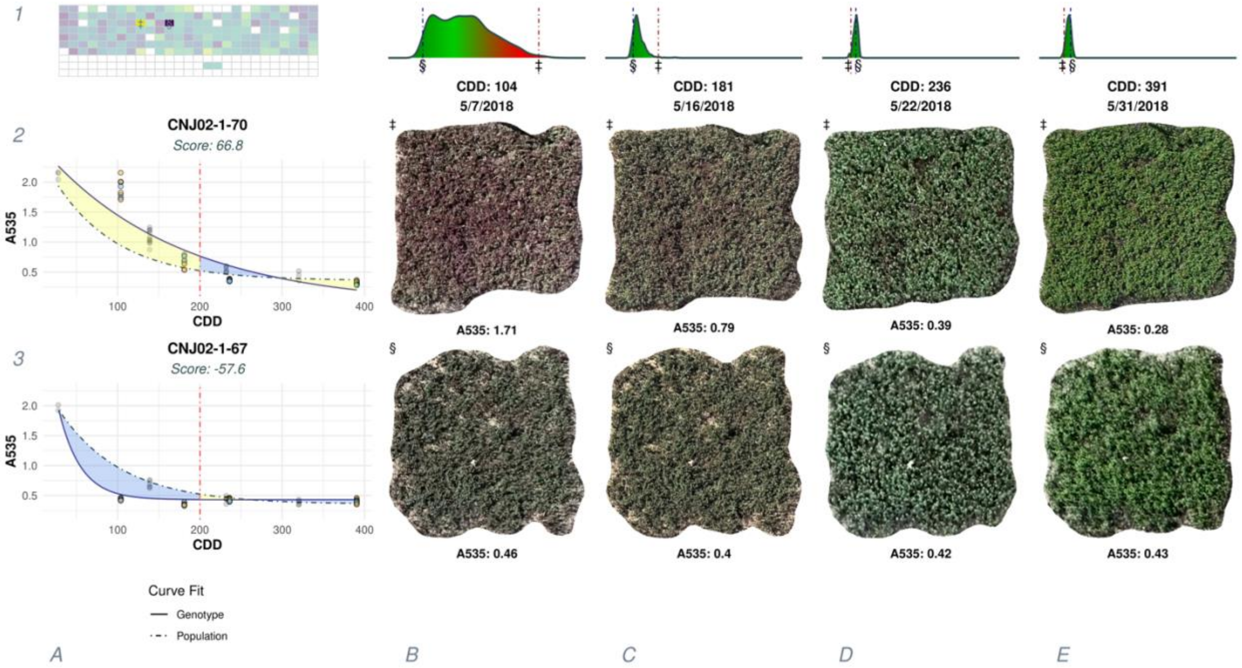


Figure III-9. Cranberry population CNJ02 leaf-color development for top- (row *2*) and bottom-performing (row *3*) genotypes in 2018. Column *A* contains functional graphs of the predicted A535 versus CDD for each genotype. These plots contain the exponential decay curve fit functions for the genotype datapoints (solid line), population datapoints (dashed line), and color-coded area between the two curve types. Yellow indicates an area that positively contributes to the selection index score, while blue indicates a negative contribution. The dashed vertical line at CDD equal to 200 heat units demarcates the inflection point at which the integral area difference function is inverted (Equation III-3). Each graph is labeled with its respective genotype exponential decay fit function and the fit curve's associated R^2, while the population's decay fit curve is described in row *1*, column *A*. Row *1*, columns *B-E* show the probability distributions of the predicted A535 for a given genotype, with labeled vertical dashed lines indicating the location of the top- (‡) and bottom-performing (§) genotypes. Each distribution graph is colored according to its A535 bin value: redder values indicating higher A535 and greener indicating lower A535. Each column of *B-E* corresponds to a given CDD value and date (month-day-year), with a representative plot image for each genotype at that timepoint. Columns *B-E* are sorted with increasing CDD from left to right to indicate leaf developmental progression. Each representative plot image is labeled below with its predicted A535. All dates are formatted as month/day/year.

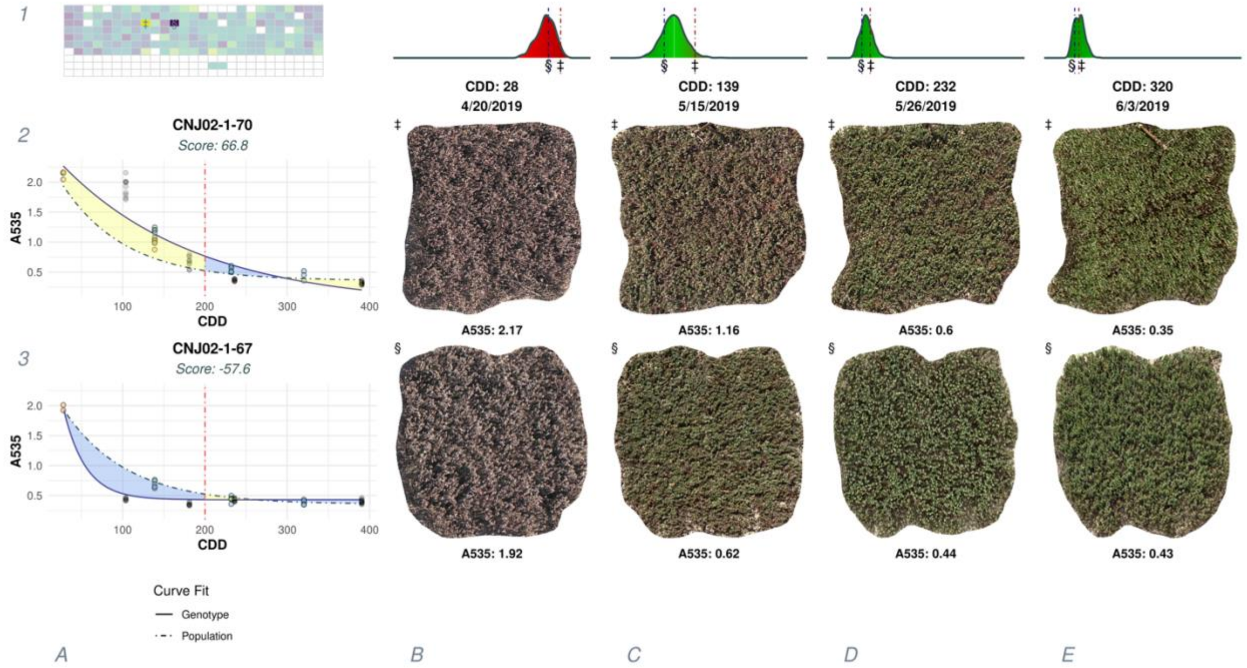


Figure III-10. Cranberry population CNJ02 leaf phenology progression for top- (row *2*) and bottom-performing (row *3*) genotypes in 2019. Column *A* contains functional graphs of the predicted A535 versus CDD for each genotype. These plots contain the exponential decay curve fit functions for the genotype datapoints (solid line) and population datapoints (dashed line), along with color-coded area between the two curve types. Yellow indicates an area that positively contributes to the selection index score, while blue indicates a negative contribution. The dashed vertical line at CDD equal to 200 heat units demarcates the inflection point at which the integral area difference function is inverted (Equation III-3). Each graph is labeled with its respective genotype exponential decay fit function and the fit curve's associated R^2, while the population's decay fit curve is described in row *1*, column *A*. Row *1*, columns *B-E* show the probability distributions of the predicted A535 for a given genotype, with labeled vertical dashed lines indicating the location of the top- (‡) and bottom-performing (§) genotypes. Each distribution graph is colored according to its A535 bin value: redder values indicating higher A535 and greener indicating lower A535. Each column of *B-E* corresponds to a given CDD value and date (month-day-year), with a representative plot image for each genotype at that timepoint. Columns *B-E* are sorted with increasing CDD from left to right to indicate leaf developmental progression. Each representative plot image is labeled below with its predicted A535. All dates are formatted as month/day/year.

P40 and P97 are characteristic top- and bottom-ranking GRYG genotypes displayed below in Figure III-11. The two graphs of column *A*, rows *2-3*, illustrate developmental distinctions between the two genotype fit curves, with selection index scores of 72.2 for P40 and -47.1 for P97. This phenological differentiation is mirrored in the plot photo progressions between P40 and P97 (rows *2*-*3*, columns *B*-*E*). The distribution of predicted A535 values in row *1* shows wider variance, and hence more segregation, between phenotypes at timepoint 5/7/2018 (column *C*).

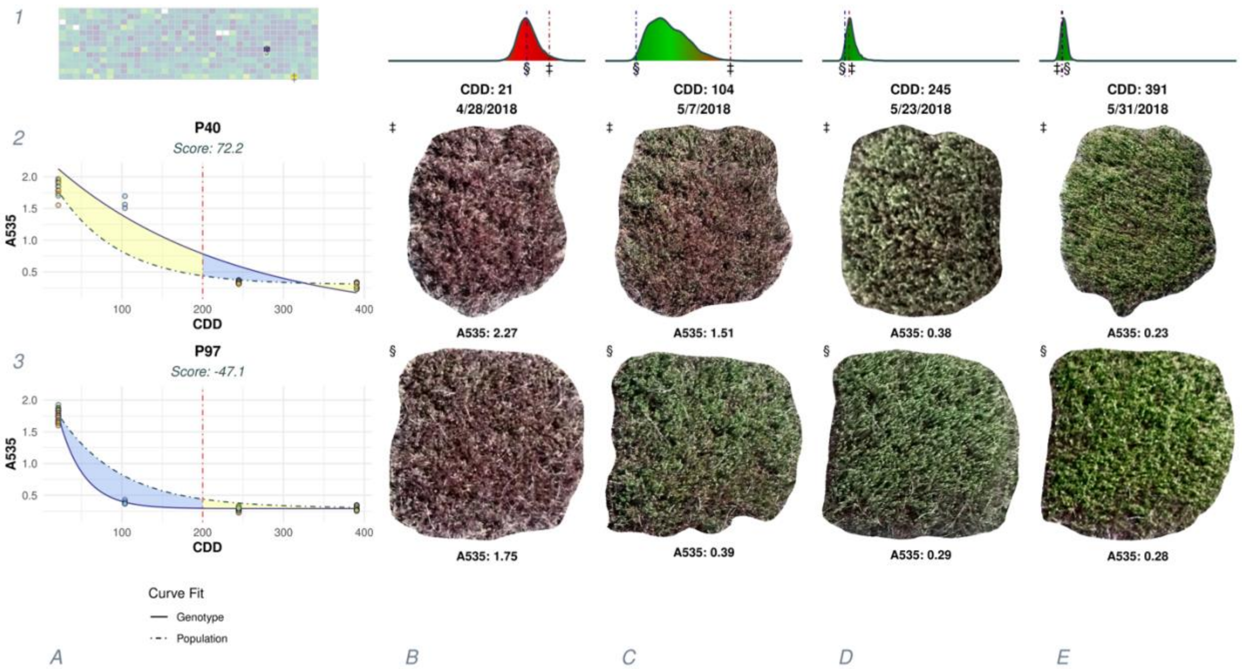


Figure III-11. Cranberry population GRYG leaf phenology progression for top- (row *2*) and bottom-performing (row *3*) genotypes in 2018. Column *A* contains functional graphs of the predicted A535 versus CDD for each genotype. These plots contain the exponential decay curve fit functions for the genotype datapoints (solid line) and population datapoints (dashed line), along with color-coded area between the two curve types. Yellow indicates an area that positively contributes to the selection index score, while blue indicates a negative contribution. The dashed vertical line at CDD equal to 200 heat units demarcates the inflection point at which the integral area difference function is inverted (Equation III-3). Each graph is labeled with its respective genotype exponential decay fit function and the fit curve's associated R^2, while the population's decay fit curve is described in row *1*, column *A*. Row *1*, columns *B*-*E* show the probability distributions of the predicted A535 for a given genotype, with labeled vertical dashed lines indicating the location of the top- (‡) and bottom-performing (§) genotypes. Each distribution graph is colored according to its A535 bin value: redder values indicating higher A535 and greener indicating lower A535. Each column of *B*-*E* corresponds to a given CDD value and date (month-day-year), with a representative plot image for each genotype at that timepoint. Columns *B*-*E* are sorted with increasing CDD from left to right in order to indicate leaf developmental progression. Each representative plot image is labeled below with its predicted A535. All dates are formatted as month/day/year.

Table III-3. Top 5 percent of genotypes (up to maximum of 10 genotypes) per cranberry population using the selection index defined in Equation III-3. The genotypes are ranked from higher to lower from top to bottom, with the associated selection index score to the right of each genotype.

GRYG		CNJ02		CNJ04	
P40	72.2	CNJ02-1-70	66.8	CNJ04-2-14	29.8

GRYG		CNJ02		CNJ04	
P131	59.3	CNJ02-1-191	59.7	CNJ04-2-8	27.8
P461	59.0	CNJ02-1-14	55.9	CNJ04-2-29	26.6
P417	52.9	CNJ02-1-201	53.4		
P431	51.5	CNJ02-1-32	43.9		
P374	48.6	CNJ02-1-20	42.5		
P338	47.7	CNJ02-1-4	42.3		
P53	47.5	CNJ02-1-15	41.9		
P212	47.2				
P391	44.6				

Table III-4. Bottom 5 percent of genotypes (up to maximum of 10 genotypes) per cranberry population using the selection index defined in Equation III-3. The genotypes are ranked from lower to higher from top to bottom, with the associated selection index score to the right of each genotype.

GRYG		CNJ02		CNJ04	
P190	-53.1	CNJ02-1-51	-60.5	CNJ04-21-28	-43.5
P97	-47.1	CNJ02-1-235	-58.6	CNJ04-21-8	-34.8
P359	-45.9	CNJ02-1-67	-57.6	CNJ04-2-20	-34.7
P343	-45.8	CNJ02-1-129	-57.0		
P269	-44.7	CNJ02-1-42	-56.0		
P94	-43.9	CNJ02-1-78	-53.3		
P150	-43.9	CNJ02-1-68	-51.6		
P363	-42.8	CNJ02-1-123	-50.9		
P356	-42.7				
P434	-42.6				

Discussion

As an evergreen broadleaf, cranberries retain most of their leaves through the winter. These leaves display a green to red color transition when entering their cold dormancy period, followed by a red to green transition in spring's exit from dormancy. This color transition affords a simple and accessible metric of cranberry spring development, providing a trait that can be selected on in breeding populations. Cranberry flower buds are susceptible to late spring frosts, which can dramatically impact fall yields. Selecting late, rapidly developing genotypes represents a powerful tool to avoid crop damage using genetics.

This research sets out to assess the utility of low-cost RGB UAS to monitor leaf color development in cranberry breeding populations. Its goal is to offer an efficient and unbiased alternative to human scorers, unlocking the potential to phenotype larger cranberry populations with high temporal and spatial resolution.

Trait Patterns

As a key metric governing physiological development in cranberry, CDD's relationship to SRs and VIs informs major image features associated with the spring phenology of cranberry leaf color (Figure III-3A). All the pairwise traits listed in Figure III-3A are consistent with empirical (formulas in Table III-1) and theoretical expectations.

The strong positive correlation of CDD and *LB* emerges out of predictable developmental patterns in leaf pigments as spring progresses and heat units accumulate. In early spring, when CDD heat value units are low, cranberry leaf anthocyanin levels are high and chlorophyll levels are low (↑r, ↓g). Anthocyanins exhibit higher absorption in the green region (500-565 nm) relative to other parts of the visible spectrum, expressed as red or violet coloration in plant tissues enriched in this pigment. Later in spring, steep increases in CDD are accompanied by leaf anthocyanin declines and chlorophyll increases (↓r, ↑g). As chlorophyll absorbs light preferentially in the blue and red range, the relative reflected proportion of incoming blue light decreases while reflected yellow light levels remain stable ($y = r + g$; ↓r, ↑g). *LB* is defined as a blue-yellow opposition color metric whereby yellow values are more positive and blue values are more negative. Stable yellow but decreasing blue levels as CDD accumulates would manifest an increase in *LB*, explaining the positive correlation between CDD and *LB*.

Of the remaining VIs that strongly correlate ($|r| > 0.75$) with CDD, their magnitude and direction of correlation emerges from the spring temporal dynamics of their associated function inputs (↓r, ↑g, ↓b; Table III-1's formula column). VIs with a strong positive relationship ($r \geq -0.75$) include *gli*, *egri*, and *vari*. The VI *gli* ($r \approx 0.85$) has a numerator of 2G-R-B, and with an increase in green and decreases in red and blue as development progresses, the numerator would increase relative to the denominator. The strong positive covariance of *egri* ($r \approx 0.84$) with CDD comes directly out of its formula: 3g-2.4r-b. The VI *vari*'s substantial correlation ($r \approx 0.84$) with CDD manifests from a rapidly increasing numerator and a nearly stable denominator ((G-R)/(G+R-B)). VIs with a strong inverse relationship ($r < -0.75$) to CDD include cive ($r \approx -0.93$), ac.cb ($r \approx -0.86$), eri ($r \approx -0.79$), and ebi ($r \approx -0.78$). Of these, ac.cb, a VI developed to differentiate anthocyanin levels, aligns with observed anthocyanin patterns. Additionally, the response variable A535, a measurement of anthocyanins and a key indicator of the red pigmentation in cranberry leaves, has a pairwise correlation of -0.81 with CDD. This aligns with the exponential decline in red pigmentation that accompanies the sharp rise in CDD during late spring (Figure III-1B).

A few VIs selected in the SA model demonstrated modest correlation with CDD. These include the vegetation index (vegi; $r \approx 0.67$), principal component analysis index (ipca; $r \approx 0.57$), HSV's hue channel (H; $r \approx -0.49$), and green area index (gai; $r \approx 0.39$). As the sum of histogram frequencies of the HSV hue channel for hue values between 60° and 180°, gai represents the greenness of the cranberry plot. Its modest correlation ($0.25 \leq r < 0.75$) with CDD tracks observed developmental increases in chlorophyll. Spectrophotometer absorbances A649 and A660, proxies for chlorophyll levels, also show similar correlation patterns with CDD ($r \approx 0.36$ and $r \approx 0.41$, respectively; Supplementary Table A-2). As for the HSV colorspace, saturation is related to chroma, or color richness, with smaller values exhibiting higher levels of grey. A possible reason for the positive correlation between the S channel and CDD is that as cranberry leaves start to green and grow in tandem with spring's progression, they cover more bare ground and grey branches, enriching the colors of the pixels circumscribed by the plot boundary. A more robust segmentation of leaves and uprights from the background soil and branches would likely reduce this correlation, as the lower-chroma elements are removed from the analysis.

