Genomic heritability estimates in sweet cherry reveal non-additive genetic variance is relevant for industry-prioritized traits

J. Piaskowski* (corresponding author)

Craig Hardner[†]

Lichun Cai[‡]

Yunyang Zhao[§]

Amy Iezzoni[‡]

Cameron Peace*

*Department of Horticulture, Washington State University, Pullman, WA 99164-6414

[†]Centre for Plant Science, Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Brisbane, Australia

[‡]Department of Horticulture, Michigan State University, East Lansing, MI 48824-1325

[§]Plant for Human Health Institute, North Carolina State University, Kannapolis, NC 28081

ABSTRACT

Background: Sweet cherry is consumed widely across the world and provides substantial economic benefits in regions where it is grown. While cherry breeding has been conducted in the Pacific Northwest for over half a century, little is known about the genetic architecture of important traits. We used a genome-enabled mixed model to predict the genetic performance of 505 individuals for 32 phenological, disease response and fruit quality traits evaluated in the RosBREED sweet cherry crop data set. Genome-wide predictions were estimated using a repeated measures model for phenotypic data across 3 years, incorporating additive, dominance and epistatic variance components. Genomic relationship matrices were constructed with high-density SNP data and were used to estimate relatedness and account for incomplete replication across years.

Results: High broad-sense heritabilities of 0.83, 0.77, and 0.75 were observed for days to maturity, firmness, and fruit weight, respectively. Epistatic variance exceeded 40% of the total genetic variance for maturing timing, firmness and powdery mildew response. Dominance variance was the largest for fruit weight and fruit size at 34% and 27%, respectively. Omission of non-additive sources of genetic variance from the genetic mode resulted in inflation of narrow-sense heritability but minimally influenced prediction accuracy of genetic values in validation. Predicted genetic rankings of individuals from single-year models were inconsistent across years, likely due to incomplete sampling of the population genetic variance.

Conclusions: Predicted breeding values and genetic values a measure revealed many high-performing individuals for use as parents and the most promising selections to advance for cultivar release consideration, respectively. This study highlights the importance of using the appropriate genetic model for calculating breeding values to avoid inflation of expected parental contribution to genetic gain. The genomic predictions obtained will enable breeders to efficiently leverage the genetic potential of North American sweet cherry germplasm by identifying high quality individuals more rapidly than with phenotypic data alone.

Keywords:

- GBLUP
- sweet cherry
- *Prunus*
- genomic selection
- non-additive genetic variation

BACKGROUND

Sweet cherry (*Prunus avium* L.) is a lucrative fresh market horticultural crop whose monetary worth is directly and indirectly determined by several horticultural and fruit traits. Worldwide, more than 2.8 million tons of sweet cherry fruit were produced in 2014 [1]. In 2015, the U.S. was the second largest producer of cherries, producing 338.6 kilotons of fruit valued at $703 million, of which 60% were grown in Washington State [2,3].

Sweet cherry cultivars must garner a positive critical reception among growers, market intermediaries (a category which includes packers, shippers, and marketers), and consumers to succeed commercially. The U.S. sweet cherry industry and consumers have previously prioritized which fruit trait thresholds are essential for a successful cultivar. Sweet cherry producers identified fruit size, flavor, firmness, and powdery mildew resistance as trait priorities in a survey conducted in 2011 [4]. Powdery mildew (causative agent *Podosphaera clandestine*) is a foliar and fruit disease with a high cost of control in susceptible cultivars. Sweetness and flavor were ranked by consumers as the most important attributes in sweet cherry, followed by firmness, shelf life, and fruit size [5]. Consumers are willing to pay more for sweet, firm cherries with an ideal balance of sweetness and acidity. Sweetness and acidity are quantified with assays for soluble solids content (SSC) and titratable acidity (TA), respectively [5–8]. Market intermediaries indicated a willingness to pay producers more per pound for fruit greater than 2.5 cm in diameter, firmness above 300 g/mm, and SSC above 18° Brix [9]. Market intermediaries also ranked fruit size as the most important trait, followed by firmness and external appearance [10]. The USDA Agriculture Marketing Service evaluates skin color, fruit size, and fruit firmness when grading sweet cherries [11], an assessment which influences market receipts for that crop.