For pairwise correlations of SR and VI traits with A535, these traits all exhibit a negative, inverse correlative relationship respective to CDD, with a Spearman rank correlation of -0.95 of A535 to CDD. These inverted associations emerge from the strong negative correlation between A535 and CDD. Interestingly, given A535 as a proxy for anthocyanins, and their association with red, the correlation of A535 with the VI r ($\mathrm{r} \approx 0.62$) would be expected to be stronger. However, anthocyanins are also associated with purples, combinations of reds and blues, and VIs that include red and blue in their formula, like ac.cb ($\mathrm{r} \approx 0.93$), gli ($\mathrm{r} \approx -0.90$), egri ($\mathrm{r} \approx -0.88$), cive ($\mathrm{r} \approx 0.84$), and ebi ($\mathrm{r} \approx 0.84$), demonstrate some of the most pronounced correlative coupling with A535.

Model Characteristics

The comparative analysis of model performance illuminates key considerations in balancing predictive accuracy with practical utility for leaf color prediction. The more complex models (PCR, PLSR, RR, and OLSE) achieve marginally higher accuracy metrics, but their additional computational complexity, modestly higher relative variance on repeated measures, and reduced interpretability present practical tradeoffs that warrant careful consideration.

The SA-derived linear regression model demonstrates that a carefully selected subset of features can achieve comparable predictive power to more complex approaches. The minimal difference in predictive accuracy (reduction in median R^2 of 0.011) coupled with higher predictive stability on untrained data suggests that the SA model captures the essential predictive relationships while avoiding potential overfitting. This enhanced generalizability is particularly valuable for field applications where imaging conditions may vary from those encountered during model development.

The effectiveness of the simplified SA model can be understood by examining the underlying structure of the vegetation indices (VIs) used in this analysis. The high degree of correlation among many VIs reflects their derivation from the same basic spectral components - primarily the red, green, and blue channels. This inherent redundancy in the feature space means that additional indices often contribute overlapping information rather than unique predictive value. While multicollinearity doesn't necessarily impair prediction accuracy during controlled testing, it can significantly impact a model's reliability when deployed in varied field conditions where input values may deviate from training data distributions. Furthermore, multicollinearity can impair the interpretability of the model predictor coefficients and their relative importance.

Figure III-5 provides insights into the contribution of each predictor variable in the SA model. Coefficient estimate magnitudes do not necessarily equate with the relative importance of their associated predictors to the model. This downside is compensated for by ANOVA effect size estimates ($\widehat{\eta_{\mathrm{p}}^{2}}$). The ANOVA effect size is normalized to values between 0 and 1, with higher values representing more meaningful relative contributions of each coefficient to the model's determinative variation.

The VI *gli* demonstrates the highest $\widehat{\eta_{\mathrm{p}}^{2}}$ at 0.88, followed by *ac.cb* ($\widehat{\eta_{\mathrm{p}}^{2}} \approx 0.35$). Despite being a VI developed to differentiate anthocyanin levels in RGB images (Del Valle et al., 2018), *ac.cb* $\left(\frac{\mathrm{B+R}}{\mathrm{G}}\right)$ is secondary to *gli* $\left(\frac{\mathrm{2G-R-B}}{\mathrm{2G+R+B}}\right)$ in its contribution to the model's predictive output. The VI *gli* was developed as a way to quantify animal grazing impacts on wheat canopy coverage by emphasizing green leaf tissues relative to soil background (Louhaichi et al., 2001). Part of *gli*'s higher relative predictive impact on anthocyanins (A535) may come out of its utility to differentiate the green channel from reds and blues while normalizing over all channels. This normalizing lends to a more continuous and stable output on the margins of the channel values, a problem with the *ac.cb* VI, which starts to saturate as $G \rightarrow 0$, no matter the value of its numerator. Moreover, as a differential VI that incorporates all three RGB channels, *gli* lends predictive power to anthocyanin levels while avoiding issues related to chlorophyll level saturation with RGB imaging.

Although the remaining SA model predictors (*S, cive, ipca, gai, and egri*) have minor effect sizes compared to *gli* and *ac.cb*, they demonstrate meaningful impact on lowering model prediction error variance. Interestingly, while *egri* has a high pairwise correlation with A535 (Figure III-1), it has the lowest estimated model effect impact (Figure III-5B). This may be a consequence of the high pairwise correlation of *egri* with both *gli* and *ac.cb*, such that the components of *gli* and *ac.cb* that contribute to the model are mostly redundant with *egri*. A similar phenomenon is likely influencing the low effect impact of *cive*, which also demonstrates a low effect estimate while having a high pairwise correlation with *gli* and *ac.cb*. The *ipca*'s small effect size is more likely a consequence of its low correlation with A535 than it providing redundant information, as it has low pairwise correlations with *gli* and *ac.cb*. Though it exhibits a minute predictive effect size, *gai* is likely important to the model not only for it acting as a proxy for green tissue, but also because it offers a unique metric derived from spectral band counts of the HSV hue channel histograms. This is evident by its low correlation to all other VIs in Figure III-3C, outside of *ggai*, which is derived from *gai* but with a narrower bandwidth.

The HSV *S* channel's modest but meaningful impact possibly is a consequence of background tissue and soil 'noise', as these image features are likely to have lower *S,* or chroma values. As the spring season progresses, these background features are covered with a flush of new green tissue, producing a modest negative effect of *S* relative to predicted A535. *S* as a predictor may not be problematic to model accuracy if it applies universally to all field plots, but some plots have less coverage (more background soil), so *S* may have less consistent prediction impact between plots.

Better segmentation of soil background could produce a model with the *S* predictor that exhibits lower prediction error variance.

Genotype Performance

Our findings demonstrate that a simple UAS equipped with an RGB camera can effectively monitor cranberry leaf anthocyanin levels during spring development. To leverage this capability, we developed a novel selection index for scoring genotypes based on their spring development patterns. This index specifically favors genotypes that maintain their red leaf coloration longer but quickly transition to green once the frost risk subsides, potentially offering a valuable tool for identifying frost-resistant cultivars. The discriminatory power of this selection index is illustrated in Figures 9 through 11, which capture the developmental contrasts between genotypes representing the upper and lower bounds of the index range.

In particular, the 2018 datasets (Figure III-9 and Figure III-11) highlight the index's discriminatory power, showing pronounced contrast between the higher-scoring genotype in row 2 and the lower-scoring genotype in row 3. This enhanced differentiation likely occurred because the 2018 imaging captured a crucial developmental window when genotypes exhibited maximum segregation in anthocyanin levels. This timing advantage is particularly apparent in the distribution plots of the CNJ02 and GRYG populations from May 7, 2018 (CDD=104). Based on these observations, we recommend that future studies concentrate data collection efforts within the CDD range of 75-150, as this window encompasses rapid developmental changes and reveals clearer segregation patterns. Moreover, with marker data and linkage maps now available for these breeding populations, subsequent research can explore the genetic underpinnings of these developmental patterns, potentially accelerating the selection of frost-resistant cultivars through marker-assisted breeding.

Conclusion

Successful integration of next-generation sequencing and dense marker data have propelled the development and utilization of cranberry genomic resources (Covarrubias-Pazaran et al., 2016, 2018; Diaz-Garcia et al., 2018b, 2019, 2021; Weiss et al., 2020; Kawash et al., 2022; Neyhart et al., 2022; Maule et al., 2024). Programs invest in crop genomic resources to streamline the breeding selection process and enhance their germplasm. The synthesis of dense genomic data and precise phenotypic measurements enables the development of more powerful predictive models, associating genetic markers with essential agronomic traits (Meuwissen et al., 2001; Moose and Mumm, 2008). These enhanced models propel breeding efficiency by informing strategic crosses that amplify selection intensity, refine accuracy, and maximize additive genetic diversity (Meuwissen et al., 2001; Heffner et al., 2010; Desta and Ortiz, 2014; Cobb et al., 2019). Boosts in breeding efficiency are particularly amplified when applied to long generation-time perennial crops, like cranberries (Bernardo, 2008; Covarrubias-Pazaran et al., 2018; Ferrão et al., 2021; Seyum et al., 2022; Adunola et al., 2023).

Accurate marker-trait association modeling depends critically on the availability of precise phenotypic data (Araus and Cairns, 2014; Zhao et al., 2019). However, the systematic and repeated collection of robust, comprehensive, and unbiased trait information presents notable challenges, as it often relies on the capabilities and constraints of human evaluators. These constraints stem from subjectivity, inconsistencies, technical errors, and limited scalability often introduced by human scorers (Araus et al., 2018). Human scorers may make judgments prone to personal biases, which can vary based on experience levels and specific capabilities (e.g., color perception differences or deficiencies) (Andersson and Prager, 2006). Inconsistencies can manifest from the difficulty in maintaining uniform evaluation criteria across different days or seasons, with score drift especially evident in long-term studies. The quality of data collection can be compromised by fatigue-induced mistakes, inattentiveness, or transcription errors (incorrect or incomplete entries) that arise under time constraints. Scalability can be challenging when coordinating and standardizing evaluations across multiple human scorers or when collecting multiple traits simultaneously. These human-centric challenges can introduce inconsistencies, inaccuracies, and bottlenecks into the phenotypic dataset, potentially impacting the reliability of subsequent marker-trait association models (White et al., 2012; Araus and Cairns, 2014; Rebetzke et al., 2019; Zhao et al., 2019). Addressing these limitations often requires a combination of training, standardization protocols (White et al., 2015), and where possible, the integration of automated phenotyping technologies to complement human efforts.

This research demonstrates that carefully selected spectral features from UAS imagery can enable robust, interpretable models for monitoring cranberry development. The SA model's effectiveness validates that parsimonious approaches can achieve the necessary predictive accuracy while maintaining practical utility for field applications. These findings establish a foundation for implementing UAS-based monitoring systems in commercial cranberry production. This study's method can be combined with the many various successful applications of RGB cameras to crop high-throughput phenotyping (HTP) include monitoring of biotic and abiotic stressors (Jin et al., 2017; De Swaef et al., 2021; Feng et al., 2021; Sarkar et al., 2021), measuring height (Watanabe et al., 2017), evaluating biomass (Johansen et al., 2020; Liu et al., 2022), assessing texture (Liu et al., 2022), inferring canopy coverage (Starý et al., 2020), estimating yield (Fernandez-Gallego et al., 2020; Johansen et al., 2020), efficiently managing inputs (Abrougui et al., 2022; Colovic et al., 2024), and obtaining other parameters related to the active photosynthetic canopy and senescence such as green area (GA) and greener green area (GGA) (Casadesús et al., 2007; Sarkar et al., 2021).

The selection index developed here provides a quantitative tool for identifying genotypes with potentially advantageous spring development patterns, particularly those that maintain protective anthocyanin pigmentation during frost-risk periods while rapidly transitioning to photosynthetically active states afterward. By concentrating data collection efforts within the critical CDD range of 75-150, where developmental changes are most rapid and genotype segregation is clearest, future studies can better use the discriminatory power of this index.

Some limitations and opportunities for further research should be acknowledged. First, while the selection index is based on the premise that prolonged leaf reddening is associated with greater frost resistance, direct validation of this relationship using controlled frost exposure experiments would strengthen the index's applicability. Second, the model's reduced accuracy for high A535 values (>2) suggests that extremely red-leaved genotypes may not be reliably differentiated, potentially limiting the index's upper-end resolution. Although this would likely have little impact on performing selections when A535 values are lower, it could limit the range of predictions required — limiting a model's power to differentiate — in mechanistic studies. Third, with marker data and linkage maps now available for these breeding populations, a logical next step would be to explore the genetic basis of these developmental patterns, potentially enabling marker-assisted selection for enhanced frost resistance. Fourth, since the LMI scoring index is calculated relative to population-specific A535 exponential decay curves, it is currently only effective for within-population selection and not useful for comparing across populations. Despite this, new exponential decay curves can always be fit across more expansive genetic scopes, regenerating an index that is comparable across genotypes. Lastly, the scope of RGB imaging for phenotyping is limited by its narrow spectral range and low spectral resolution, which particularly limits its scope as an indicator of biotic and abiotic stressors and deficient nutrient status (Tanaka et al., 2024). Moreover, RGB cameras record their values as digital numbers, which limits their applicability across different types of crops or under different imaging conditions (Tanaka et al., 2024).

Despite these limitations, the methods and insights presented here offer valuable tools for cranberry breeders seeking to develop cultivars with improved frost tolerance and yield stability. The application of RGB cameras and image sensors to reliably and efficiently track leaf pigment changes, particularly anthocyanins, has been demonstrated repeatedly (Vina and Gitelson, 2011; Del Valle et al., 2018; Xie et al., 2018; Li and Huang, 2021). UAS-derived RGB phenotyping and modeling has further established its value to high throughput cranberry phenomics through successful application of deep neural networks to predict cranberry heat damage or assess segregating ripening patterns in breeding populations (Akiva et al., 2022; Johnson et al., 2023). This research extends phenomics applications in cranberry research and breeding by uniquely leveraging a low-cost RGB camera-mounted UAS to identify genotypes with spring temporal development patterns that confer improved frost tolerance.

References

Abdallah, A. Y. (1989). Biochemical and biophysical parameters coinciding with initiation of fruit ripening and with seasonal changes in freezing stress resistance of the cranberry plant. United States -- Wisconsin: The University of Wisconsin - Madison. Available at: https://ezproxy.library.wisc.edu/login?url=https://www.proquest.com/dissertations-theses/biochemical-biophysical-parameters-coinciding/docview/303756517/se-2?accountid=465

Abrougui, K., Boughattas, N., Belhaj, M., Buchaillot, M., Segarra, J., Dorbolo, S., et al. (2022). Assessing Phytosanitary Application Efficiency of a Boom Sprayer Machine Using RGB Sensor in Grassy Fields. *Sustainability* 14, 3666. doi: 10.3390/su14063666

Adunola, P., Ferrão, M. A. G., Ferrão, R. G., Da Fonseca, A. F. A., Volpi, P. S., Comério, M., et al. (2023). Genomic selection for genotype performance and environmental stability in *Coffea canephora*. *G3 Genes Genomes Genet.* 13, jkad062. doi: 10.1093/g3journal/jkad062

Akiva, P., Dana, K., Oudemans, P., and Mars, M. (2020). Finding Berries: Segmentation and Counting of Cranberries using Point Supervision and Shape Priors., in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, (Seattle, WA, USA: IEEE), 219–228. doi: 10.1109/CVPRW50498.2020.00033

Akiva, P., Planche, B., Roy, A., Oudemans, P., and Dana, K. (2022). Vision on the bog: Cranberry crop risk evaluation with deep learning. *Comput. Electron. Agric.* 203, 107444. doi: 10.1016/j.compag.2022.107444

Andersson, S., and Prager, M. (2006). "Quantifying Colors," in *Bird Coloration, Volume 1*, eds. G. E. Hill and K. J. McGraw (Harvard University Press), 41–89. doi: 10.2307/j.ctv22jnscm.5

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Araus, J. L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S., and Cairns, J. E. (2018). Translating High-Throughput Phenotyping into Genetic Gain. *Trends Plant Sci.* 23, 451–466. doi: 10.1016/j.tplants.2018.02.001

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: Learning from the last 20 years. *Crop Sci.* 48, 1649–1664. doi: 10.2135/cropsci2008.03.0131

Berry, J. C., Fahlgren, N., Pokorny, A. A., Bart, R. S., and Veley, K. M. (2018). An automated, high-throughput method for standardizing image color profiles to improve image-based plant phenotyping. *PeerJ* 6, e5727. doi: 10.7717/peerj.5727

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (2017). *Classification And Regression Trees.*, 1st Edn. Routledge. doi: 10.1201/9781315139470

Carter, G. A., and Knapp, A. K. (2001). Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. *Am. J. Bot.* 88, 677–684. doi: 10.2307/2657068

Casadesús, J., Kaya, Y., Bort, J., Nachit, M. M., Araus, J. L., Amor, S., et al. (2007). Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. *Ann. Appl. Biol.* 150, 227–236. doi: 10.1111/j.1744-7348.2007.00116.x

Chandler, F., Wilcox, R., Bain, H., Bergman, H., and Dermen, H. (1947). Cranberry breeding investigation of the US Dept. of Agriculture. *Cranberries* 12, 6–9.

Cobb, J. N., Biswas, P. S., and Platten, J. D. (2019). Back to the future: revisiting MAS as a tool for modern plant breeding. *Theor. Appl. Genet.* 132, 647–667. doi: 10.1007/s00122-018-3266-4

Colovic, M., Stellacci, A. M., Mzid, N., Di Venosa, M., Todorovic, M., Cantore, V., et al. (2024). Comparative Performance of Aerial RGB vs. Ground Hyperspectral Indices for Evaluating Water and Nitrogen Status in Sweet Maize. *Agronomy* 14, 562. doi: 10.3390/agronomy14030562

Covarrubias-Pazaran, G., Diaz-Garcia, L., Schlautman, B., Deutsch, J., Salazar, W., Hernandez-Ochoa, M., 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. doi: 10.1186/s12864-016-2802-3

Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., et al. (2018). Multivariate GBLUP Improves Accuracy of Genomic Selection for Yield and Fruit Weight in Biparental Populations of *Vaccinium macrocarpon* Ait. *Front. Plant Sci.* 9, 1–13. doi: 10.3389/fpls.2018.01310

Davies, K. M., Landi, M., Van Klink, J. W., Schwinn, K. E., Brummell, D. A., Albert, N. W., et al. (2022). Evolution and function of red pigmentation in land plants. *Ann. Bot.* 130, 613–636. doi: 10.1093/aob/mcac109

De Swaef, T., Maes, W. H., Aper, J., Baert, J., Cougnon, M., Reheul, D., et al. (2021). Applying RGB- and Thermal-Based Vegetation Indices from UAVs for High-Throughput Field Phenotyping of Drought Tolerance in Forage Grasses. *Remote Sens.* 13, 147. doi: 10.3390/rs13010147

Del Valle, J. C., Gallardo-López, A., Buide, M. L., Whittall, J. B., and Narbona, E. (2018). Digital photography provides a fast, reliable, and noninvasive method to estimate anthocyanin pigment concentration in reproductive and vegetative plant tissues. *Ecol. Evol.* 8, 3064–3076. doi: 10.1002/ece3.3804

Desta, Z. A., and Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19, 592–601. doi: 10.1016/j.tplants.2014.05.006

Diaz-Garcia, L., Covarrubias-Pazaran, G., Schlautman, B., Grygleski, E., and Zalapa, J. (2018a). Image-based phenotyping for identification of QTL determining fruit shape and size in American cranberry (*Vaccinium macrocarpon* L.). *PeerJ* 6, e5461. doi: 10.7717/peerj.5461

Diaz-Garcia, L., Covarrubias-Pazaran, G., Schlautman, B., and Zalapa, J. (2016). GiNA, an Efficient and High-Throughput Software for Horticultural Phenotyping. *PLOS ONE* 11, e0160439. doi: 10.1371/journal.pone.0160439

Diaz-Garcia, L., Garcia-Ortega, L. F., González-Rodríguez, M., Delaye, L., Iorizzo, M., and Zalapa, J. (2021). Chromosome-Level Genome Assembly of the American Cranberry (*Vaccinium macrocarpon* Ait.) and Its Wild Relative *Vaccinium microcarpum*. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.633310

Diaz-Garcia, L., Rodriguez-Bonilla, L., Smith, T., and Zalapa, J. (2019). Pacbio Sequencing Reveals Identical Organelle Genomes Between American Cranberry (*Vaccinium macrocarpon* Ait.) and a Wild Relative. *Genes* 10, 1–15. doi: 10.3390/genes10040291

Diaz-Garcia, L., Schlautman, B., Covarrubias-Pazaran, G., Maule, A., Johnson-Cicalese, J., Grygleski, E., et al. (2018b). Massive phenotyping of multiple cranberry populations reveals novel QTLs for fruit anthocyanin content and other important chemical traits. *Mol. Genet. Genomics* 293, 1379–1392. doi: 10.1007/s00438-018-1464-z

Federer, W. T., and Raghavarao, D. (1975). On Augmented Designs. *Biometrics* 31, 29–35. doi: 10.2307/2529707

Federer, W. T., Unit, C. U. B., Biometrics, C. U. D. of, and Biology, C. U. D. of B. S. and C. (1956). Augmented (or Hoonuiaku) Designs. Available at: https://hdl.handle.net/1813/32841 (Accessed April 29, 2025).

Feng, Z., Song, L., Duan, J., He, L., Zhang, Y., Wei, Y., et al. (2021). Monitoring Wheat Powdery Mildew Based on Hyperspectral, Thermal Infrared, and RGB Image Data Fusion. *Sensors* 22, 31. doi: 10.3390/s22010031

Fernandez-Gallego, J. A., Kefauver, S. C., Gutierrez, N. A., Nieto-Taladriz, M. T., and Araus, J. L. (2020). Implications of Very Deep Super-Resolution (VDSR) on RGB imagery for grain yield assessment in wheat., in *2020 Virtual Symposium in Plant Omics Sciences (OMICAS)*, (Bogotá, Colombia: IEEE), 1–5. doi: 10.1109/OMICAS52284.2020.9535654

Ferrão, L. F. V., Amadeu, R. R., Benevenuto, J., De Bem Oliveira, I., and Munoz, P. R. (2021). Genomic Selection in an Outcrossing Autotetraploid Fruit Crop: Lessons From Blueberry Breeding. *Front. Plant Sci.* 12, 676326. doi: 10.3389/fpls.2021.676326

Fuleki, T., and Francis, F. J. (1968). Quantitative Methods for Anthocyanins.: 3. Purification of Cranberry Anthocyanins. *J. Food Sci.* 33, 266–274. doi: 10.1111/j.1365-2621.1968.tb01365.x

Gamon, J. A., and Surfus, J. S. (1999). Assessing leaf pigment content and activity with a reflectometer. *New Phytol.* 143, 105–117. doi: 10.1046/j.1469-8137.1999.00424.x

Gehan, M. A., Fahlgren, N., Abbasi, A., Berry, J. C., Callen, S. T., Chavez, L., et al. (2017). PlantCV v2: Image analysis software for high-throughput plant phenotyping. *PeerJ* 5, e4088. doi: 10.7717/peerj.4088

Gitelson, A. A., Chivkunova, O. B., and Merzlyak, M. N. (2009). Nondestructive estimation of anthocyanins and chlorophylls in anthocyanic leaves. *Am. J. Bot.* 96, 1861–1868. doi: 10.3732/ajb.0800395

Gitelson, A. A., Kaufman, Y. J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80, 76–87. doi: 10.1016/S0034-4257(01)00289-9

Guijarro, M., Pajares, G., Riomoros, I., Herrera, P. J., Burgos-Artizzu, X. P., and Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Comput. Electron. Agric.* 75, 75–83. doi: 10.1016/j.compag.2010.09.013

Hague, T., Tillett, N. D., and Wheeler, H. (2006). Automated Crop and Weed Monitoring in Widely Spaced Cereals. *Precis. Agric.* 7, 21–32. doi: 10.1007/s11119-005-6787-1

Hashimoto, H., Uragami, C., and Cogdell, R. J. (2016). "Carotenoids and photosynthesis," in *Carotenoids in nature: Biosynthesis, regulation and function*, ed. C. Stange (Cham: Springer International Publishing), 111–139. doi: 10.1007/978-3-319-39126-7_4

Heffner, E. L., Lorenz, A. J., Jannink, J., and Sorrells, M. E. (2010). Plant Breeding with Genomic Selection: Gain per Unit Time and Cost. *Crop Sci.* 50, 1681–1690. doi: 10.2135/cropsci2009.11.0662

Hlatshwayo, S. T., Mutanga, O., Lottering, R. T., Kiala, Z., and Ismail, R. (2019). Mapping forest aboveground biomass in the reforested Buffelsdraai landfill site using texture combinations computed from SPOT-6 pan-sharpened imagery. *Int. J. Appl. Earth Obs. Geoinformation* 74, 65–77. doi: 10.1016/j.jag.2018.09.005

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Huang, Y., Liu, N., Wagner Hokanson, E., Hansen, N., and Townsend, P. A. (2024). Exploring the potential of multi-source satellite remote sensing in monitoring crop nutrient status: A multi-year case study of cranberries in Wisconsin, USA. *Int. J. Appl. Earth Obs. Geoinformation* 132, 104063. doi: 10.1016/j.jag.2024.104063

Hughes, N. M. (2011). Winter leaf reddening in 'evergreen' species. *New Phytol.* 190, 573–581. doi: 10.1111/j.1469-8137.2011.03662.x

Hughes, N. M., Neufeld, H. S., and Burkey, K. O. (2005). Functional role of anthocyanins in high-light winter leaves of the evergreen herb *Galax urceolata*. *New Phytol.* 168, 575–587. doi: 10.1111/j.1469-8137.2005.01546.x

Jeranyama, P., and Kennedy, C. D. (2021). Advancements in spring frost protection to sustain cranberry production in Massachusetts. *Agron. J.* 113, 4560–4567. doi: 10.1002/agj2.20928

Jin, X., Liu, S., Baret, F., Hemerlé, M., and Comar, A. (2017). Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Remote Sens. Environ.* 198, 105–114. doi: 10.1016/j.rse.2017.06.007

Johansen, K., Morton, M. J. L., Malbeteau, Y., Aragon, B., Al-Mashharawi, S., Ziliani, M. G., et al. (2020). Predicting Biomass and Yield in a Tomato Phenotyping Experiment Using UAV Imagery and Random Forest. *Front. Artif. Intell.* 3, 28. doi: 10.3389/frai.2020.00028

Johnson, F., Lowry, J., Dana, K., and Oudemans, P. (2023). Vision-Based Cranberry Crop Ripening Assessment. doi: 10.48550/arXiv.2309.00028

Kataoka, T., Kaneko, T., Okamoto, H., and Hata, S. (2003). Crop growth estimation system using machine vision., in *Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003)*, (Kobe, Japan: IEEE), b1079–b1083. doi: 10.1109/AIM.2003.1225492

Kawash, J., Colt, K., Hartwick, N. T., Abramson, B. W., Vorsa, N., Polashock, J. J., et al. (2022). Contrasting a reference cranberry genome to a crop wild relative provides insights into adaptation, domestication, and breeding. *PLOS ONE* 17, e0264966. doi: 10.1371/journal.pone.0264966

Kawashima, S. (1998). An Algorithm for Estimating Chlorophyll Content in Leaves Using a Video Camera. *Ann. Bot.* 81, 49–54. doi: 10.1006/anbo.1997.0544

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science* 220, 671–680. doi: 10.1126/science.220.4598.671

Kuhn, Max (2008). Building predictive models in R using the caret package. *J. Stat. Softw.* 28, 1–26. doi: 10.18637/jss.v028.i05

Lee, H.-C., 1950- (2005). *Introduction to color imaging science*. Cambridge ; New York : Cambridge University Press, 2005. Available at: https://search.library.wisc.edu/catalog/9910138871402121

Li, Y., and Huang, J. (2021). Leaf Anthocyanin Content Retrieval with Partial Least Squares and Gaussian Process Regression from Spectral Reflectance Data. *Sensors* 21, 3078. doi: 10.3390/s21093078

Liland, K. H., Mevik, B.-H., and Wehrens, R. (2023). pls: Partial least squares and principal component regression. Available at: https://CRAN.R-project.org/package=pls

Lin, C. S., and Poushinsky, G. (1983). A Modified Augmented Design for an Early Stage of Plant Selection Involving a Large Number of Test Lines without Replication. *Biometrics* 39, 553–561. doi: 10.2307/2531083

Liu, N., Wagner Hokanson, E., Hansen, N., and Townsend, P. A. (2023). Multi-year hyperspectral remote sensing of a comprehensive set of crop foliar nutrients in cranberries. *ISPRS J. Photogramm. Remote Sens.* 205, 135–146. doi: 10.1016/j.isprsjprs.2023.10.003

Liu, Y., Feng, H., Yue, J., Jin, X., Li, Z., and Yang, G. (2022). Estimation of potato above-ground biomass based on unmanned aerial vehicle red-green-blue images with different texture features and crop height. *Front. Plant Sci.* 13, 938216. doi: 10.3389/fpls.2022.938216

Loarca, J., Wiesner-Hanks, T., Lopez-Moreno, H., Maule, A. F., Liou, M., Torres-Meraz, M. A., et al. (2024). BerryPortraits: Phenotyping Of Ripening Traits cranberry (*Vaccinium macrocarpon* Ait.) with YOLOv8. *Plant Methods* 20, 172. doi: 10.1186/s13007-024-01285-1

Louhaichi, M., Borman, M. M., and Johnson, D. E. (2001). Spatially Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat. *Geocarto Int.* 16, 65–70. doi: 10.1080/10106040108542184

Maule, A. F., Loarca, J., Diaz-Garcia, L., Lopez-Moreno, H., Johnson-Cicalese, J., Vorsa, N., et al. (2024). Of buds and bits: a meta-QTL study identifies stable QTL for berry quality and yield traits in cranberry mapping populations (*Vaccinium macrocarpon* Ait.). *Front. Plant Sci.* 15, 1294570. doi: 10.3389/fpls.2024.1294570

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157, 1819–1829. doi: 10.1093/genetics/157.4.1819

Meyer, G. E., Hindman, T. W., and Laksmi, K. (1999). Machine vision detection parameters for plant species identification., eds. G. E. Meyer and J. A. DeShazer (Boston, MA), 327–335. doi: 10.1117/12.336896

Meyer, G. E., and Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* 63, 282–293. doi: 10.1016/j.compag.2008.03.009

Moose, S. P., and Mumm, R. H. (2008). Molecular Plant Breeding as the Foundation for 21st Century Crop Improvement. *Plant Physiol.* 147, 969–977. doi: 10.1104/pp.108.118232

National Centers for Environmental Information (2025). Frost Dates: First and last frost dates for Mather 3 Nw, WI. *Natl. Gard. Assoc.* Available at: https://garden.org/apps/frost-dates/Mather%203%20Nw,%20WI/ (Accessed June 9, 2025).

Neyhart, J. L., Kantar, M. B., Zalapa, J., and Vorsa, N. (2022). Genomic-environmental associations in wild cranberry (*Vaccinium macrocarpon* Ait.). *G3 GenesGenomesGenetics* 12, jkac203. doi: 10.1093/g3journal/jkac203

Osorio, F., and Ogueda, A. (2024). Fast computation of some matrices useful in statistics. Available at: https://faosorios.github.io/fastmatrix/

Patil, I. (2021). Visualizations with statistical details: The "ggstatsplot" approach. *J. Open Source Softw.* 6, 3167. doi: 10.21105/joss.03167

Pérez, A. J., López, F., Benlloch, J. V., and Christensen, S. (2000). Colour and shape analysis techniques for weed detection in cereal fields. *Comput. Electron. Agric.* 25, 197–212. doi: 10.1016/S0168-1699(99)00068-X

Rebetzke, G. J., Jimenez-Berni, J., Fischer, R. A., Deery, D. M., and Smith, D. J. (2019). Review: High-throughput phenotyping to enhance the use of crop genetic resources. *Plant Sci.* 282, 40–48. doi: 10.1016/j.plantsci.2018.06.017

Saberioon, M. M., Amin, M. S. M., Anuar, A. R., Gholizadeh, A., Wayayok, A., and Khairunniza-Bejo, S. (2014). Assessment of rice leaf chlorophyll content using visible bands at different growth stages at both the leaf and canopy scale. *Int. J. Appl. Earth Obs. Geoinformation* 32, 35–45. doi: 10.1016/j.jag.2014.03.018

Sarkar, S., Ramsey, A. F., Cazenave, A.-B., and Balota, M. (2021). Peanut Leaf Wilting Estimation From RGB Color Indices and Logistic Models. *Front. Plant Sci.* 12, 658621. doi: 10.3389/fpls.2021.658621

Seyum, E. G., Bille, N. H., Abtew, W. G., Munyengwa, N., Bell, J. M., and Cros, D. (2022). Genomic selection in tropical perennial crops and plantation trees: a review. *Mol. Breed.* 42, 58. doi: 10.1007/s11032-022-01326-4

Sims, D. A., and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81, 337–354. doi: 10.1016/S0034-4257(02)00010-X

Song, P., Wang, J., Guo, X., Yang, W., and Zhao, C. (2021). High-throughput phenotyping: Breaking through the bottleneck in future crop breeding. *Crop J.* 9, 633–645. doi: 10.1016/j.cj.2021.03.015

Starý, K., Jelínek, Z., Kumhálová, J., Chyba, J., and Balážová, K. (2020). Comparing RGB - based vegetation indices from UAV imageries to estimate hops canopy area. 830.9Kb. doi: 10.15159/AR.20.169

Tucker, C. J. (1979). Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens.* 8, 127–150. doi: 10.1016/0034-4257(79)90013-0

Vander Kloet, S. P. (1988). *The genus Vaccinium in North America*. [Ottawa] : Research Branch, Agriculture Canada, 1988. Available at: https://search.library.wisc.edu/catalog/999600224502121

Vander Kloet, S. P., and Avery, T. S. (2010). *Vaccinium* on the Edge. *Edinb. J. Bot.* 67, 7. doi: 10.1017/S0960428609990199

Vina, A., and Gitelson, A. A. (2011). Sensitivity to Foliar Anthocyanin Content of Vegetation Indices Using Green Reflectance. *IEEE Geosci. Remote Sens. Lett.* 8, 464–468. doi: 10.1109/LGRS.2010.2086430

Vorsa, N., and Zalapa, J. (2019). "Domestication, Genetics, and Genomics of the American Cranberry," in *Plant Breeding Reviews*, (John Wiley & Sons, Ltd), 279–315. doi: https://doi.org/10.1002/9781119616801.ch8

Watanabe, K., Guo, W., Arai, K., Takanashi, H., Kajiya-Kanegae, H., Kobayashi, M., et al. (2017). High-Throughput Phenotyping of Sorghum Plant Height Using an Unmanned Aerial Vehicle and Its Application to Genomic Prediction Modeling. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00421

Wei, T., and Simko, V. (2021). R package "corrplot": Visualization of a correlation matrix. Available at: https://github.com/taiyun/corrplot

Weiss, M., Jacob, F., and Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* 236, 111402. doi: 10.1016/j.rse.2019.111402

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crops Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003

White, T. E., Dalrymple, R. L., Noble, D. W. A., O'Hanlon, J. C., Zurek, D. B., and Umbers, K. D. L. (2015). Reproducible research in the study of biological coloration. *Anim. Behav.* 106, 51–57. doi: 10.1016/j.anbehav.2015.05.007

Wintermans, J. F. G. M., and De Mots, A. (1965). Spectrophotometric characteristics of chlorophylls a and b and their phenophytins in ethanol. *Biochim. Biophys. Acta BBA - Biophys. Photosynth.* 109, 448–453. doi: 10.1016/0926-6585(65)90170-6

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., and Mortensen, D. A. (1993). Plant species identification, size, and enumeration using machine vision techniques on near-binary images., eds. J. A. DeShazer and G. E. Meyer (Boston, MA), 208–219. doi: 10.1117/12.144030

Woebbecke, D., Meyer, G., Von Bargen, K., and Mortensen, D. (1995). Color Indices for Weed Identification Under Various Soil, Residue, and Lighting Conditions. *Trans. ASAE* 38, 259–269. doi: 10.13031/2013.27838

Workmaster, B. A. A. (2001). Cold hardiness, ice nucleation, and growth modeling in the cranberry plant. 2001. Available at: https://search.library.wisc.edu/catalog/999926188702121

Workmaster, B. A. A., and Palta, J. P. (2006). Shifts in Bud and Leaf Hardiness during Spring Growth and Development of the Cranberry Upright: Regrowth Potential as an Indicator of Hardiness. *J. Am. Soc. Hortic. Sci.* 131, 327–337. doi: 10.21273/JASHS.131.3.327

Workmaster, B. A. A., and Palta, J. P. (2009). Frost Hardiness of Cranberry Plant: A guide to manage the crop during critical periods in spring and fall. University of Wisconsin-Madison. Available at: https://fruit.webhosting.cals.wisc.edu/wp-content/uploads/sites/36/2016/02/Workmaster-and-Palta-2009-Frost-Hardiness-of-Cranberry-Plant-UW-Madison.pdf

Xie, Y., Civco, D. L., and Silander, J. A. (2018). Species-specific spring and autumn leaf phenology captured by time-lapse digital cameras. *Ecosphere* 9, e02089. doi: 10.1002/ecs2.2089

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., et al. (2017). Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Front. PLANT Sci.* 8. doi: 10.3389/fpls.2017.01111

Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., et al. (2019). Crop Phenomics: Current Status and Perspectives. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00714

Zheng, H., Cheng, T., Zhou, M., Li, D., Yao, X., Tian, Y., et al. (2019). Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. *Precis. Agric.* 20, 611–629. doi: 10.1007/s11119-018-9600-7

Chapter IV Mapping the Genetic Basis of Temporal Segregation of Spring Greening in Cranberry

Abstract

Cranberries (*Vaccinium macrocarpon* Ait.) are vulnerable to late spring frosts during dormancy emergence when cold hardiness rapidly declines. This study presents the first genetic mapping of spring leaf color development in cranberry using the late maturity index (LMI), a trait quantifying spring leaf color changes associated with dormancy exit timing. Quantitative trait loci (QTL) mapping was performed on three populations using a high-density linkage map containing 6,073 single nucleotide polymorphism (SNP) and simple sequence repeat (SSR) markers. Mixed models incorporating spatial effects and genomic relationships estimated breeding values and heritability. QTL mapping in population CNJ02, where LMI exhibited high narrow-sense genomic heritability ($h^2 = 0.74$), yielded eleven significant QTL were identified across linkage groups 6-12, explaining 1.8-9.1% of genetic variance each. The quantitative architecture suggests LMI is controlled by multiple genes of modest effect, consistent with complex dormancy traits. Candidate gene analysis revealed homologs of key dormancy regulators, including circadian clock components (LHY, PRR95), photoperiod sensors (CRY1, COL12), vernalization genes (SVP, AGL24), and flowering integrators (SPL15, FTIP1, FTIP3). These findings provide the first genetic framework for cranberry spring developmental timing and establish molecular targets for breeding frost-resilient cultivars without compromising harvest timing.