Many of the trait thresholds identified by consumers and the cherry industry alike have been individually met or exceeded through genetic improvement. Beginning with the 1952 release of ‘Rainier’, a highly popular sweet cherry cultivar, the Washington State University sweet cherry program (formerly USDA-ARS) has released several dozen cultivars with improved flavor, size, and firmness in each subsequent release [12,13]. This program and others have largely relied on phenotypic selection complemented with trait-predictive DNA tests for high heritability traits, such as fruit skin color and self-compatibility [13–16]. The Washington State University breeding program has seen genetic gains in fruit

dimensions, firmness and other traits of breeding relevance due to moderate heritability of those traits [17–19].

Sweet cherry has a juvenility period of three to five years before a tree is capable of flowering and producing fruit [20]. Therefore, the pace of cultivar release is slow, taking 15 to 25 years between making a cross to cultivar release [16]. Sweet cherry breeding is structured like many other crops: an initial set of crosses is made, followed by evaluation of a large number of offspring. After a rapid screening, the majority of these offspring is discarded, and the remaining selections are evaluated more extensively in replicated trials. Selections are clonally propagated in subsequent evaluations. Consequently, the genetic potential identified in F1 seedlings remains fixed throughout the evaluative phases of a breeding program and is not lost during recombination and segregation.

Understanding the genetic architecture of crop traits can help plant geneticists and allied scientists maximize genetic gain and elucidate the genetic potential of seedlings and parents. Best linear unbiased prediction (BLUP) is an analysis tool that is used to estimate the genetic potential of each individual from unbalanced trials by modeling genetic effects as a random effect in a mixed model [21]. It requires prior estimation of genetic variance components, which are obtained through maximum likelihood, restricted maximum likelihood (REML) or Bayesian approaches [22,23]. Pedigree-based BLUPs have been developed to leverage information from related individuals. This is used to estimate the genetic potential that a parent can pass to its offspring and is termed “breeding value” [24]. Genomic BLUPs (GBLUPs) are an extension of pedigree-based BLUPS, using DNA marker information instead of pedigree information to construct a realized relationship matrix between individuals in a population. The realized relationship matrix can more accurately estimate relatedness, particularly among full siblings, than the pedigree-based relationship matrix [25–27]. The resultant breeding values are expected to more closely mirror the true genetic potentials of individuals [28–30].

Breeding values derived from BLUPs have been used to successfully identify superior individuals in several rosaceous crops including apple, peach, raspberry, and strawberry [31–37]. Extensive work has been done in apple to estimate the breeding values from unreplicated trials [31,33,38,39]. Breeders have observed enhanced genetic gain using both pedigree-based and genome estimated breeding values in other perennial tree crops, including citrus, rubber and *Eucalyptus* [40–43]. Sweet cherry shares

104 many of the breeding scheme challenges of apple and other perennial tree crops: unbalanced trials and a
105 long juvenility period. Hence, the same methodologies can be utilized.

106 Additive effects are considered to be the largest component of genetic variance that is passed to
107 progeny [44]. While many genome-wide approaches including GBLUPs have been employed to estimate
108 breeding values across crops, these methods are almost solely focused on estimating additive effects
109 alone as a proxy for total genetic effects. Few studies have examined non-additive genetic variance
110 components in rosaceous crops [45]. Kumar et al. [45] reported on a comprehensive study estimating
111 sources of genetic variance for 32 traits in apple across 17 families and two locations using GBLUPs.

112 In cherry, there are few published accounts that utilize BLUPs or other genome-wide DNA-
113 enabled approaches for estimating the genotypic value of individuals. The only published genome-wide
114 study in sweet cherry estimated breeding values for cherry fruit size in U.S.-relevant germplasm from
115 large-effect QTLs in a Bayesian analysis, but it did not include genetic background effects [18]. There is
116 no published information on the genome-wide additive and non-additive variance components and
117 prediction of the genetic value of individuals for any sweet cherry trait.