Introduction

Cranberries (*Vaccinium macrocarpon* Ait.) are native to North America where wild stands grow in low elevation niches, such as acidic peat bogs and marshes, where dense cold air can settle and make the plant and fruit susceptible to frost pockets (Dana, 1990). However, cranberries exhibit physiological adaptations such as the supercooling and dehydration of reproductive buds that allow fully dormant buds to survive temperatures lower than –22 °C (Abdallah and Palta, 1989; Workmaster and Palta, 2006; Villouta et al., 2020).

The time from mid-April to late May in Wisconsin is a particularly vulnerable time, as late frost events can damage sensitive cranberry tissues, particularly reproductive buds. By mid-April, cranberries in Wisconsin have acquired the necessary chilling time to break dormancy, where chilling hours are the integrated hours of temperatures between 0°C and 10°C (Eck, 1990). As such, cranberry's leaves and terminal buds will readily respond to favorable conditions such as longer days and warmer temperatures, beginning the physiological and anatomical changes necessary for leaves and buds to exit quiescence and resume perennial growth. This emergence from dormancy is linked with a rapid drop in freezing stress resistance in plant tissues, with leaves around 1-2 °C hardier than buds (Workmaster, 2001; Workmaster and Palta, 2006; Villouta et al., 2020).

The ability of cranberries to emerge from dormancy early in May represents an evolutionary adaptation to resource-limited environments, where wild stands must balance growth and reproduction within a compressed growing season. This timing strategy allows older leaves to provide newly manufactured sugars to emerging buds, enabling photoassimilate accumulation necessary for fruit and seed production, while positioning berry ripening to coincide with fall floods that disperse their buoyant fruits (Vander Kloet, 1988; Hagidimitriou and Roper, 1995; Vorsa and Zalapa, 2019). In favorable years, this enables substantial fruit development, seed maturation in the wild and successful genetic dispersal, while in unfavorable years, frost events severely reduce reproductive output.

However, this natural phenological response can make cranberry, as a crop, vulnerable to late spring frosts, risking harvest losses to an industry that requires consistently high year-to-year yields. The late spring frost period represents a critical tipping point in crop productivity, where growers must diligently monitor weather conditions and implement protective management measures such as reflooding or sprinkler irrigation. These interventions carry significant tradeoffs, including equipment costs, operational demands, and new risks such as flooding-induced anoxic damage to actively growing crops (Eck, 1990).

While scientific research insights on cranberry dormancy and hardiness have informed industry to more strategically manage crop risk, studies in the molecular and genetic mechanisms underlying cranberry exit-from-dormancy are limited or non-existent. Additionally, no studies have been performed to map the genetics of spring developmental patterns associated with dormancy exit in a cranberry population. This research aims to performing QTL mapping of late maturity index (*LMI*), a trait developed using UAV images to track spring leaf color changes linked with dormancy exit. The primary goal of this mapping is to guide breeding decisions that will add genetics to the toolkit for mitigating spring freezing damage. A downstream goal is to perform candidate gene analysis on meaningful QTL, informed by molecular and genetic studies of dormancy in related perennial crops. This work will benefit both growers and breeders seeking to mitigate yield losses to late spring frosts by providing a method to phenotype and model spring greening patterns and select cultivars that stay dormant longer and avoid frost damage.

Materials and Methods

Trait

To develop a phenotype scoring index, predicted A535 values were plotted against cumulative growth degree days (CDD) for all cranberry plot images. Exponential decay functions were subsequently fit on a per population and a per genotype level, combining predicted values from 2018 and 2019. Scores were calculated as the conditional integral difference between the genotype-fit and population-fit curves, as defined in Equation IV-1. The last week of May in central Wisconsin, where most cranberries are grown, is a time when the probability of a frost event drops below 30% (National Centers for Environmental Information, 2025). For this study, the last week of May corresponds to around 200 accumulated heat units, denoted as x_c in Equation IV-1. The accumulated integral difference before the x_c threshold was combined with the inverted accumulated integral difference after x_c. This selection index, subsequently referred to as *leaf maturity index* (LMI), scores genotypes with delayed spring development higher: greater red pigmentation before the critical period threshold and less pigmentation after, favoring rapid developmental post-frost.

Equation IV-1. Leaf maturity index (LMI) calculation, where $\mathrm{s_g(x)}$ is the score at CDD x, $\widehat{\mathrm{r_g}}(\mathrm{x})$ is the genotype-fit curve A535 value at CDD x, $\widehat{\mathrm{r_p}}(\mathrm{x})$ is the population-fit curve A535 value at CDD x, x_c is the critical frost period threshold CDD, $\mathrm{S_g}$ is the selection index of genotype g, x_s and x_f are the starting and ending CDD values that bound integration of $\mathrm{s_g(x)}$.

$$s_g(x) = \begin{cases} \widehat{r_g}(x) - \widehat{r_p}(x), x \leq x_c \\ \widehat{r_p}(x) - \widehat{r_g}(x), x > x_c \end{cases}$$

$$S_g = \int_{x_s}^{x_f} s_g(x) d_x$$

Populations

An investigation into the genetic basis of LMI was performed on three cranberry breeding populations planted across two locations. Two half-sib populations, CNJ02 and CNJ04 (n = 69), were planted in an augmented design with four parental checks at the marsh of Cranberry Creek Cranberries, Inc., in Necedah, Wisconsin, USA (Federer et al., 1956; Federer and Raghavarao, 1975; Lin and Poushinsky, 1983). A third population, GRYG (n = 494), was planted at Valley Corporation in Tomah, WI, USA. GRYG was assessed in the year 2018, while CNJ02 and CN0J4 (n = 69) were assessed in the years 2018 and 2019. CNJ02 was derived from a cross between seed-bearing parent, CNJ97_105_3 (Mullica Queen®) and pollen-donating parent, NJS98_23 (Crimson Queen®); CNJ04 was derived from reciprocal crosses between CNJ97_105_3 (Mullica Queen®) and Stevens. GRYG (n = 494) was derived from a cross between seed-bearing parent, [BGx(BLxNL)]95, and the pollen-donating parent, GH1x35.

Linkage Maps

QTL were mapped onto a high-density cranberry composite linkage map containing 6073 SNP and SSR markers corresponding to 1560 unique positions, or bins (Schlautman et al., 2017). As the composite map derived from the three populations in this study, its use in QTL mapping facilitates direct cross-population QTL comparison. Further composite linkage map details are described in Schlautman et al. (2017), while the application and motivation of its use can be found in Maule et al. (2024).

The biparental consensus linkage maps for each population were each respectively used to statistically reinforce the fit of the mixed models with an additive genomic covariance relationship matrix (Endelman, 2011), as described below in *Breeding Values and Heritability*.

Breeding Values and Heritability

To infer breeding values and estimate heritability, mixed models were constructed to regress the phenotype with random effects for genotype, row, column, and 2D-spline spatial patterns. Equation IV-2 shows the mixed model used to estimate breeding values, or best linear unbiased predictors (BLUPs) (Henderson, 1975). The equation variables are defined as follows: y = phenotype value, μ = intercept (global mean of trait), Z_g = genotype random-effect incidence matrix, g = genotypic effects (BLUPs), Z_r = row random effect incidence matrix, r = row effect, Z_c = column random effect incidence matrix, c = column effect, Z_s = 2D-spline random effect incidence matrix, s = spline effect, ε = residuals, G = genotype variance-covariance matrix (de los Campos et al., 2015), A = additive genomic relationship matrix (Endelman, 2011), σ_a^2 = additive genomic variance, I = identity matrix, σ_r^2 = row variance, σ_c^2 = column variance, σ_s^2 = 2D-spline variance, σ_ε^2 = residual error variance.

Equation IV-2. Mixed model for estimating BLUPs within years.

$$y = \mu + Z_g g + Z_r r + Z_c c + Z_s s +,$$

$$where\, g \sim \mathcal{N}(0, G), G = A\sigma_a^2, r \sim \mathcal{N}(0, I\sigma_r^2), c \sim \mathcal{N}(0, I\sigma_c^2), s \sim \mathcal{N}(0, I\sigma_s^2), and\, \varepsilon \sim \mathcal{N}(0, I\sigma_\varepsilon^2).$$

As the phenotype scores were calculated from the combined year fits for CNJ02 and CNJ04, and GRYG had only one year of data, year effects and genotype-by-year effects were not modeled. The availability of linkage map markers for all genotypes enabled robust estimation of additive genetic variance and narrow sense heritability. Narrow-sense genomic heritability calculated from Equation IV-3.

Equation IV-3. Narrow sense heritability calculation for within-year mixed model.

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\varepsilon^2}$$

Model selection was performed from a full model search on the random terms of the model defined in Equation IV-2 using the *Akaike Information Criteria* (AIC) (Akaike, 1974). The model with the lowest AIC was subsequently used to estimate BLUPs, balancing model complexity with parsimony.

QTL Mapping

QTL mapping was performed with the CRAN package r/QTL, a software toolkit for mapping experimental crosses (Broman et al., 2003). BLUPs were used in lieu of phenotypes to map QTL where mixed models successfully converged, or on the unadjusted phenotype values when the mixed models failed to fit. QTL were detected using a model selection approach with Haley-Knott regression to model inter-marker QTL (Haley and Knott, 1992). The model selection approach used the `stepwiseqtl()` function, which runs a forward/backward model search algorithm by which additive and interacting terms are successively added to the model, followed by "backward" pruning of other model terms that optimizes a penalized LOD score. The penalized LOD score used 95th percentile thresholds derived from running `scantwo()` against 10,000 permutations of the phenotypes to limit the false positive rate percent to less than 5 percent (Broman and Sen, 2009).

As the `stepwiseqtl()` function performs a forward-backward model search to find the principal trait QTL, it is limited to finding locally optimal solutions. Consequently, consistent QTL derived from 500 repeated calls to `stepwiseqtl()` provided robust consensus QTL. To find consensus QTL, inter-QTL distances (centiMorgans, or cM) were calculated, and hierarchical agglomerative clustering was applied with a cut-tree length of 5cM. All QTL within a given cluster were condensed into a single consensus QTL, where the position was adjusted based on the weighted mean position of all QTLs within a cluster. Weights were derived from the normalized percent variance explained (PVE) of the QTL. The final QTL model included only those loci detected across the majority (>50%) of the `stepwiseqtl()` repetitions. QTL locations and model statistics were fine-tuned using the `fitqtl()` function parameterized with consensus QTL positions.

Results

Quantile-quantile (QQ) and distributions plots of *LMI* are shown for all three populations in Supplementary Figure B-1 panels A and B. All three *LMI* distributions approximated a normal distribution per the QQ plot fits in Supplementary Figure B-1A, fulfilling the normality assumption of the linear model define in Equation IV-2. Of the three populations, only CNJ02 exhibited convergence while fitting the mixed model defined in Equation IV-2. Consequently, QTL mapping was conducted using unadjusted phenotypes for CNJ04 and GRYG and model-adjusted breeding values for CNJ02 mapping. QTL mapping for *LMI* failed to find any significant loci in the CNJ04 and GRYG populations. Many significant QTL were found by mapping onto the *LMI* BLUPs in population CNJ02.

CNJ02 LMI distributions and summary statistics of the original phenotypes are shown vis-à-vis their model-adjusted BLUPs in Figure IV-1. Boxplot distributions and summary statistics of Crimson Queen and Mullica Queen phenotypes are also shown due to repeated plantings of parental checks and their recurrent measures. Estimated *LMI* genomic heritability was $h^2 = 0.74$. Crimson Queen's observed *LMI* scores range from -5.5 to 24.9, with a median of 15.8 and a mean of 18.7. Mullica Queen's observed *LMI* scores range from -5.1 to 30.8, with a median of 13.2 and a mean of 13. Crimson Queen and Mullica Queen's estimated BLUPs are -1.3 and -1, respectively. CNJ02's observed *LMI* phenotypes range from -60.5 (genotype CNJ02_1_51) to 66.8 (genotype CNJ01_1_70), with a median of -1.4 and a mean of -3.5. CNJ02's BLUP *LMI* estimates range from -40.5 (genotype CNJ02_1_226) to 29.5 (genotype CNJ02_1_226), with a median of -0.5 and a mean of -2.2.

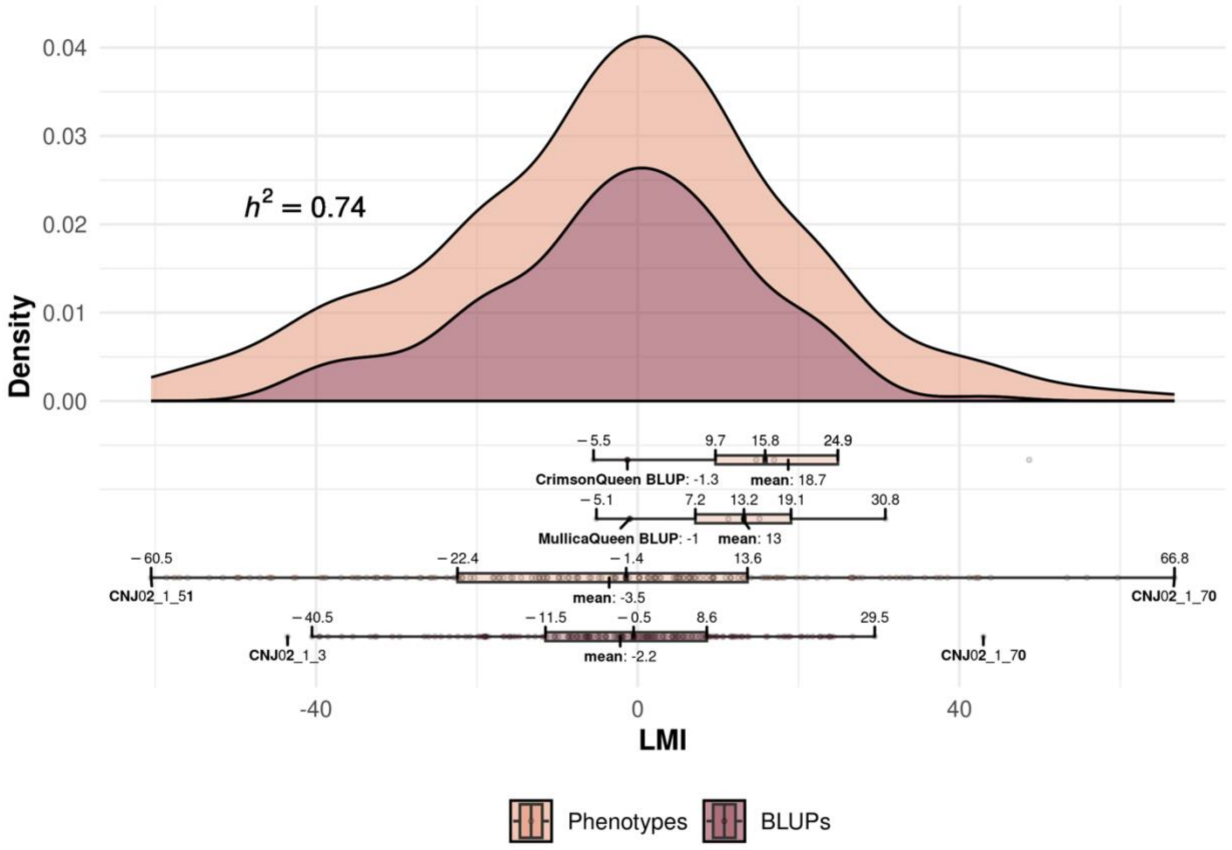


Figure IV-1. Distributions and summary statistics of the original *LMI* trait values and their fitted BLUPs for cranberry population CNJ02.

CNJ02's Consensus QTL derived from 500 stepwiseqtl() iterations are displayed in Table IV-1. In total, 11 significant QTL were found on linkage groups (LG) 6 through 12. The maximum absolute maternal-paternal effect size is an *LMI* effect size of 26.9, and the maximum mean genotype effect size is 11.4. Percent variance explained of QTL (R_q^2) range from 1.8% to 9.1%. Four QTL have R_q^2 above 5%.

Table IV-1. Cranberry population CNJ02 consensus QTL for *LMI*, sorted from highest to lowest R_q^2.

LG	**1.5-LOD min**[a]	**Position**[b]	**1.5-LOD max**[c]	**R_q^2**[d]	**A.-B.**[e]	**.C-.D**[f]	**Int**[g]	**eff_{AC}**[h]	**eff_{AD}**[i]	**eff_{BC}**[j]	**eff_{BD}**[k]
11	43.5	47.6	54.8	9.1%	-6.7	26.9	4.0	6.0±2.4	-9.5±2.5	7.3±2.3	-4.1±2.1
10	43.2	51.8	52.8	7.5%	-20.3	2.8	5.1	-3.5±2.7	-7.4±2.4	4.1±2.2	5.2±2.4
12	33.4	42.9	46.3	7.2%	14.1	-24.7	-2.8	-2.4±2.3	11.4±2.5	-8.0±2.0	3.0±2.4
7	12.3	16.8	19.2	6.1%	17.4	18.2	-4.1	7.4±2.4	0.4±2.3	0.8±2.2	-10.4±2.6
6	56.4	61.0	79.2	4.5%	7.6	18.4	1.1	6.7±2.4	-3.1±2.5	2.3±2.7	-6.3±2.8
8	47.7	50.0	51.7	3.4%	-10.5	17.6	1.1	1.8±2.3	-7.5±2.3	6.6±2.3	-1.7±2.7
11	4.6	11.1	37.2	3.2%	4.2	21.0	2.8	7.0±2.7	-4.9±2.2	3.5±2.2	-5.6±2.7
8	5.3	25.2	41.1	3.0%	-17.6	1.5	-0.6	-4.4±2.4	-4.8±2.5	4.8±2.2	3.7±2.8
12	4.8	7.9	14.6	2.5%	6.2	-13.9	7.2	0.1±2.4	3.5±2.4	-6.6±2.4	4.0±2.7
7	42.6	51.1	57.3	2.0%	12.3	6.6	-12.8	1.2±2.5	4.3±2.4	1.5±2.5	-8.2±2.6
9	4.8	52.5	75.2	1.8%	-0.7	-15.3	-3.4	-4.8±2.3	4.5±2.7	-2.7±2.9	3.2±2.2

[a] 1.5 below peak LOD left position (cM).

[b] QTL position of peak LOD (cM).

[c] 1.5 below peak LOD right position (cM).

[d] Percent of additive genetic variance explained by QTL.

[e] Maternal effect size: (AC+AD)-(BC+BD)

[f] Paternal effect size: (AC+BC)–(AD+BD)

[g] Interaction effect size: (AC+BD)-(AD+BC)

[h] Effect size of AC genotype.

[i] Effect size of AD genotype.

[j] Effect size of BC genotype.

[k] Effect size of BD genotype.

A circos plot of the significant QTL found in Table IV-1 is displayed in Figure IV-2, along with effect plots of the top four QTL by R_q^2. Effect plots are titled with their respective QTL, with genotype effects in the top and bottom and respective maternal-paternal-interaction effects on the left and right. QTL are represented as solid circles in the outer ring, sized relative to their R_q^2. Just inside the QTL are their associated LOD-1.5 intervals. Continuing inward towards the center are the linkage map axes, linkage map markers, and a ring segment representing the proportion of genetic variation explained by the QTL model. At the center of the graph is a pedigree of the CNJ02 population, annotated with parental abbreviations represented as nodes with relevant edges annotated with parental haplotypes.

For QTL 11@47.6 cM, the Stevens haplotype (A) contributed a small negative effect (-6.7) relative to Ben Lear's haplotype (B), while LeMunyon's haplotype (C) produced a substantial positive effect (26.9) compared to haplotype D (#35). The interaction effect was 4.0. Maximum and minimum effects were 7.3 (genotype BC) and -9.5 (genotype AD).

QTL 10@51.8 cM also displayed contrasting patterns of the maternal and paternal effects relative to the corresponding effects on QTL 11@47.6 cM. Here, Stevens’ haplotype conferred a large negative effect (-20.3) versus the Ben Lear haplotype, whereas LeMunyon's haplotype produced only a miniscule increase (2.8) relative to the #35 haplotype (D). The interaction effect was 5.0. Genotype BD presented the maximum effect (5.2), while genotype AD showed the minimum (-7.4).

At QTL 12@42.9 cM, the Stevens haplotype provided a meaningful positive contribution (14.1) compared to Ben Lear's haplotype, but LeMunyon's haplotype caused a very large decrease (-24.7) versus the #35 haplotype. A small interaction effect of 2.8 was found. The maximum effect occurred in genotype AD (11.4) and minimum in genotype BC (-8.0).

QTL 7@16.8 cM showed similar maternal and paternal effect magnitudes (17.4 and 18.2) but with a negative interaction effect (-4.1). That is, the Stevens haplotype and the LeMunyon haplotype contributed large positive effects relative to the Ben Lear and #35 haplotypes, respectively. Genotype AC produced the maximum effect (7.4) while genotype BD yielded the minimum (-10.4).

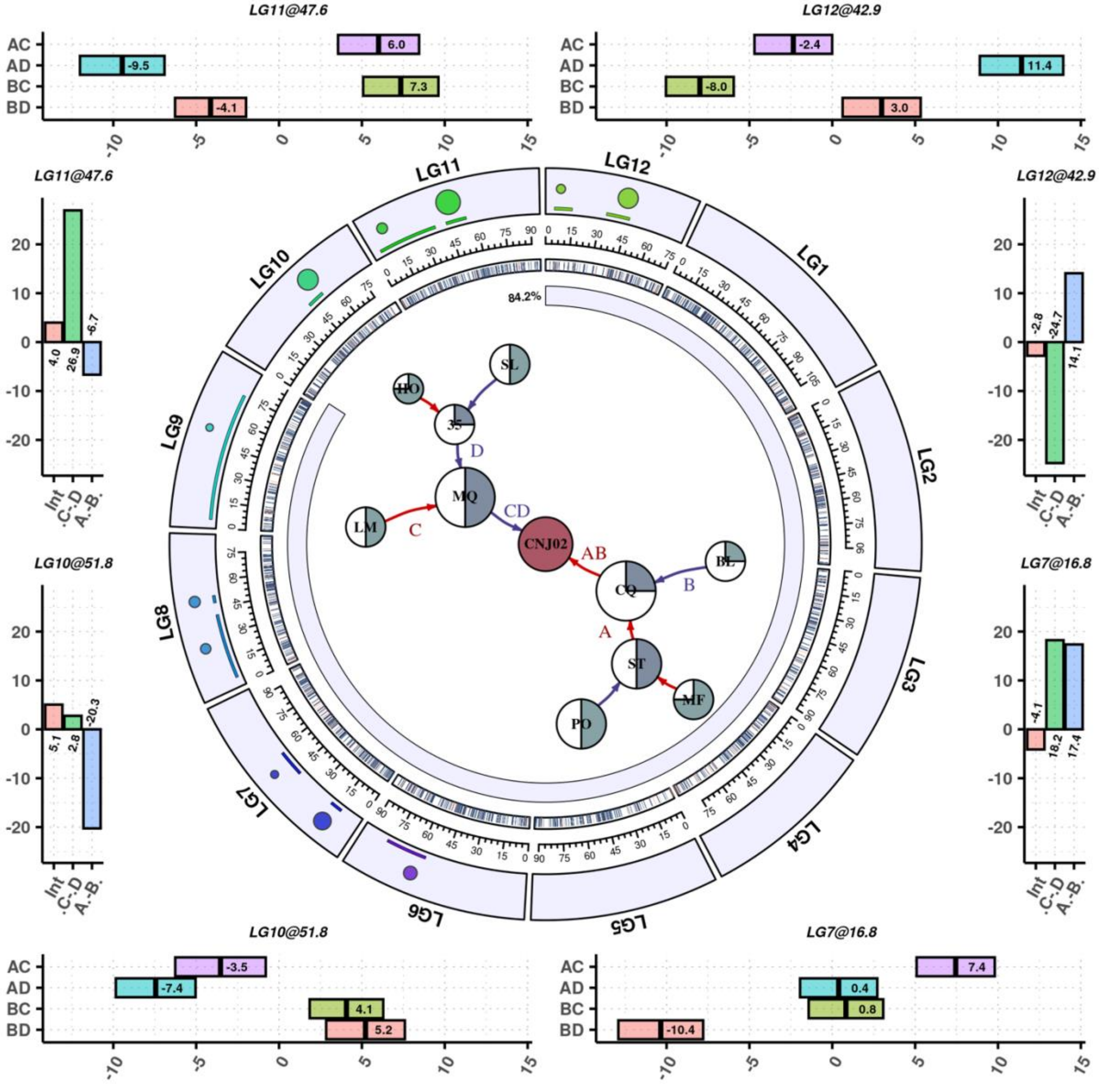


Figure IV-2. Significant QTL found for trait *LMI* in cranberry population CNJ02. Effect size plots for the four highest-ranking QTL by R^2 are displayed along the figure margins. Crossbar plots of mean effect sizes across four possible QTL genotypes are shown at the top and bottom of the figure, with their associated standard error. Bar plots of the average maternal, paternal, and interaction effect sizes are shown at the left and right margins. The outer ring of the circos plot shows the QTL (circles) and their associated LOD-1.5 interval, with QTL sizes proportional to their respective R^2. The center ring displays the linkage map marker locations, and the innermost circos sector shows the percent variance explained by the QTL model. The center plot demonstrates the CNJ02 population's pedigree, where each node is labeled with its cultivar abbreviation, with edge arrows indicating the crossing direction. Cultivar abbreviations are as follows: HO = Howes, SL = Searles, BL = Ben Lear, MF = McFarlin, PO = Potters Favorite, LM = LeMunyon, 35 = #35, CQ = Crimson Queen, ST = Stevens, MQ = Mullica Queen. Red edges represent the ovule parent, dark blue the pollen parent. Edge letters, where specified, are the QTL haplotypes associated with the respective grandparent.

Discussion

This research represents the first exploration of the genetics underlying spring leaf color development in cranberries, advancing our understanding of dormancy emergence to inform breeding decisions to increase and protect yield for cranberry growers. Building upon the development of a leaf maturity index (LMI)—a trait that distinguishes phenotypes based on favorable spring developmental patterns—this work identifies genetic markers that can serve as selection tools or targets for further investigation of the genetic components associated with spring leaf color shifts and dormancy exit patterns. The primary goal of this mapping is to guide breeding decisions that will add genetics to the toolkit for mitigating spring freezing damage. A downstream goal is to perform candidate gene analysis on meaningful QTL, informed by molecular and genetic studies of dormancy in related perennial crops. This work will benefit both growers and breeders seeking to mitigate yield losses to late spring frosts by providing a method to phenotype and model spring greening patterns and select cultivars that stay dormant longer and avoid frost damage.

Model Fit and Heritabilities

The mixed model as defined in Equation IV-2 failed to converge on solutions for populations CNJ04 and GRYG. Consequently, GBLUP predictions, or breeding value inferences, were never found for these two populations. Additionally, model variance components could not be inferred, meaning that genomic heritability estimates could not be calculated. There could be several explanations for failed model fit outside of a lack of underlying genetic variation for the LMI trait. For instance, the GRYG population only had data available for one year of data. As the LMI trait is derived from data fit to three-parameter exponential curves, incomplete data may fail to properly capture and calculate LMI values for genotypes (overfit on one-year), saturating the GBLUP models with error variance such that genetic signals cannot be decomposed from the model noise. As a smaller population, CNJ04 could lack the power to detect genetic signals tied to LMI, made more pronounced by a lack of repeated plantings.

QTL Patterns

CNJ02 exhibited high genomic heritability for LMI ($h^2 \approx 0.74$). Since the genetic variation captures and represents more of the available selectable phenotypic variation for CNJ02, QTL mapping models had the power to detect 11 significant QTL explaining around 84% of the GBLUP variation, spread across linkage groups 6 to 12. None of these QTL were classified as major ($R_q^2 > 0.1$), but four QTL were found to have modest contributions ($R_q^2 > 0.05$). These four QTL exhibit no consistent genotype or haplotype trends in the magnitude and direction of their effects. This lack of a coherent pattern could manifest from the underlying trait complexity -- LMI does not simply capture how slow or fast leaf color changes but also rates of acceleration. Also, although LMI is ultimately tied to a measurement for leaf anthocyanin pigmentation, it could additionally include the complex genetic mechanisms associated with dormancy emergence in perennials.

The pedigree at the center of Figure IV-2 includes each ancestor's harvest window class. Specifically, the four QTL haplotypes A, B, C, and D represent cultivars Stevens, Ben Lear, Le Munyon, and #35, respectively. These four cultivars are roughly classified into the following harvest windows: Stevens (mid-season), Ben Lear (early-season), Le Munyon (mid-season), and #35 (early-season). If harvest window class associated with LMI, then we'd expect consistency of harvest window with QTL effects. No general association appears to exist between harvest window and LMI, which suggests that selecting on LMI will not impact the harvest window in CNJ02. This likely avoids linkage or pleiotropy between the traits, which is ideal because growers favor early-harvest cultivars. Despite a general lack of correspondence between harvest window and LMI, investigating inter-trait patterns in specific LMI QTL is warranted when using marker-assistant selection.

Candidate Genes

Although LMI more directly measures the dynamics of leaf anthocyanin loss in cranberries, the genetics of variation in this pigment's kinetics could be a consequence of upstream differences in molecular pathways governing dormancy emergence. Dormancy in perennials is a dynamic state that requires a complex choreography of anatomical, physiological, and molecular changes needed to harden a plant to survive seasonal bottlenecks like winter. A plant senses and integrates environmental cues like photoperiod and temperature to synchronize dormancy entry and exit to aptly prioritize survival or growth.

Characterization of the molecular pathways governing dormancy in perennial trees has demonstrated a high correspondence with the components that control floral initiation and timing in annuals like Arabidopsis (Zhao et al., 2025). Flowering synchrony in annuals integrates the same environmental cues as dormancy, sensing photoperiod and chilling hours (vernalization) through photoreceptors that interact with circadian rhythm genes or through cold-induced chemokinetic processes that reduce repression of floral induction signal pathways (Wilkie et al., 2008). These molecular processes, through gene duplication events and subsequent specialization, have been retooled to control and calibrate dormancy (André et al., 2022). Additionally, as with many plant developmental processes, phytohormones are important endogenous signals in orchestrating dormancy processes (Eriksson and Moritz, 2002; Olsen, 2010; Cooke et al., 2012).

Motivated by literature characterizing floral development and the multifaceted processes underlying dormancy emergence, a preliminary gene candidate search was performed using data from cranberry genome assemblies. This included searching for dormancy-related, regulatory gene analogs like those reviewed by Zhao et al. (2025), in addition to searching for downstream gene pathways – genes involved in anthocyanin synthesis and degradation; carbohydrate, lipid, and protein allocation and formulation changes; abscisic acid and gibberellin signaling; and callose degradation as it relates to plasmodesmata microstructural shifts (Welling and Palva, 2006; Maule, 2008; Zhang et al., 2011; Ren et al., 2019; Li et al., 2020; Wang et al., 2022; Hu et al., 2024).

A summary of characterized genes within the 1.5 LOD interval is shown in Supplementary Table B-1. These initial candidate genes are outlined in Supplementary Table B-2, ordered by from top to bottom by decreasing QTL R_q^2. The discussion that follows will highlight developmentally critical genes, with particular attention to those mediating circadian rhythm calibration, photoperiod detection, chilling hour-dependent flowering pathways, and dormancy regulation in perennial systems. As important upstream regulatory elements, small genetic variation in these genes are more likely to manifest and affect multiple downstream components of dormancy emergence, including the leaf color changes mapped in this study. Genetic pathways that act downstream of the main regulatory machinery that orchestrates dormancy emergence are included in Supplementary Table B-2, but not mentioned further.

For QTL 11@47.62 ($R_q^2 \approx 9.12\%$), both the Stevens and Ben Lear describe a candidate gene with homology to Agamous-like MADS-box protein mads1, a MADS-domain-containing transcription factor that demonstrates a regulatory role in floral and berry development in *Vitis vinifera* (Boss et al., 2001). LHY is an additional gene (Stevens) found near QTL 11@47.62 that is a core part of the molecular feedback machinery governing circadian rhythm calibrating and photoperiod sensing (Schaffer et al., 1998; Mizoguchi et al., 2002), acting as an important mediator of shoot growth in poplar by acting on FT2 (Ramos-Sánchez et al., 2019). PRR95, a putative two-component plant response regulator, is located very close to the LHY gene (Stevens). Like LHY, PRR95 may be at the basis of the endogenous circadian clock and demonstrates control of photoperiodic flowering response in *Oryza sativa subsp. japonica* (Murakami et al., 2003; MURAKAMI et al., 2005, 2007).

QTL10@51.79 ($R_q^2 \approx 7.5\%$) has three important floral development genes that operate in the photoperiod and vernalization molecular signaling networks. The gene bHLH93 (Ben Lear), sometimes labeled as NFL in *Arabidopsis*, is a transcription factor that effects short-day floral meristem induction by regulating downstream gibberellin pathways (Sharma et al., 2016; Poirier et al., 2018). AGL24 (Ben Lear) is a transcriptional activator known to mediate photoperiod and vernalization signals in determining floral identity in *Arabidopsis* (Yu et al., 2002; Michaels et al., 2003). SVP (Stevens) is a well-documented, MADS-box transcription factor that represses floral transition in *Arabidopsis* meristems (Hartmann et al., 2000; Gregis et al., 2006, 2008), with analogs in several perennial crops showing evidence of its role in repressing bud break (Wu et al., 2017a, 2017b; Singh et al., 2018, 2019). CRY1 (Stevens), or cryptochrome-1, is a member of blue-light photosensors that have been shown to regulate photomorphogenesis in *Arabidopsis* (Yu et al., 2010), while a paralog in poplar has been shown to regulate shoot growth under short day conditions (Wei et al., 2024b).

Although the QTL 6@60.98 explains less than 5% of the model variance, it contains several important regulatory genes that could explain its contribution of phenotypic variation for LMI. For Ben Lear, two FT-interacting protein homologs are annotated – FT-interacting protein 1 (FTIP1) and FT-interacting protein 3 (FTIP3). FTIP1 is involved in the export of FT from the phloem companion cells to the sieve elements through the plasmodesmata in *Arabidopsis*, regulating flowering time under long days (Liu et al., 2012). FTIP3 mediates the proliferation and differentiation of shoot stem cells in the shoot apical meristem (SAM) by controlling intracellular trafficking and subcellular localization of stem cell identity factors (Liu et al., 2018).

Stevens lists two important candidate genes at QTL 6@60.98 – COL12 and SPL15. COL12, or CONSTANS-LIKE 12, is a zinc finger transcription factor that affects photoperiod-based control of flowering by repressing expression of FT in *Arabidopsis* (Griffiths et al., 2003; Ordoñez-Herrera et al., 2018). SPL15, or squamosa promoter-binding-like protein 15, is a plant-specific transcription factor involved in various developmental processes, particularly the transition from vegetative to reproductive growth. It has demonstrated its role in age-based vernalization in the perennial *Arabidopsis alpina* through interaction with miR156, a well-documented microRNA involved in regulatory pathways downstream of photoperiod and vernalization signals (Zhao et al., 2025). SPL16 and SPL23, homologs of SPL15, mediate shoot growth in short days in poplar by direct repression of FT2 (Wei et al., 2024a). Genes that closely interact with *Flowering Locus T* (FT) paralogs are particularly interesting as they have demonstrated fundamental roles in releasing endodormancy (FT1) and roles in promotive growth processes (FT2) (Hsu et al., 2011; André et al., 2022). More promising, an FT-like gene in northern highbush blueberry (*Vaccinium corymbosum*) (*Vc*FT) was shown to be a flowering activator that reverses photoperiodic and chilling requirements in blueberry in overexpressed transgenic line (Song et al., 2013; Song, 2025).

Conclusion and Future Directions

As the first exploration of the genetics underlying spring leaf color development in cranberries, this study lays the foundation for several future studies and practical breeding applications. The highly quantitative nature of LMI in population CNJ02 is consistent with the hypothesis that LMI is a complex trait likely associated with genetic variation in upstream dormancy regulatory processes. A literature review on dormancy and dormancy-adjacent genetic regulation and molecular signaling in perennials, combined with preliminary gene analysis on two cranberry genome assemblies, demonstrate provisional evidence that the principal QTL are near important dormancy regulatory elements.

To strengthen evidence of these QTL's roles in dormancy regulation, future work should look at gene enrichment relative to the genome, incorporating gene ontology, protein families, and important regulatory domains into the analyses. Uncharacterized proteins in the Ben Lear and Stevens genome assemblies could warrant a more thorough homology search to add and amend existing annotations, including important non-gene regulatory elements identified as important to dormancy emergence. Since the genomic range of many of these QTL and their 1.5 LOD intervals can be rather large, narrowing the genomic search window around the predicted QTL position may provide more relevant gene candidates.

To validate and characterize these QTL and their prospective genetic elements, transcriptomic and metabolomic studies could be performed on genotypes stratified to the lowest and highest BLUPs. Performing GWAS mapping on a diversity panel of cranberries could also broaden our understanding of the standing genetic diversity of the LMI trait. Candidate gene analysis on significant GWAS loci would subsequently inform the design of downstream transcriptomic and metabolomic studies.

Although the scope of the current study is limited to the genetic variation endemic in CNJ02, it demonstrates distinct potential in its application to breeding and to understanding the genetic mechanisms underpinning dormancy emergence in cranberries. At a minimum, the QTL defined herein can inform meaningful crosses from the CNJ02 pedigree to develop genotypes with more inherent resilience to late spring frosts, without compromising the timing of ripening of berries in fall. More meaningfully, functional studies guided by QTL, GWAS, or genomic prediction results could serve to enlighten the evolutionary underpinnings of dormancy emergence in cranberries specifically, as well as in economically important *Vaccinium* species such as highbush blueberry (*V. corymbosum*), lowbush blueberry (*V. angustifolium*), rabbiteye blueberry (*V. virgatum*), lingonberry (*V. vitis-ideae*), and bilberry (*V. myrtillus*). Candidate gene analysis lends support or at least understanding to the genetic variation for this trait in CNJ02. Functional characterization would require more sophisticated studies such as studying the transcriptomes/metabolomes of some select genotypes to better understand regulation differences, along with possible cloning and expression altering in candidate genes.