118 This study addresses a deficiency of published information on genetic parameters for sweet
119 cherry breeding-relevant traits beyond those influenced primarily by large-effect QTLs by obtaining robust
120 estimates of genetic variance components. To ensure wide applicability of the study for cherry, we used a
121 large set of sweet cherry breeding germplasm. These data were gathered from germplasm in public
122 sweet cherry breeding programs as part of RosBREED project [46]. Our objectives were to: (1) estimate
123 variance components across a broad spectrum of traits in sweet cherry germplasm important to North
124 American breeders and producers, and (2) assess the predictive accuracy of obtained genome-estimated
125 breeding values (GEBVs) for a subset of the most valuable traits. Previous studies show that the
126 genome-estimated breeding values of individuals that are robust across years and families can increase
127 the pace and efficiency of breeding. Specifically, valuable cherry parents can be identified more quickly
128 and with greater confidence than those obtained through phenotypic data alone.
129
130
131

METHODS

Germplasm

We used all individuals from the RosBREED sweet cherry Crop Reference Set with genome-wide SNP data, totaling 505 individuals (Additional File 1). This set consisted of cultivars (n = 42), wild accessions (n = 3), unreleased selections (n = 24), and unselected offspring (n = 436) from 66 families. The unselected offspring category includes 77 F1 offspring derived from a wild parent and 359 F1 offspring derived from existing cultivars. Trees were grown at two sites in Washington State (U.S.A.) located approximately 0.5 km apart: the Irrigated Agriculture Research and Extension Center of Washington State University Roza Unit, (46 □29'N and 119 □73'W) and at Pear Acres (46 □29'N and 119 □75'W). Each tree was planted in 2006, 2007, or 2008 and managed using conventional orchard management practices. Unselected offspring were grown on their own roots, and the remaining germplasm were grown on Gisela 6 rootstock [47]. A single tree was used for each individual. The Crop Reference Set was established to represent North American sweet cherry breeding germplasm for QTL identification and validation and other quantitative genetics endeavors [48].

Phenotypic data

This study used the sweet cherry phenotypic data set previously described in Chavoshi et al. [49] obtained in the RosBREED project. This data set consisted of 32 traits evaluated in 2010, 2011, and 2012. Standardized phenotyping protocols for sweet cherry [49] were used. For individual fruit traits, the five largest fruit without blemish were measured and averaged. In the case of pitting and cracking, the proportion of fruit observed with symptoms out of 25 fruit was recorded. Bulked fruit traits (bulked fruit weight, bulked firmness, bulked SSC, and bulked TA) were reported as the average of measurements over 25 fruit.

Nine traits of the 32 were focused on here because of their importance in new sweet cherry cultivars: time to bloom, time to maturity, pedicel-fruit retention force (PFRF), fruit dimensions, fruit weight, firmness, SSC, TA, and powdery mildew incidence. Time to bloom and time to maturity were measured both in Julian calendar days starting from January 1st of the calendar year and in growing degree days (GDD). The force required to pull a ripe cherry fruit from its pedicel, PFRF, and fruit weight were both measured in grams. Firmness, SSC, and TA were measured in units of g/mm, Brix°, and percentage,

160 respectively. Foliar powdery mildew incidence was scored in August of each year, immediately after the
161 fruiting season, on a 0-5 scale, where 0 is no infection and 5 is highly infected leaves. These nine traits
162 are referred to as “focus traits” for the rest of the study. All trait data were measured over three years
163 except for powdery mildew incidence, which was not assayed in 2010. Results from the other traits are
164 given in the supplementary material, but not discussed.
165 Several transformations of the trait data were performed for the focus traits. “Fruit dimensions”
166 was determined newly here as the first component from a principal component analysis between fruit
167 length and fruit width, which are both end-to-end fruit measurements in millimeters. The first principal
168 component summarized 95.4% of total phenotypic variation for fruit length and width. Growing degree
169 days was calculated for an alternative measure of phenological traits. Climatic data was obtained from
170 Washington State University’s AgWeatherNet using the “Roza” station [50], using a base temperature of
171 4.5 °C and maximum of 30 °C. Daily maximum temperatures above 30 °C were reduced to 30 °C, and
172 negative temperatures were set to zero, following McMaster and Wilhelm [51]. Erroneous data points,
173 defined as those larger than twice the next largest value or less than one-half of the next smallest value
174 and having a studentized residual with an absolute value greater than 5, were removed. Such data were
175 assumed to be data entry errors. There were 97 individuals with no phenotypic data: 13 selections and 84
176 unselected progeny. These individuals were used in the model-building and prediction steps for all
177 models except for cross validation.