References

Abdallah, A. Y., and Palta, J. P. (1989). Changes in the Freezing Stress Resistance of the Cranberry Leaf, Flower Bud, and Fruit During Growth and Development., in *Acta Horticulturae*, (International Society for Horticultural Science (ISHS), Leuven, Belgium), 273–276. doi: 10.17660/ActaHortic.1989.241.46

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19, 716–723. doi: 10.1109/TAC.1974.1100705

André, D., Marcon, A., Lee, K. C., Goretti, D., Zhang, B., Delhomme, N., et al. (2022). *FLOWERING LOCUS T* paralogs control the annual growth cycle in *Populus* trees. *Curr. Biol.* 32, 2988-2996.e4. doi: 10.1016/j.cub.2022.05.023

Boss, P. K., Vivier, M., Matsumoto, S., Dry, I. B., and Thomas, M. R. (2001). A cDNA from grapevine (Vitis vinifera L.), which shows homology to AGAMOUS and SHATTERPROOF, is not only expressed in flowers but also throughout berry development. *Plant Mol. Biol.* 45, 541–553. doi: 10.1023/A:1010634132156

Broman, K. W., and Sen, S. (2009). *A Guide to QTL Mapping with R/qtl*. New York, NY: Springer New York. doi: 10.1007/978-0-387-92125-9

Broman, K. W., Wu, H., Sen, S., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19, 889–890. doi: 10.1093/bioinformatics/btg112

Cooke, J. E. K., Eriksson, M. E., and Junttila, O. (2012). The dynamic nature of bud dormancy in trees: environmental control and molecular mechanisms. *Plant Cell Environ.* 35, 1707–1728. doi: 10.1111/j.1365-3040.2012.02552.x

Dana, M. N. (1990). "Cranberry management," in *Small fruit crop management*, eds. G. J. Galletta and D. G. Himelrick (Englewood Cliffs, N.J. : Prentice-Hall, [1990] ©1990), 334–362. Available at: https://search.library.wisc.edu/catalog/999610034602121 de los Campos, G., Sorensen, D., and Gianola, D. (2015). Genomic Heritability: What Is It? *PLoS Genet.* 11, 1–21. doi: 10.1371/journal.pgen.1005048

Eck, P. (1990). *The American cranberry*. New Brunswick : Rutgers University Press, [1990] ©1990. Available at: https://search.library.wisc.edu/catalog/999621210702121

Endelman, J. B. (2011). Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024

Eriksson, M., and Moritz, T. (2002). Daylength and spatial expression of a gibberellin 20-oxidase isolated from hybrid aspen (Populus tremula L. × P. tremuloides Michx.). *Planta* 214, 920–930. doi: 10.1007/s00425-001-0703-3

Federer, W. T., and Raghavarao, D. (1975). On Augmented Designs. *Biometrics* 31, 29–35. doi: 10.2307/2529707

Federer, W. T., Unit, C. U. B., Biometrics, C. U. D. of, and Biology, C. U. D. of B. S. and C. (1956). Augmented (or Hoonuiaku) Designs. Available at: https://hdl.handle.net/1813/32841 (Accessed April 29, 2025).

Gregis, V., Sessa, A., Colombo, L., and Kater, M. M. (2006). AGL24, SHORT VEGETATIVE PHASE, and APETALA1 Redundantly Control AGAMOUS during Early Stages of Flower Development in Arabidopsis. *Plant Cell* 18, 1373–1382. doi: 10.1105/tpc.106.041798

Gregis, V., Sessa, A., Colombo, L., and Kater, M. M. (2008). AGAMOUS-LIKE24 and SHORT VEGETATIVE PHASE determine floral meristem identity in Arabidopsis. *Plant J.* 56, 891–902. doi: 10.1111/j.1365-313X.2008.03648.x

Griffiths, S., Dunford, R. P., Coupland, G., and Laurie, D. A. (2003). The Evolution of CONSTANS-Like Gene Families in Barley, Rice, and Arabidopsis. *Plant Physiol.* 131, 1855–1867. doi: 10.1104/pp.102.016188

Hagidimitriou, M., and Roper, T. R. (1995). Seasonal changes in CO2 assimilation of cranberry leaves. *Sci. Hortic.* 64, 283–292. doi: 10.1016/0304-4238(95)00835-7

Haley, C. S., and Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69, 315–324. doi: 10.1038/hdy.1992.131

Hartmann, U., Höhmann, S., Nettesheim, K., Wisman, E., Saedler, H., and Huijser, P. (2000). Molecular cloning of SVP: a negative regulator of the floral transition in Arabidopsis. *Plant J.* 21, 351–360. doi: 10.1046/j.1365-313x.2000.00682.x

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31, 423–47.

Hsu, C.-Y., Adams, J. P., Kim, H., No, K., Ma, C., Strauss, S. H., et al. (2011). FLOWERING LOCUS T duplication coordinates reproductive and vegetative growth in perennial poplar. *Proc. Natl. Acad. Sci.* 108, 10756–10761. doi: 10.1073/pnas.1104713108

Hu, Z., Wu, Z., Zhu, Q., Ma, M., Li, Y., Dai, X., et al. (2024). Multilayer regulatory landscape and new regulators identification for bud dormancy release and bud break in Populus. *Plant Cell Environ.* 47, 3181–3197. doi: 10.1111/pce.14938

Li, G., Lin, R., Egekwu, C., Blakeslee, J., Lin, J., Pettengill, E., et al. (2020). Seasonal nitrogen remobilization and the role of auxin transport in poplar trees. *J. Exp. Bot.* 71, 4512–4530. doi: 10.1093/jxb/eraa130

Lin, C. S., and Poushinsky, G. (1983). A Modified Augmented Design for an Early Stage of Plant Selection Involving a Large Number of Test Lines without Replication. *Biometrics* 39, 553–561. doi: 10.2307/2531083

Liu, L., Li, C., Song, S., Teo, Z. W. N., Shen, L., Wang, Y., et al. (2018). FTIP-Dependent STM Trafficking Regulates Shoot Meristem Development in *Arabidopsis*. *Cell Rep.* 23, 1879–1890. doi: 10.1016/j.celrep.2018.04.033

Liu, L., Liu, C., Hou, X., Xi, W., Shen, L., Tao, Z., et al. (2012). FTIP1 is an essential regulator required for florigen transport. *PLoS Biol.* 10, e1001313. doi: 10.1371/journal.pbio.1001313

Maule, A. F., Loarca, J., Diaz-Garcia, L., Lopez-Moreno, H., Johnson-Cicalese, J., Vorsa, N., et al. (2024). Of buds and bits: a meta-QTL study identifies stable QTL for berry quality and yield traits in cranberry mapping populations (*Vaccinium macrocarpon* Ait.). *Front. Plant Sci.* 15, 1294570. doi: 10.3389/fpls.2024.1294570

Maule, A. J. (2008). Plasmodesmata: structure, function and biogenesis. *Curr. Opin. Plant Biol.* 11, 680–686. doi: 10.1016/j.pbi.2008.08.002

Michaels, S. D., Ditta, G., Gustafson-Brown, C., Pelaz, S., Yanofsky, M., and Amasino, R. M. (2003). AGL24 acts as a promoter of flowering in Arabidopsis and is positively regulated by vernalization. *Plant J.* 33, 867–874. doi: 10.1046/j.1365-313X.2003.01671.x

Mizoguchi, T., Wheatley, K., Hanzawa, Y., Wright, L., Mizoguchi, M., Song, H.-R., et al. (2002). *LHY* and *CCA1* Are Partially Redundant Genes Required to Maintain Circadian Rhythms in *Arabidopsis*. *Dev. Cell* 2, 629–641. doi: 10.1016/S1534-5807(02)00170-3

Murakami, M., Ashikari, M., Miura, K., Yamashino, T., and Mizuno, T. (2003). The Evolutionarily Conserved OsPRR Quintet: Rice Pseudo-Response Regulators Implicated in Circadian Rhythm. *Plant Cell Physiol.* 44, 1229–1236. doi: 10.1093/pcp/pcg135

MURAKAMI, M., MATSUSHIKA, A., ASHIKARI, M., YAMASHINO, T., and MIZUNO, T. (2005). Circadian-Associated Rice Pseudo Response Regulators (OsPRRs): Insight into the Control of Flowering Time. *Biosci. Biotechnol. Biochem.* 69, 410–414. doi: 10.1271/bbb.69.410

MURAKAMI, M., TAGO, Y., YAMASHINO, T., and MIZUNO, T. (2007). Characterization of the Rice Circadian Clock-Associated Pseudo-Response Regulators in Arabidopsis thaliana. *Biosci. Biotechnol. Biochem.* 71, 1107–1110. doi: 10.1271/bbb.70048

National Centers for Environmental Information (2025). Frost Dates: First and last frost dates for Mather 3 Nw, WI. *Natl. Gard. Assoc.* Available at: https://garden.org/apps/frost-dates/Mather%203%20Nw,%20WI/ (Accessed June 9, 2025).

Olsen, J. E. (2010). Light and temperature sensing and signaling in induction of bud dormancy in woody plants. *Plant Mol. Biol.* 73, 37–47. doi: 10.1007/s11103-010-9620-9

Ordoñez-Herrera, N., Trimborn, L., Menje, M., Henschel, M., Robers, L., Kaufholdt, D., et al. (2018). The Transcription Factor COL12 Is a Substrate of the COP1/SPA E3 Ligase and Regulates Flowering Time and Plant Architecture. *Plant Physiol.* 176, 1327–1340. doi: 10.1104/pp.17.01207

Poirier, B. C., Feldman, M. J., and Lange, B. M. (2018). bHLH093/NFL and bHLH061 are required for apical meristem function in Arabidopsis thaliana. *Plant Signal. Behav.* 13, e1486146. doi: 10.1080/15592324.2018.1486146

Ramos-Sánchez, J. M., Triozzi, P. M., Alique, D., Geng, F., Gao, M., Jaeger, K. E., et al. (2019). LHY2 Integrates Night-Length Information to Determine Timing of Poplar Photoperiodic Growth. *Curr. Biol.* 29, 2402-2406.e4. doi: 10.1016/j.cub.2019.06.003

Ren, J., Liu, Z., Chen, W., Xu, H., and Feng, H. (2019). Anthocyanin Degrading and Chlorophyll Accumulation Lead to the Formation of Bicolor Leaf in Ornamental Kale. *Int. J. Mol. Sci.* 20, 603. doi: 10.3390/ijms20030603

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carré, I. A., et al. (1998). The *late elongated hypocotyl* Mutation of *Arabidopsis* Disrupts Circadian Rhythms and the Photoperiodic Control of Flowering. *Cell* 93, 1219–1229. doi: 10.1016/S0092-8674(00)81465-8

Schlautman, B., Covarrubias-Pazaran, G., Diaz-Garcia, L., Iorizzo, M., Polashock, J., Grygleski, E., 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 GenesGenomesGenetics* 7, 1177–1189. doi: 10.1534/g3.116.037556

Sharma, N., Xin, R., Kim, D.-H., Sung, S., Lange, T., and Huq, E. (2016). NO FLOWERING IN SHORT DAY (NFL) is a bHLH transcription factor that promotes flowering specifically under short-day conditions in Arabidopsis. *Development* 143, 682–690. doi: 10.1242/dev.128595

Singh, R. K., Maurya, J. P., Azeez, A., Miskolczi, P., Tylewicz, S., Stojkovič, K., et al. (2018). A genetic network mediating the control of bud break in hybrid aspen. *Nat. Commun.* 9, 4173. doi: 10.1038/s41467-018-06696-y

Singh, R. K., Miskolczi, P., Maurya, J. P., and Bhalerao, R. P. (2019). A Tree Ortholog of *SHORT VEGETATIVE PHASE* Floral Repressor Mediates Photoperiodic Control of Bud Dormancy. *Curr. Biol.* 29, 128-133.e2. doi: 10.1016/j.cub.2018.11.006

Song, G. (2025). Two Lines Enable FasTrack Breeding in Blueberry. *J Amer Soc Hort Sci* 150, 28–33. doi: 10.21273/JASHS05447-24

Song, G., Walworth, A., Zhao, D., Jiang, N., and Hancock, J. F. (2013). The Vaccinium corymbosum FLOWERING LOCUS T-like gene (VcFT): a flowering activator reverses photoperiodic and chilling requirements in blueberry. *Plant Cell Rep.* 32, 1759–1769. doi: 10.1007/s00299-013-1489-z

Vander Kloet, S. P. (1988). *The genus Vaccinium in North America*. [Ottawa] : Research Branch, Agriculture Canada, 1988. Available at: https://search.library.wisc.edu/catalog/999600224502121

Villouta, C., Workmaster, B. A., Bolivar-Medina, J., Sinclair, S., and Atucha, A. (2020). Freezing stress survival mechanisms in Vaccinium macrocarpon Ait. terminal buds. *Tree Physiol.* 40, 841–855. doi: 10.1093/treephys/tpaa028

Vorsa, N., and Zalapa, J. (2019). "Domestication, Genetics, and Genomics of the American Cranberry," in *Plant Breeding Reviews*, (John Wiley & Sons, Ltd), 279–315. doi: https://doi.org/10.1002/9781119616801.ch8

Wang, W., Talide, L., Viljamaa, S., and Niittylä, T. (2022). Aspen growth is not limited by starch reserves. *Curr. Biol.* 32, 3619-3627.e4. doi: 10.1016/j.cub.2022.06.056

Wei, H., Luo, M., Deng, J., Xiao, Y., Yan, H., Liu, H., et al. (2024a). SPL16 and SPL23 mediate photoperiodic control of seasonal growth in Populus trees. *New Phytol.* 241, 1646–1661. doi: 10.1111/nph.19485

Wei, H., Sun, F., Mo, J., Hu, B., and Luo, K. (2024b). Overexpression of *CRYPTOCHROME 2* enhances shoot growth and wood formation in poplar under growth-restrictive short days. *J. Genet. Genomics* 51, 1310–1313. doi: 10.1016/j.jgg.2024.08.003

Welling, A., and Palva, E. T. (2006). Molecular control of cold acclimation in trees. *Physiol. Plant.* 127, 167–181. doi: 10.1111/j.1399-3054.2006.00672.x

Wilkie, J. D., Sedgley, M., and Olesen, T. (2008). Regulation of floral initiation in horticultural trees. *J. Exp. Bot.* 59, 3215–3228. doi: 10.1093/jxb/ern188

Workmaster, B. A. A. (2001). Cold hardiness, ice nucleation, and growth modeling in the cranberry plant. 2001. Available at: https://search.library.wisc.edu/catalog/999926188702121

Workmaster, B. A. A., and Palta, J. P. (2006). Shifts in Bud and Leaf Hardiness during Spring Growth and Development of the Cranberry Upright: Regrowth Potential as an Indicator of Hardiness. *J. Am. Soc. Hortic. Sci.* 131, 327–337. doi: 10.21273/JASHS.131.3.327

Wu, R., Tomes, S., Karunairetnam, S., Tustin, S. D., Hellens, R. P., Allan, A. C., et al. (2017a). SVP-like MADS Box Genes Control Dormancy and Budbreak in Apple. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00477

Wu, R., Wang, T., Warren, B. A. W., Allan, A. C., Macknight, R. C., and Varkonyi-Gasic, E. (2017b). Kiwifruit SVP2 gene prevents premature budbreak during dormancy. *J. Exp. Bot.* 68, 1071–1082. doi: 10.1093/jxb/erx014

Yu, H., Xu, Y., Tan, E. L., and Kumar, P. P. (2002). AGAMOUS-LIKE 24, a dosage-dependent mediator of the flowering signals. *Proc. Natl. Acad. Sci.* 99, 16336–16341. doi: 10.1073/pnas.212624599

Yu, X., Liu, H., Klejnot, J., and Lin, C. (2010). The Cryptochrome Blue Light Receptors. *Arab. Book Am. Soc. Plant Biol.* 8, e0135. doi: 10.1199/tab.0135

Zhang, Y., Zheng, S., Liu, Z., Wang, L., and Bi, Y. (2011). Both HY5 and HYH are necessary regulators for low temperature-induced anthocyanin accumulation in *Arabidopsis* seedlings. *J. Plant Physiol.* 168, 367–374. doi: 10.1016/j.jplph.2010.07.025

Zhao, Y., Ma, Y., Qiu, H., Zhou, L., He, K., and Ye, Y. (2025). Wake up: the regulation of dormancy release and bud break in perennial plants. *Front. Plant Sci.* 16. doi: 10.3389/fpls.2025.1553953

Chapter V Conclusions & Recommendations for Future Research

From Upright Trait Assessment to High-Throughput Phenomics: Validation and New Applications in Breeding

This research demonstrates a progressive advancement in cranberry phenotyping methods and their application to genetic mapping and breeding. It first established a critical validation between traditional per-upright measurements and higher-throughput, higher-accuracy methods such as whole-plot assessment approaches and traits extracted from digital images. Cross-method, cross-study meta-QTL analysis revealed composite-trait associations linking traditional measurements such as upright berry width to image-derived berry width traits (Diaz-Garcia et al., 2016; Maule et al., 2024). By establishing these connections between established and emerging assessment methodologies, the research provides a foundation for enhanced phenotyping efficiency in large-scale breeding operations. Subsequently, the work developed UAV-based RGB imaging protocols to predict and monitor spring developmental patterns across breeding populations, and developed a novel selection index, the Late Maturity Index (LMI), to score genotypes on a continuous scale that balances delayed development for frost resilience with rapid post-frost growth characteristics to favor earlier harvest windows. Finally, successful genetic mapping of LMI in the CNJ02 cranberry breeding population (n = 176) identified quantitative trait loci that can guide breeding decisions, while preliminary candidate gene analysis provides evidence for the dormancy regulation mechanisms underlying these genetic regions.

Limitations and Areas for Improvement

Several constraints emerged that limit broader application of these methodologies. Mixed model convergence failures for populations CNJ04 and GRYG when analyzing LMI likely resulted from insufficient temporal data resolution combined with reduced population sizes. This highlights the critical importance of adequate data collection intensity and frequency for complex developmental traits that exhibit rapid temporal changes. In particular, the LMI trait appeared to exhibit the largest trait segregation pattern in a narrow window between 75-150 cumulative degree day heat units, so higher sampling rates in this window would better serve to tease apart genetic variation where more phenotypic variation is present.