178 *SNP data*

179 The SNP data were obtained from the RosBREED project using the RosBREED cherry 6K SNP array v1
180 (an Illumina Infinium® II array) [52]. The SNP curation pipeline is described in Cai et al [53]. Missing data
181 were imputed with Beagle as implemented in SynBreed [54,55] using the hidden Markov model and a
182 minor allele frequency of 0.05. Individuals or SNPs missing more than 25% data were removed from
183 analysis and the SNP. In total, a genome-wide set of 1615 SNPs was used.

184 *Statistical modeling*

185 Variance components were estimated with R-ASReml 3.0 [56], and additional statistical analyses were
186 conducted in R v3.4 [57]. The following model was used for initial estimates of genetic effects for a single
187 trait, $\mathbf{Y}$:

188 $$\mathbf{Y} = \mathbf{Xb} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{d} + \mathbf{Z}_3\mathbf{i} + \mathbf{Z}_4\mathbf{a}_\mathbf{Y} + \mathbf{Z}_5\mathbf{d}_\mathbf{Y} + \mathbf{Z}_6\mathbf{i}_\mathbf{Y} + \mathbf{e}$$

189 where $\mathbf{a}, \mathbf{d}, \mathbf{i}, \mathbf{a}_\mathbf{Y}, \mathbf{d}_\mathbf{Y}$ and $\mathbf{i}_\mathbf{Y}$ are the random variables for additive effects, dominance effects, effects

190 from additive-by-additive epistatic, additive-by-year effects, dominance-by-year effects, and epistasis-by-

191 year effects, respectively. Variables $\mathbf{Z}_{1\text{-}3}$ and $\mathbf{Z}_{4\text{-}6}$ are design matrices for main effects and interaction

192 terms, respectively. Dimensions of $\mathbf{Z}_{1\text{-}3}$ are $nY \times Y$ and $\mathbf{Z}_{4\text{-}6}$ are $nY \times nY$, where *n* is the number of

193 individuals and *Y* is the number of years with trait data for an individual. Year was treated as a fixed

194 effect, where **X** is the design matrix relating observations to years and **b** is a vector of fixed effects due to

195 year. In a preliminary analysis, the effect of location was evaluated as a fixed effect using a Wald test.

196 Location did not have a significant effect on the focus traits (p-value > 0.10) and was omitted from the

197 model. Random variables were assumed to follow a normal distribution:

198 $$\mathbf{a} \sim N(0, \mathbf{G}_\mathbf{a}\sigma_a^2), \mathbf{d} \sim N(0, \mathbf{D}\sigma_d^2), \mathbf{i} \sim N(0, \mathbf{G}_{\mathbf{aa}}\sigma_{aa}^2),$$
$$\mathbf{a}_\mathbf{Y} \sim N(0, \mathbf{I}_\mathbf{Y} \otimes \mathbf{G}_\mathbf{a}\sigma_{aY}^2), \mathbf{d}_\mathbf{Y} \sim N(0, \mathbf{I}_\mathbf{Y} \otimes \mathbf{D}\sigma_{dY}^2), \mathbf{i}_\mathbf{Y} \sim N(0, \mathbf{I}_\mathbf{Y} \otimes \mathbf{G}_{\mathbf{aa}}\sigma_{aaY}^2)$$
$$\mathbf{e} \sim N(0, \mathbf{R})$$

199 The covariance structure for year was modeled as a repeated measure: $\mathbf{R} = \mathbf{I}_{Individual} \otimes \mathbf{e}_\mathbf{Y}$ where

200 $\mathbf{I}_{Individual}$ is an identity matrix of individuals included in the study and $\mathbf{e}_\mathbf{Y}$ is a 3×3 matrix of year error

201 terms using a general correlation structure implemented in ASReml. The genomic additive relationship

202 matrix was computed with R/rrBLUP [58] using the VanRaden method [59]:

203 $$\mathbf{G}_\mathbf{a} = \frac{\mathbf{HH'}}{2\sum p_i(1-p_i)}$$

204 where p_i is frequency of the positive allele for a single marker column, and **H** was computed as equal to

205 centered marker data, $\{H\}_{ij} = \{M\}_{ij} - 2(p_i - 0.5)$. **M** is an *n* x *m* marker matrix with *n* individuals and *m*

206 markers expressed as (-1,0,1) frequency. The dominance relationship matrix was computed using

207 normalized matrices described by Su et al. [60] and implemented using a custom R program [61]:

208 $$\mathbf{D} = \frac{\mathbf{ZZ'}}{\sum_i 2p_i(1-p_i)(1-2p_i(1-p_i))}$$

where the **Z** matrix is a transformation of the marker matrix, **M**:

$$\{Z\}_{ij} = \begin{cases} -2p_i(1-p_i) & \text{if } m_{ij} = -1 \\ 1-2p_i(1-p_i) & \text{if } m_{ij} = 0 \\ -2p_i(1-p_i) & \text{if } m_{ij} = 1 \end{cases}$$

The epistatic relationship matrix for additive by additive effects was computed by taking the Hadamard product between $\mathbf{G_a}$, the additive genomic relationship matrix, and itself: $\mathbf{G_{aa}} = \mathbf{G_a} \circ \mathbf{G_a}$. When a relationship matrix was not positive definite, a small constant of $1e^{-6}$ was added to the first eigenvector, and the matrix was inverted.

The full model included additive, dominance, and epistatic main effects and their interactions with year and is also called the "ADI model" in this paper. Model fit was assessed by checking for model convergence, examining studentized residuals for each trait by year combination, and examining the extended hat matrix for influential observations. The default model convergence criteria for ASReml were used, in which the final iteration must satisfy the following conditions: a change log likelihood less than 0.002 * previous log likelihood, and the variance parameters estimates change less than 1% from the previous iteration. The extended hat matrix for linear mixed models is:

$$\mathbf{WC^{-1}W'}$$

$$\text{Where } \mathbf{C} = \mathbf{W'R^{-1}W} + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G^{-1}} \end{pmatrix} \text{ and } \mathbf{W} = \begin{bmatrix} \mathbf{X} & \mathbf{Z} \end{bmatrix}$$

Influential data points were those with a value greater than 2 times the average value of the diagonal of the hat matrix excluding zeros.

The statistical significance of main effects and interactions were tested by first generating reduced models and then performing log-likelihood ratio tests between full and reduced models. To account for positively-bound variance component estimates, a mixture of Chi-square distributions as implemented in the R package asremlPlus [62] was used. Non-significant values from the log likelihood ratio tests were interpreted as the reduced models being as effective as the full model in modeling the response variable. Heritability numerators were estimated as σ^2_a for narrow-sense heritability (h^2) and as $\sigma^2_a + \sigma^2_d + \sigma^2_{aa}$ for broad-sense heritability (H^2); both were divided by the sum of the variance components

for final heritability estimates. Genetic values were computed as the sum of main effects for **a, d** and **i** for an individual, following the methodology of Kumar et al. [45]. Genotype-by-year effects are the sum of $\mathbf{a}_y$, $\mathbf{d}_y$, and $\mathbf{i}_y$ when all years were used in the estimation.

Model validation

Five-fold cross validation was used where the data set was randomly divided into 5 equal-sized parts ("folds"), a single fold (20% of the individuals) was removed across all years, and the remaining observations were used for variance component estimation and prediction of genetic values. The resultant model was used to predict genetic values of those removed individuals. This process was repeated for all 5 folds. Observations lacking phenotypic information for a specific year and trait were excluded from the model-building and validation. Because predictions can be affected by sampling variance, 5-fold cross validation was repeated 25 times using different randomly generated folds for each iteration. In addition, cross validation was performed, omitting each of the 66 full-sib families or a year as validation populations. These latter situations were intended to reflect the situation of predicting genetic performance for previously unphenotyped individuals that are related to the training population, and for predicting performance for an unobserved year. Prediction accuracy was assessed by computing correlation coefficients between predicted genetic values and observed data adjusted for fixed effects.