The highly quantitative genetic architecture of LMI, evident by individual QTL explaining only 1.8 to 9.1 percent of genetic variance, indicates that other germplasm improvement approaches may prove more suitable than marker-assisted selection (MAS). Previous reviews have indicated that when QTL explain less than 10 percent of genetic variation, MAS is particularly ineffectual (Goddard and Hayes, 2007; Hayes et al., 2009). Moreover, constrained program resources present in many small crop breeding programs limit population sizes, in turn limiting the accuracy of QTL and their effect estimates (Dekkers and Hospital, 2002; Schön et al., 2004). The effectiveness of QTL and downstream MAS are additionally restricted by the scope of a few mapping populations: allelic diversity and genetic background effects available to a breeding program will not be captured with a single or a few biparental populations (Heffner et al., 2009). Validation and effect re-estimation in different genetic backgrounds, important for expanding genetic scope, further limit the efficiency of MAS derived from biparental populations (Holland, 2004). The complex polygenic architecture of LMI in cranberry population CNJ02 indicates that phenotypic selection will continue to be an important component to the genetic advancement of LMI in cranberry, absent applications of newer statistical tools, larger and more diverse populations, and more robust phenotype datasets (Bernardo and Charcosset, 2006; Bernardo, 2008).

As many of the LMI QTL are of lower effect size, their confidence intervals (LOD-1.5) proved relatively wide, reducing precision for candidate gene identification. Combining higher marker density approaches with association mapping on populations that represent more of the species standing genetic variation would better pinpoint the functional genetic elements responsible for differences in spring development (Xu et al., 2017). These genetic insights could both serve to validate the QTL in the existing germplasm resources while providing useful targets if gene editing becomes a viable genetic improvement approach. Even without a mechanistic understanding of the genetics controlling LMI variation, genomic prediction and selection methods supplemented with higher density maps are advanced methods that offer paths to rapid improvement int cranberries (Meuwissen et al., 2001; Desta and Ortiz, 2014).

Higher fidelity genetic maps combined with higher population genetic diversity are insufficient to effectively locate and estimate meaningful QTL or perform genomic prediction and selection. A high quality phenotyping dataset representing multiple years and environments of data is also critical, integrated with field designs meant to better decompose spatial variation and enhance accuracy of breeding value estimates . While the current research utilizes a large dataset captured over several growing seasons, it could benefit from a few optimizations and additions. The identification of specific developmental windows where phenotypic segregation reaches maximum levels suggests that strategic intensive sampling during key periods may provide greater value than uniform temporal coverage throughout growing seasons. Higher spatial and spectral resolution sensors would benefit the phenotyping dataset by providing more robust segmentation and isolation of discrete anatomical features (spatial resolution) while exhibiting higher sensitivity to physiological shifts driven by development and the environment (spectral resolution) (Aasen et al., 2018, 2020; Tanaka et al., 2024). The integration of proximal data from high density field data loggers (temperature, solar irradiance, humidity, and soil moisture) could serve to reinforce models and validate physiological responses to conditions, improving estimates of genetic and genetic-by-environment effects and variances.

Despite these limitations, a few features of this study's mapping lend to a more robust interpretation of QTL and their application to MAS. As the QTL were mapped onto genomic basic linear unbiased predictors (GBLUPs) instead of the original phenotypes, a practice not necessarily done at the time these critiques of MAS were published, one-to-one comparisons are not necessarily appropriate. These GBLUPs were derived from mixed models amended with marker-derived additive covariance matrices, allowing for better dissection of additive genetic variance and breeding values. Moreover, as GBLUPs are fit using L2-norm regularized mixed effect models, breeding value estimates are conservatively dampened, potentially under-representing genotypic effects (Meuwissen et al., 2001; Heffner et al., 2011). As such, QTL variances reported under 10 percent in this study would potentially be reported higher when mapped to the original phenotypes. Also, a few promising gene candidates regulating dormancy developmental were found in the top five LMI QTL, and strategically selecting QTL that represent high-impact upstream regulatory elements could be effective MAS targets.

Near-Term Research Priorities

Genetic Architecture Analysis and Validation

Immediate research priorities should focus on expanding the genetic analysis framework established for spring developmental traits. Candidate gene enrichment analysis for LMI quantitative trait loci should examine gene ontology patterns, protein family distributions, and regulatory domain enrichment within QTL intervals compared to genome-wide expectations. This systematic analysis could strengthen evidence for the dormancy-related regulatory genes identified and reveal additional functional categories relevant to spring development.

Multi-Trait Integration and Analysis

The research identified substantial additional UAV imagery and ground reference data for flower timing, berry set, berry sizing, berry coloration, and yield characteristics. These datasets represent valuable resources for modeling key cranberry developmental processes, particularly in their relation to yield. Inter-trait correlation analyses could provide crucial insights into relationships between spring development patterns and subsequent reproductive performance. Integrating these relationships across time could also increase the power to find QTL and refine the accuracy of their respective loci (Miao et al., 2020; Adak et al., 2023).

Following the methodological framework established in mapping the LMI trait, QTL can be mapped for flower timing, berry set, berry sizing, berry coloration, and yield. Further guided by inter-trait correlation blocks, comprehensive metaQTL analysis should help identify loci that co-locate across multiple traits (Maule et al., 2024). These collocated QTL could represent regulatory genomic regions that coordinate multiple aspects of cranberry development and reproduction, offering targets for simultaneous improvement of multiple characteristics. As such, these metaQTL will likely inform the most stable and broadly applicable genetic targets for breeding applications.

Longer-Term Research Objectives

Advanced Phenotyping and Data Integration

The insights gained from model development and QTL mapping should inform more sophisticated data collection strategies that leverage emerging genomic resources for *Vaccinium* species, including the recently developed pangenome resources that reveal substantial genetic diversity with 48.4% auxiliary orthogroups in highbush blueberry and 47.0% auxiliary genes in cranberry (Yocca et al., 2023). Auxiliary genes are often dispensable and are usually enriched for proteins involved in stress response, environmental adaptation, metabolic flexibility, or specialized functions that become important under particular conditions (Tettelin et al., 2005; Matthews et al., 2024). The identification of critical developmental windows suggests that targeted intensive sampling during key periods, combined with pangenomic information, may better illuminate the genetic architecture of environmental adaptation in cranberries. These insights can both help in QTL validation while also allowing for optimized breeding.

Data fusion approaches integrating multiple sensor modalities could address limitations identified with RGB-only imaging systems. While the research demonstrated effective capture of visible pigment changes, spectral resolution constraints limit detection of subtle physiological changes or stress responses that may be relevant for comprehensive phenotyping applications. Research has already demonstrated the value of monitoring leaf canopy nutrition in cranberry beds using hyperspectral sensors, but increasing the spatial resolution to characterize small breeding plots will be necessary (Liu et al., 2023; Huang et al., 2024). Thermal imaging could provide a detailed picture of drought and heat stress responses in different genotypes. Low cost sensors placed at high density could provide more detailed maps of field spatial variation and identify key environmental distinctions and extreme weather conditions. These low cost sensors combined with RGB cameras have already demonstrated utility in creating accurate crop risk models in cranberry, providing a warning system to growers to manage heat scald risk effectively (Akiva et al., 2020, 2021).

Population Genetics and Genomic Selection Implementation

Characterizing standing genetic variation through diversity panel studies represents a crucial step for translating research findings into practical breeding applications, particularly given successful genomic selection implementation in blueberry breeding programs that has demonstrated improved accuracy and reduced breeding cycle times (Ferrão et al., 2021). The current research focused on biparental populations that capture only a fraction of genetic diversity available in cranberry germplasm collections, although this limitation is partly alleviated by the use of a three-population composite linkage map for evaluating cross-population QTL (Schlautman et al., 2017).

Diversity panels would enable genome-wide association studies that could identify additional QTL and assess broader applicability of findings across cranberry genetic backgrounds, building upon coordinated transdisciplinary research approaches being developed through initiatives like VacciniumCAP for marker-assisted selection implementation in *Vaccinium* breeding programs (Edger et al., 2022).

Genomic Prediction Model Development

The development of genomic prediction models integrating high-throughput phenotyping data with diversity panels offers substantial potential for transforming cranberry breeding efficiency, particularly given advances in target capture sequencing methods that enable accurate allele dosage estimation and consistent read depths across samples and loci (Edger et al., 2022). Such models could enable selection decisions based on genetic markers and early-season phenotyping data, potentially reducing the six to eight year evaluation cycles that currently constrain genetic progress.

Preliminary candidate gene analysis provides several targets for functional validation studies, potentially through transcriptomic analysis of genotypes with contrasting LMI values or through comparative studies with related *Vaccinium* species where functional genomics tools may be more advanced, building upon the remarkable genome collinearity observed across *Vaccinium* species that facilitates cross-species research translation (Diaz-Garcia et al., 2021; Yocca et al., 2023).

Community Resource Development and Integration

Future research should leverage the Genome Database for *Vaccinium* as a community platform for integrating genomic, transcriptomic, metabolomic and proteomic datasets, which is essential for understanding how gene-metabolite networks determine specific traits relevant to breeding applications (Edger et al., 2022). The high-throughput phenotyping approaches developed in this research could contribute valuable datasets to such community resources.

The methodological frameworks established for UAV-based phenotyping and genetic mapping provide templates that could be adapted for other economically important traits and extended to related *Vaccinium* species. Such expansion could accelerate breeding progress across the broader *Vaccinium* community while enabling comparative studies that enhance understanding of trait evolution and adaptation mechanisms.

This research establishes important methodological foundations for integrating advanced phenotyping technologies with genetic analysis in cranberry breeding. The successful development of cost-effective UAV-based approaches for monitoring complex developmental traits, combined with robust genetic mapping of spring phenological patterns, provides tools and insights that can accelerate development of climate-resilient cranberry cultivars while maintaining production efficiency and fruit quality characteristics demanded by modern markets.

References

Aasen, H., Honkavaara, E., Lucieer, A., and Zarco-Tejada, P. J. (2018). Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: A review of sensor technology, measurement procedures, and data correction workflows. *Remote Sens.* 10, 1–42. doi: 10.3390/rs10071091

Aasen, H., Kirchgessner, N., Walter, A., and Liebisch, F. (2020). PhenoCams for Field Phenotyping: Using Very High Temporal Resolution Digital Repeated Photography to Investigate Interactions of Growth, Phenology, and Harvest Traits. *Front. Plant Sci.* 11, 593. doi: 10.3389/fpls.2020.00593

Adak, A., Murray, S. C., and Anderson, S. L. (2023). Temporal phenomic predictions from unoccupied aerial systems can outperform genomic predictions. *G3 GenesGenomesGenetics* 13, jkac294. doi: 10.1093/g3journal/jkac294

Akiva, P., Dana, K., Oudemans, P., and Mars, M. (2020). Finding Berries: Segmentation and Counting of Cranberries using Point Supervision and Shape Priors., in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, (Seattle, WA, USA: IEEE), 219–228. doi: 10.1109/CVPRW50498.2020.00033

Akiva, P., Planche, B., Roy, A., Dana, K., Oudemans, P., and Mars, M. (2021). AI on the Bog: Monitoring and Evaluating Cranberry Crop Risk., in *2021 IEEE Winter Conference on*

Applications of Computer Vision (WACV), (Waikoloa, HI, USA: IEEE), 2492–2501. doi: 10.1109/WACV48630.2021.00254

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: Learning from the last 20 years. *Crop Sci.* 48, 1649–1664. doi: 10.2135/cropsci2008.03.0131

Bernardo, R., and Charcosset, A. (2006). Usefulness of Gene Information in Marker-Assisted Recurrent Selection: A Simulation Appraisal. *Crop Sci.* 46, 614–621. doi: 10.2135/cropsci2005.05-0088

Dekkers, J. C. M., and Hospital, F. (2002). The use of molecular genetics in the improvement of agricultural populations. *Nat. Rev. Genet.* 3, 22–32. doi: 10.1038/nrg701

Desta, Z. A., and Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19, 592–601. doi: 10.1016/j.tplants.2014.05.006

Diaz-Garcia, L., Garcia-Ortega, L. F., González-Rodríguez, M., Delaye, L., Iorizzo, M., and Zalapa, J. (2021). Chromosome-Level Genome Assembly of the American Cranberry (*Vaccinium macrocarpon* Ait.) and Its Wild Relative *Vaccinium microcarpum*. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.633310

Edger, P. P., Iorizzo, M., Bassil, N. V., Benevenuto, J., Ferrão, L. F. V., Giongo, L., et al. (2022). There and back again; historical perspective and future directions for *Vaccinium* breeding and research studies. *Hortic. Res.* 9, uhac083. doi: 10.1093/hr/uhac083

Ferrão, L. F. V., Amadeu, R. R., Benevenuto, J., de Bem Oliveira, I., and Munoz, P. R. (2021). Genomic Selection in an Outcrossing Autotetraploid Fruit Crop: Lessons From Blueberry Breeding. *Front. Plant Sci.* 12, 676326. doi: 10.3389/fpls.2021.676326

Goddard, M. E., and Hayes, B. J. (2007). Genomic selection. *J. Anim. Breed. Genet.* 124, 323–330. doi: 10.1111/j.1439-0388.2007.00702.x

Hayes, B. J., Bowman, P. J., Chamberlain, A. J., and Goddard, M. E. (2009). *Invited review*: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* 92, 433–443. doi: 10.3168/jds.2008-1646

Heffner, E. L., Jannink, J.-L., Iwata, H., Souza, E., and Sorrells, M. E. (2011). Genomic Selection Accuracy for Grain Quality Traits in Biparental Wheat Populations. *Crop Sci.* 51, 2597–2606. doi: 10.2135/cropsci2011.05.0253

Heffner, E. L., Sorrells, M. E., and Jannink, J.-L. (2009). Genomic Selection for Crop Improvement. *Crop Sci.* 49, 1–12. doi: 10.2135/cropsci2008.08.0512

Holland, J. B. (2004). Implementation of molecular markers for quantitative traits in breeding programs - challenges and opportunities.

Huang, Y., Liu, N., Wagner Hokanson, E., Hansen, N., and Townsend, P. A. (2024). Exploring the potential of multi-source satellite remote sensing in monitoring crop nutrient status: A multi-year case study of cranberries in Wisconsin, USA. *Int. J. Appl. Earth Obs. Geoinformation* 132, 104063. doi: 10.1016/j.jag.2024.104063

Liu, N., Wagner Hokanson, E., Hansen, N., and Townsend, P. A. (2023). Multi-year hyperspectral remote sensing of a comprehensive set of crop foliar nutrients in cranberries. *ISPRS J. Photogramm. Remote Sens.* 205, 135–146. doi: 10.1016/j.isprsjprs.2023.10.003

Matthews, C. A., Watson-Haigh, N. S., Burton, R. A., and Sheppard, A. E. (2024). A gentle introduction to pangenomics. *Brief. Bioinform.* 25, bbae588. doi: 10.1093/bib/bbae588

Maule, A. F., Loarca, J., Diaz-Garcia, L., Lopez-Moreno, H., Johnson-Cicalese, J., Vorsa, N., et al. (2024). Of buds and bits: a meta-QTL study identifies stable QTL for berry quality and yield traits in cranberry mapping populations (*Vaccinium macrocarpon* Ait.). *Front. Plant Sci.* 15, 1294570. doi: 10.3389/fpls.2024.1294570

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157, 1819–1829. doi: 10.1093/genetics/157.4.1819

Miao, C., Xu, Y., Liu, S., Schnable, P. S., and Schnable, J. C. (2020). Increased Power and Accuracy of Causal Locus Identification in Time Series Genome-wide Association in Sorghum. *Plant Physiol.* 183, 1898–1909. doi: 10.1104/pp.20.00277

Schlautman, B., Covarrubias-Pazaran, G., Diaz-Garcia, L., Iorizzo, M., Polashock, J., Grygleski, E., 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 GenesGenomesGenetics* 7, 1177–1189. doi: 10.1534/g3.116.037556

Schön, C. C., Utz, H. F., Groh, S., Truberg, B., Openshaw, S., and Melchinger, A. E. (2004). Quantitative Trait Locus Mapping Based on Resampling in a Vast Maize Testcross Experiment and Its Relevance to Quantitative Genetics for Complex Traits. *Genetics* 167, 485–498. doi: 10.1534/genetics.167.1.485

Tanaka, T. S. T., Wang, S., Jørgensen, J. R., Gentili, M., Vidal, A. Z., Mortensen, A. K., et al. (2024). Review of Crop Phenotyping in Field Plot Experiments Using UAV-Mounted Sensors and Algorithms. *Drones* 8, 212. doi: 10.3390/drones8060212

Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., et al. (2005). Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: Implications for the microbial "pan-genome." *Proc. Natl. Acad. Sci. U. S. A.* 102, 13950–13955. doi: 10.1073/pnas.0506758102

Xu, Y., Li, P., Yang, Z., and Xu, C. (2017). Genetic mapping of quantitative trait loci in crops. *Crop J.* 5, 175–184. doi: 10.1016/j.cj.2016.06.003

Yocca, A. E., Platts, A., Alger, E., Teresi, S., Mengist, M. F., Benevenuto, J., et al. (2023). Blueberry and cranberry pangenomes as a resource for future genetic studies and breeding efforts. *bioRxiv*, 2023.07.31.551392. doi: 10.1101/2023.07.31.551392

Appendix A Supplementary Data for Chapter III

Supplementary Table A-1. Pairwise correlations of all VIs used when fitting regression models. These correlations are derived from all UAS images of cranberry plots, including those without ground truth reference data.