Other statistics

The statistical significance of year on the models was checked with the Wald test. Genetic-by-year effects were further explored by estimating genetic values and genetic variance components using a single year of data. Spearman's rank-order correlations were conducted to evaluate changes in rank of genetic values of individuals across years. Pairwise Pearson (r) and Spearman (ρ) correlations between traits were assessed for the multi-year ADI model. Principal component analyses were conducted on correlation matrix of genetic values calculated from (1) all individuals used in this study, and (2) only the cultivars and ancestors (n=48), using 8 independent traits: bloom time, harvest time, pedicel-fruit retention force, fruit weight, firmness, SSC, TA, and powdery mildew incidence. The first and second principal components were graphed on a biplot [63], where the rotations for plotting the variables were scaled by the first eigenvalue.

RESULTS

Distribution of phenotypic data

All trait distributions (consisting of 600-755 data points for each trait) were influenced by the year of data collection (Fig. 1). Wald test results for year were consistently highly significant for all focus traits across all models ($p < 0.001$ in all cases).

The 2010 data visually differed most from the other years, particularly for bloom date, fruit dimensions, fruit weight, firmness, and SSC. Data in 2010 were also the most sparse compared to data from other years (Additional File 2). Fruit dimensions and fruit weight had similar distributions across years. Although the distributions of bloom date and bloom time seemed to differ, the accumulation of GDD remained relatively stable over the three years. However, GDD accumulation was higher in early 2010 than other years during the critical period of flower bloom (data not shown).

Statistical assumptions and model fit

All models for the focus traits converged. Inspection of the residual plots and quantile-quantile plots signal that the error terms were independently and identically distributed (results not shown). The extended hat matrix revealed no influential data points for any of the models. Appropriate residual patterns were observed for all models and traits (results not shown), demonstrating no major departures from the assumption of homoscedasticity. Moderate correlations were observed between the additive, dominance and epistatic effects within a trait for the full model ($r = 0.3 – 0.7$). Population structure was observed among the individuals. In a principal component analysis of the correlation matrix of the SNP data, the first two components summarized 14% of the variation. There was distinct grouping of the wild accessions and offspring derived from those wild accessions along the second principal component (data not shown). Visual inspection of the diagonals and off-diagonals from the realized relationship implies a single Gaussian distribution of the matrix elements. Thus, the population structure had minimal impact on the genomic additive relationship matrix (Additional File 3).

Log likelihood ratio tests comparing reduced models with the full ADI model demonstrated that the full model was not necessary to describe trait variance for any focus trait (Table 1). The main effects-only model that included only additive, dominance, and epistatic effects was significantly different from

289 the full model (p-values <0.05) for all focus traits, except for powdery mildew incidence and SSC, which
290 had notable p-values defined as less than 0.10. Reduced models consisting of single main effects
291 (additive, dominance or epistatic) or single main effects plus their year interaction term (e.g., additive and
292 additive-by-year) were highly significant for all traits. This demonstrates that the reduced models did not
293 adequately capture variation compared to the full model. For most focus traits, genetic models that
294 included additive, epistatic, additive-by-year and epistasis-by-year effects were not statistically different
295 from the full model. Thus, dominance and dominance-by-year could be dropped from their genetic models
296 without significant loss of information. Traits that were exceptions to the above were fruit weight, fruit
297 dimensions, and bloom date, for which optimal fit was obtained by including dominance in the model. For
298 all traits, dominance-by-year and epistasis-by-year effects could be removed from the model without
299 much loss of information. Additive-by-year effects had a statistically significant effect on bloom date,
300 bloom time, and PFRF (p<0.01).

Table 1: Log-Likelihood ratio test statistics for reduced models.

Model	df	Bloom Date	Bloom Time	Harvest Date	Harvest Time	PFRF	Fruit Dimensions	Fruit Weight	Firmness	SSC	TA	Powdery Mildew
a, **d**, **i**, **a_Y**, **d_Y**, i_Y	1	0.08	0.46	3.43*	0.28	0.09	2.48‡	0.11	0.07	0.00	4.82*	0.62
a, **d**, **i**, **a_Y**, d_Y, **i_Y**	1	4.92*	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.70
a, **d**, **i**, a_Y, **d_Y**, **i_Y**	1	9.80***	7.80**	0.58	3.36*	6.57**	1.23	1.24	4.80*	2.94*	0.94	0.35
a, **d**, i, **a_Y**, **d_Y**, i_Y	2	8.02**	16.18***	66.80***	64.86***	6.48*	6.78*	6.12*	65.72***	16.35***	27.20***	24.11***
a, d, **i**, **a_Y**, d_Y, **i_Y**	2	8.15**	2.18	3.67‡	3.27‡	3.01‡	27.26***	28.59**	0.10	0.00	0.00	0.70
a, **d**, **i**, a_Y, **d_Y**, **i_Y**	2	38.47***	27.03***	21.42***	26.17***	18.28***	36.39***	23.16***	20.15***	17.33***	16.92***	13.67***
a, d, i, **a_Y**, d_Y, i_Y	4	44.16***	38.83***	151.67***	148.47***	28.3***	60.34***	88.48***	93.05***	17.86***	32.11***	63.25***
a, **d**, i, a_Y, **d_Y**, i_Y	4	69.96***	70.76***	114.46***	112.39***	33.86***	66.94***	40.09***	109.78***	45.02***	63.2***	43.75***
a, d, **i**, a_Y, d_Y, **i_Y**	4	65.60***	41.18***	66.61***	73.36***	45.52***	143.04***	132.90***	47.06***	28.43***	21.12***	33.52***
a, **d**, **i**, a_Y, d_Y, i_Y	3	55.39***	27.71***	9.86**	7.55*	20.12***	14.62***	6.66*	25.72***	5.10‡	12.73**	4.93‡
a, d, i, a_Y, d_Y, i_Y	5	76.22***	53.20***	151.67***	148.47***	41.51***	63.50***	88.83***	101.90***	20.63***	33.93***	63.25***
a, **d**, i, a_Y, d_Y, i_Y	5	105.34***	75.84***	114.46***	112.39***	39.95***	71.64***	42.22***	116.05***	45.32***	63.59***	44.52***
a, d, **i**, a_Y, d_Y, i_Y	5	106.64***	65.35***	79.69***	80.30***	62.65***	159.6***	139.02***	62.46***	32.09***	35.70***	38.6***

Log-likelihoods are expressed relative to the full model (**a**, **d**, **i**, **a_Y**, **d_Y**, **i_Y**). Statistical significance is labeled as ‡ = p < 0.10, * = p < 0.05, ** = p < 0.01, *** = p < 0.001, marking if the reduced model is statistically different from the full model using the chi-square distribution (df = degrees of freedom). The terms in the models, **a**, **d**, **i** refer to effects from additive, dominance, and epistatic sources, respectively. The terms **a_Y**, **d_Y**, **i_Y** refer to additive-by-year, dominance-by-year, and epistasis-by-year effects, respectively. The bolded terms in the column "Model" indicate components included in the reduced model, while grey terms have been excluded.

Genetic variance and predictive ability of full model

Variance component estimates from the full model indicated moderate to high broad-sense heritabilities across the focus traits, ranging from 0.47 for pedicel-fruit retention force to 0.83 for harvest date (Table 2). Narrow-sense heritabilities ranged from 0.20 for PFRF to 0.37 for fruit dimensions. Epistasis was the single largest genetic variance component for most traits: bloom time (28%), harvest date (48%), harvest time (48%), firmness (49%), SSC (27%), TA (33%), and powdery mildew incidence (42%). Additive variance was the largest component for bloom date (37%), PFRF (20%), and fruit dimensions (37%). Dominance was the largest variance component only for fruit weight (34%); in contrast, dominance represented less than 1% of trait variance for firmness, SSC, TA, and powdery mildew incidence. Genotype-by-year effects were less than 10% for all traits except bloom date ($\mathbf{a_Y}$ = 11%) and TA ($\mathbf{i_Y}$ = 14%). Residual variance of most traits was less than 25% of phenotypic variance, except for PFRF (45%) and SSC (48%). Variances and standard errors for all components and traits, and variance percentages, are provided in Additional Files 4 and 5, respectively.