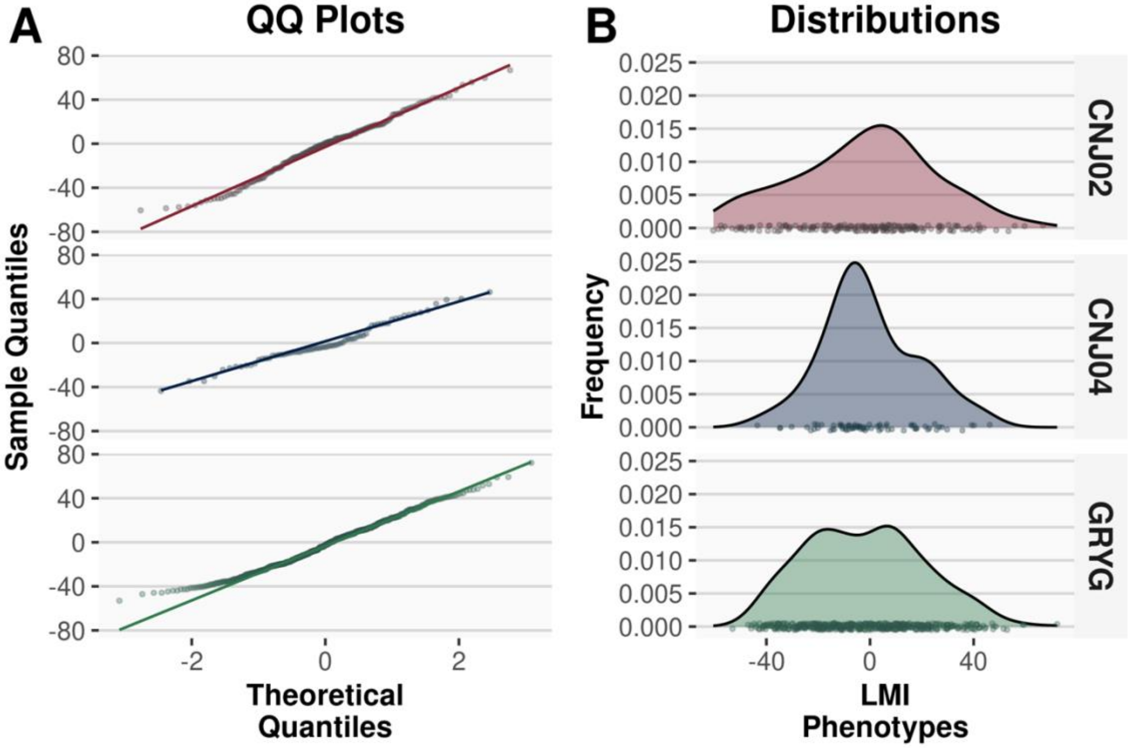
	gli	**ebi**	**ac.cb**	**egiMeri**	**cive**	**eri**	**lab.b**	**vari**	**woi**	**veg**	**rgb.B**	**rgb.R**	**hsv.h**	**ac.cr**	**hsv.s**	**gga**	**ga**	**ipca**	**kai**
gli	1.00	0.91	0.95	0.99	0.92	0.93	0.76	0.88	0.76	0.72	0.64	0.73	0.54	0.63	0.39	0.43	0.38	0.31	0.10
ebi	0.91	1.00	0.87	0.84	0.81	0.69	0.77	0.68	0.82	0.84	0.90	0.39	0.66	0.62	0.73	0.37	0.38	0.35	0.50
ac.cb	0.95	0.87	1.00	0.93	0.91	0.87	0.77	0.82	0.82	0.73	0.63	0.68	0.60	0.46	0.34	0.45	0.44	0.32	0.11
egiMeri	0.99	0.84	0.93	1.00	0.92	0.97	0.72	0.93	0.69	0.64	0.52	0.83	0.46	0.64	0.27	0.43	0.35	0.28	0.05
cive	0.92	0.81	0.91	0.92	1.00	0.88	0.89	0.89	0.70	0.63	0.54	0.73	0.52	0.51	0.30	0.40	0.35	0.54	0.01
eri	0.93	0.69	0.87	0.97	0.88	1.00	0.63	0.94	0.57	0.49	0.31	0.94	0.32	0.58	0.05	0.41	0.31	0.22	0.28
lab.b	0.76	0.77	0.77	0.72	0.89	0.63	1.00	0.65	0.73	0.58	0.63	0.43	0.58	0.37	0.43	0.32	0.33	0.83	0.24
vari	0.88	0.68	0.82	0.93	0.89	0.94	0.65	1.00	0.54	0.51	0.34	0.86	0.29	0.58	0.13	0.39	0.25	0.26	0.22
woi	0.76	0.82	0.82	0.69	0.70	0.57	0.73	0.54	1.00	0.69	0.73	0.32	0.57	0.29	0.50	0.37	0.40	0.37	0.41
veg	0.72	0.84	0.73	0.64	0.63	0.49	0.58	0.51	0.69	1.00	0.81	0.21	0.57	0.45	0.65	0.32	0.38	0.21	0.52
rgb.B	0.64	0.90	0.63	0.52	0.54	0.31	0.63	0.34	0.73	0.81	1.00	0.05	0.67	0.46	0.92	0.24	0.31	0.32	0.82
rgb.R	0.73	0.39	0.68	0.83	0.73	0.94	0.43	0.86	0.32	0.21	0.05	1.00	0.09	0.44	0.29	0.34	0.20	0.11	0.60
hsv.h	0.54	0.66	0.60	0.46	0.52	0.32	0.58	0.29	0.57	0.57	0.67	0.09	1.00	0.15	0.45	0.26	0.37	0.26	0.45
ac.cr	0.63	0.62	0.46	0.64	0.51	0.58	0.37	0.58	0.29	0.45	0.46	0.44	0.15	1.00	0.43	0.29	0.18	0.13	0.14
hsv.s	0.39	0.73	0.34	0.27	0.30	0.05	0.43	0.13	0.50	0.65	0.92	0.29	0.45	0.43	1.00	0.11	0.16	0.27	0.90
gga	0.43	0.37	0.45	0.43	0.40	0.41	0.32	0.39	0.37	0.32	0.24	0.34	0.26	0.29	0.11	1.00	0.95	0.11	0.00
ga	0.38	0.38	0.44	0.35	0.35	0.31	0.33	0.25	0.40	0.38	0.31	0.20	0.37	0.18	0.16	0.95	1.00	0.12	0.13
ipca	0.31	0.35	0.32	0.28	0.54	0.22	0.83	0.26	0.37	0.21	0.32	0.11	0.26	0.13	0.27	0.11	0.12	1.00	0.18
kai	0.10	0.50	0.11	0.05	0.01	0.28	0.24	0.22	0.41	0.52	0.82	0.60	0.45	0.14	0.90	0.00	0.13	0.18	1.00

Supplementary Table A-2. Pairwise correlations of VIs, spectrophotometer absorbances, and estimated CDD on reference data gathered from cranberry plots.

	CDD	**A535nmMean**
ipca	0.57	-0.40
hsv.h	-0.49	0.69
kai	0.08	-0.21
rgb.B	-0.55	0.66
hsv.s	0.33	-0.42
ga	0.39	-0.39
gga	0.43	-0.39
lab.b	0.87	-0.76
CDD	1.00	-0.81
lab.a	-0.91	0.85
cive	-0.93	0.84
rgbvi	0.80	-0.81
ebi	-0.78	0.84
ndyi	0.76	-0.82
A649nm_A535nm	0.83	-0.84
A6XXnm_A535nm	0.84	-0.84
A660nm_A535nm	0.84	-0.85
A535nmMean	-0.81	1.00
ac.cb	-0.86	0.93
egiMeri	0.84	-0.88
gli	0.85	-0.90
rgb.G	0.85	-0.89
egi	0.85	-0.89
woi	0.72	-0.75
veg	0.67	-0.81
A660nmMean	0.41	-0.35
A649nmMean	0.36	-0.29
csi	-0.66	0.64
rgb.R	-0.62	0.62
vari	0.81	-0.76
eri	-0.79	0.82
ndi	0.82	-0.85
ac.cr	0.46	-0.41
mgrvi	0.62	-0.57

	CDD	A535nmMean
r:g	-0.63	0.58

Appendix B Supplementary Data for Chapter IV.



Supplementary Figure B-1**.** Leaf maturity index (LMI) trait QQ (**A**) and distribution (**B**) plots of three cranberry breeding populations CNJ02, CNJ04, and GRYG.

Supplementary Table B-1. Gene counts in the genomic neighborhood of QTL for the LMI trait in cranberry population CN02. Neighborhood is defined as the genomic regions associated with the LOD 1.5 intervals of QTL. These counts are for all gene models that have homology to any characterized genes or motifs in the scientific literature. Gene models were taken from annotations in the Stevens (Diaz-Garcia et al., 2021) and Ben Lear (Kawash et al., 2022) cranberry genome assemblies.

LG	**Position (cM)**	**R_q^2 (%)**	**Gene Count (Ben Lear)**	**Gene Count (Stevens)**
11	47.62	9.12	58	213
10	51.79	7.50	21	117
12	42.91	7.22	34	103
7	16.77	6.14	27	110
6	60.98	4.45	83	263
8	49.95	3.37	6	56
11	11.07	3.20	121	403
8	25.16	3.02	180	513
12	7.91	2.49	67	303
7	51.14	1.98	86	226
9	52.54	1.78	342	1,037

Supplementary Table B-2. Putative cranberry candidate genes relevant to dormancy processes that are in the genetic neighborhood of QTL for the LMI trait in cranberry population CNJ02. Sequences of the flanking 1.5 LOD interval markers were aligned to the Stevens (Diaz-Garcia et al., 2021) and Ben Lear (Kawash et al., 2022) annotated cranberry genome assemblies.

QTL (LG@ cM)	**R_q^2 (%)**	**Genome**	**Genome Start (MBP)**	**Genome End (MBP)**	**Gene Model ID**	**Putative Product**	**Gene Ontology**	**DB Refs**
11@47.62	9.12	Ben Lear	16.657	16.659	Vmac_052489	Protein RADIALIS-like 5	GO:0003677	PF00249, IPR001005, IPR017884, IPR009057
			17.848	17.851	Vmac_052633	Agamous-like MADS-box protein mads1	GO:0003700, GO:0006355, GO:0005634	PF01486, IPR002487
			19.036	19.043	Vmac_052783	Alkaline/neutral invertase cinv2,CAZy:GH100	GO:0003824, GO:0033926	PF12899, IPR012341, IPR024746, IPR008928
			27.241	27.241	Vmac_053786	DNA topoisomerase 2	GO:0003918, GO:0003677, GO:0005524, GO:0006265	IPR002205, IPR001241, IPR013760
		Stevens	15.397	15.399	vmacro12405	ASR2: Abscisic stress-ripening protein 2		
			15.403	15.404	vmacro12406	ASR1: Abscisic stress-ripening protein 1		
			15.894	15.898	vmacro12431	FLS: Flavonol synthase/flavanone 3-hydroxylase		
			16.437	16.441	vmacro12448	MADS1: Agamous-like MADS-box protein MADS1		
			21.837	21.855	vmacro12638	PRR95: Two-component response regulator-like PRR95		
			22.054	22.057	vmacro12634	LHY: Protein LHY		
			24.925	24.931	vmacro12552	FL: Flavonol synthase/flavanone 3-hydroxylase		

QTL (LG@ cM)	R_q^2 (%)	Genome	Genome Start (MBP)	Genome End (MBP)	Gene Model ID	Putative Product	Gene Ontology	DB Refs
10@51.79	7.50	Ben Lear	10.085	10.087	Vmac_047478	Transcription factor bHLH93	GO:0046983	PF00010, IPR011598, IPR036638
			10.796	10.796	Vmac_047550	MADS-box protein agl24	GO:0046983, GO:0003677	PF00319, IPR002100, IPR036879
		Stevens	9.678	9.685	vmacro21337	SVP: MADS-box protein SVP		
			12.867	12.873	vmacro21266	CRY1: Cryptochrome-1		
7@16.77	6.14	Stevens	5.546	5.555	vmacro13560	CYP75A3: Flavonoid 3'%2C5'-hydroxylase 2		
6@60.98	4.45	Ben Lear	27.137	27.139	Vmac_030491	FT-interacting protein 1		PF00168, PF08372, IPR000008, IPR035892, IPR013583
			27.145	27.148	Vmac_030492	FT-interacting protein 3		PF00168, PF08372, IPR000008, IPR035892, IPR013583
			27.329	27.331	Vmac_030511	Delta(12)-fatty-acid desaturase	GO:0016717, GO:0055114, GO:0006629	PF00487, IPR021863, IPR005804
			29.153	29.156	Vmac_030731	Peroxidase 52,CAZy:AA2	GO:0020037, GO:0004601, GO:0006979, GO:0055114, GO:0042744	PF00141, IPR000823, IPR002016, IPR019794, IPR019793, IPR033905, IPR010255
			31.554	31.555	Vmac_031035	Peroxidase 3,CAZy:AA2	GO:0020037, GO:0004601,	PF00141, IPR000823, IPR002016,

QTL (LG@ cM)	R_q^2 (%)	Genome	Genome Start (MBP)	Genome End (MBP)	Gene Model ID	Putative Product	Gene Ontology	DB Refs
							GO:0006979, GO:0055114	IPR019794, IPR010255
			31.589	31.590	Vmac_031037	Peroxidase 3,CAZy:AA2,SECRETED:SignalP(1-23)	GO:0020037, GO:0004601, GO:0006979, GO:0055114, GO:0042744	PF00141, IPR000823, IPR002016, IPR019793, IPR019794, IPR033905, IPR010255
		Stevens	8.141	8.143	vmacro09323	COL12: Zinc finger protein CONSTANS-LIKE 12		
			8.532	8.536	vmacro09295	Anthocyanin 5-aromatic acyltransferase		
			9.536	9.539	vmacro09252	CYP75A3: Flavonoid 3'%2C5'-hydroxylase 2		
			9.708	9.709	vmacro09245	GASA10: Gibberellin-regulated protein 10		
			10.991	10.993	vmacro09184	GT1: Anthocyanidin 3-O-glucosyltransferase 1		
			11.980	11.989	vmacro09117	FAD2: Delta(12)-fatty-acid desaturase FAD2		
			12.111	12.113	vmacro09109	SPL15: Squamosa promoter-binding-like protein 15		
11@11.07	3.20	Ben Lear	6.156	6.158	Vmac_051257	Anthocyanidin 3-O-glucoside 2'''-O-xylosyltransferase	GO:0016758	IPR002213
			7.628	7.630	Vmac_051435	Gibberellin 2-beta-dioxygenase 1	GO:0016491, GO:0055114	PF03171, PF14226, IPR027443, IPR026992, IPR005123
		Stevens	6.974	6.977	vmacro12010	GA2OX1: Gibberellin 2-beta-dioxygenase 1		

QTL (LG@cM)	R_q^2 (%)	Genome	Genome Start (MBP)	Genome End (MBP)	Gene Model ID	Putative Product	Gene Ontology	DB Refs
			7.811	7.817	vmacro12062	COL10: Zinc finger protein CONSTANS-LIKE 10		
			11.159	11.163	vmacro12230	COL12: Zinc finger protein CONSTANS-LIKE 12		
8@25.16	3.02	Ben Lear	3.537	3.537	Vmac_036921	Abscisic acid receptor pyl4		PF03364, PF10604, IPR023393, IPR019587
			6.207	6.209	Vmac_037276	Anthocyanidin 3-O-galactosyltransferase f3gt1	GO:0016758	IPR002213
			6.214	6.215	Vmac_037277	Anthocyanidin 3-O-galactosyltransferase f3gt1	GO:0016758	IPR002213, IPR035595
			6.218	6.222	Vmac_037278	Anthocyanidin 3-O-galactosyltransferase f3gt1,CAZy:GT1	GO:0016758	IPR002213, IPR035595
			6.223	6.230	Vmac_037279	Anthocyanidin 3-O-galactosyltransferase f3gt1,CAZy:GT1	GO:0016758	IPR002213
			6.250	6.261	Vmac_037284	Anthocyanidin 3-O-galactosyltransferase f3gt1	GO:0016758	IPR002213
		Stevens	4.097	4.099	vmacro10037	GA2OX2: Gibberellin 2-beta-dioxygenase 2		
			5.618	5.620	vmacro04862	F3GT1: Anthocyanidin 3-O-galactosyltransferase F3GT1		
			5.626	5.628	vmacro04861	F3GT1: Anthocyanidin 3-O-galactosyltransferase F3GT1		
			5.630	5.636	vmacro04860	F3GT1: Anthocyanidin 3-O-galactosyltransferase F3GT1		
			5.641	5.642	vmacro04859	F3GT1: Anthocyanidin 3-O-galactosyltransferase F3GT1		
			5.660	5.663	vmacro04858	F3GT1: Anthocyanidin 3-O-galactosyltransferase F3GT1		
			5.670	5.672	vmacro04857	UGT78D2: Flavonol 3-O-glucosyltransferase		
			10.784	10.792	vmacro10528	LAR: Leucoanthocyanidin reductase		

QTL (LG@ cM)	R_q^2 (%)	Genome	Genome Start (MBP)	Genome End (MBP)	Gene Model ID	Putative Product	Gene Ontology	DB Refs
12@7.91	2.49	Ben Lear	2.313	2.316	Vmac_055985	Mitogen-activated protein kinase ntf3	GO:0005524, GO:0004672, GO:0004707, GO:0006468	PF06293, PF00069, PF07714, IPR000719, IPR008271, IPR003527, IPR017441, IPR011009
			2.386	2.388	Vmac_055996	Peroxidase 9,CAZy:AA2,SECRETED:SignalP(1-24)	GO:0004601, GO:0020037, GO:0006979, GO:0055114, GO:0042744	PF00141, IPR002016, IPR000823, IPR019794, IPR033905, IPR010255
		Stevens	2.005	2.008	vmacro04354	FHT: Flavanone 3-dioxygenase		
			3.331	3.334	vmacro04446	ABF2: ABSCISIC ACID-INSENSITIVE 5-like protein 5		
7@51.14	1.98	Ben Lear	17.211	17.212	Vmac_034128	Abscisic acid receptor pyl1		PF03364, PF10604, IPR023393, IPR019587
			20.620	20.621	Vmac_034511	Abscisic acid receptor pyl2		PF10604, IPR023393, IPR019587
		Stevens	18.536	18.540	vmacro14236	CYP707A2: Abscisic acid 8'-hydroxylase CYP707A2		
			19.583	19.583	vmacro14253	PYL2: Abscisic acid receptor PYL2		
			21.201	21.210	vmacro14281	DPBF3: ABSCISIC ACID-INSENSITIVE 5-like protein 2		
			23.389	23.391	vmacro22508	CET2: CEN-like protein 2		
9@52.54	1.78	Ben Lear	21.892	21.894	Vmac_043770	Peroxidase 42,CAZy:AA2,SECRETED:SignalP(1-25)	GO:0020037, GO:0004601, GO:0006979,	PF00141, IPR000823, IPR002016,

QTL (LG@ cM)	R_q^2 (%)	Genome	Genome Start (MBP)	Genome End (MBP)	Gene Model ID	Putative Product	Gene Ontology	DB Refs
							GO:0055114, GO:0042744	IPR033905, IPR010255
			25.269	25.273	Vmac_044169	FT-interacting protein 1		PF00168, PF08372, IPR000008, IPR013583, IPR035892
			25.283	25.286	Vmac_044172	FT-interacting protein 4		PF00168, PF08372, IPR000008, IPR013583, IPR035892
			36.607	36.608	Vmac_045583	Gibberellin-regulated protein 12		PF02704, IPR003854
		Stevens	3.670	3.671	vmacro02655	F3H-2: Flavanone 3-dioxygenase 2		
			4.065	4.066	vmacro02689	ASR3: Abscisic stress-ripening protein 3		
			4.078	4.078	vmacro02690	ASR3: Abscisic stress-ripening protein 3		
			4.085	4.086	vmacro02691	ASR3: Abscisic stress-ripening protein 3		
			4.103	4.104	vmacro02692	ASR3: Abscisic stress-ripening protein 3		
			4.264	4.266	vmacro02708	PDLP2: Plasmodesmata-located protein 2		
			4.565	4.573	vmacro02732	PIF1: Transcription factor PIF1		
			4.912	4.913	vmacro02769	GASA4: Gibberellin-regulated protein 4		
			8.389	8.397	vmacro03044	CYP707A2: Abscisic acid 8'-hydroxylase 2		
			20.076	20.077	vmacro03708	GA2OX8: Gibberellin 2-beta-dioxygenase 8		

QTL (LG@ cM)	R_q^2 (%)	Genome	Genome Start (MBP)	Genome End (MBP)	Gene Model ID	Putative Product	Gene Ontology	DB Refs
			20.083	20.083	vmacro03707	GA2OX8: Gibberellin 2-beta-dioxygenase 8		