Table 2: Variance components (%), narrow-sense heritability (h^2), broad-sense heritability (H^2), the coefficient of correlation (r), the coefficient of correlation after cross validation (r_{CV}), and the total number of observations for model building (N).

	Bloom Date	Bloom Time	Harvest Date	Harvest Time	PFRF	Fruit Dimensions	Fruit Weight	Firmness	SSC	TA	Powdery Mildew
	variance component (%)										
additive (A)	33.20	25.45	27.39	27.87	19.83	37.40	30.76	27.49	21.59	27.19	28.31
dominance (D)	10.80	11.48	7.73	6.68	11.10	26.80	33.61	0.42	0.00	0.00	0.00
epistasis (I)	17.47	27.84	47.66	47.90	15.62	8.36	12.08	48.96	26.76	32.81	41.52
A x Year	11.16	8.10	1.18	2.89	6.30	2.98	2.26	4.57	4.08	3.42	1.57
D x Year	4.23	0.00	0.00	0.00	0.00	0.00	0.52	0.00	0.00	0.00	1.19
I x Year	2.06	3.63	5.71	1.65	2.31	8.03	1.14	0.94	0.00	14.15	4.33
error	21.08	23.48	10.33	13.02	44.84	16.42	19.62	17.62	47.56	22.43	23.07
	trait heritability and genome-estimated breeding values accuracy										
h^2	0.33	0.25	0.27	0.28	0.20	0.37	0.31	0.27	0.22	0.27	0.28
H^2	0.61	0.65	0.83	0.82	0.47	0.73	0.76	0.77	0.48	0.60	0.70
r	0.88	0.90	0.97	0.97	0.83	0.94	0.95	0.94	0.82	0.88	0.93
r_{CV}, 5-fold	0.56	0.48	0.78	0.79	0.59	0.78	0.77	0.69	0.46	0.42	0.68
r_{CV}, -year	0.58	0.48	0.88	0.88	0.58	0.82	0.83	0.76	0.47	0.50	0.74
r_{CV}, -family	0.55	0.46	0.74	0.74	0.55	0.76	0.70	0.66	0.38	0.31	0.58
N	644	644	665	665	759	774	764	763	768	577	604

Correlations between adjusted phenotypic data and genetic values from the ADI model were high, 0.82-0.97 for all focus traits (Table 2). Coefficients of correlation under cross validation were very similar for 5-fold cross validation and when a year was left out. Correlations for cross validation that omitted full-sib families were the lowest among the three cross validation scenarios. Across all cross-validation scenarios, those traits with the highest broad-sense heritabilities, fruit dimensions, fruit weight, firmness, harvest date, and harvest time, had the most consistently high prediction accuracies ($r > 0.65$). The lowest prediction accuracies were observed for SSC and TA, which never exceeded 0.50.

Heritability and predictive ability of reduced models

Broad-sense heritability was largely unchanged across the reduced models (ADI to AI and AD) for all focus traits (Fig. 2). Narrow-sense heritability gradually increased with decreasing model complexity for all focus traits, from the full model to the AD model and from the AD to the A model. Narrow-sense heritability was highly similar in the AI and ADI models for all traits except for fruit dimensions and fruit weight, in which the AI h^2 was noticeably higher in the AI model compared to the ADI and AD models (Fig. 2). In the additive effects-only model (A), H^2 was similar in value to the h^2 of the other models.

Predictive power, as measured by r^2, was consistent between the ADI model and the AI model for all traits (Fig. 2). The predictive power decreased slightly for the AD model compared to the full model, and decreased slightly more for the A model compared to the AD model. The r^2 values under 5-fold cross validation varied little across genetic models for all traits, only decreasing slightly in the AD and A reduced models for harvest date, harvest time, and firmness. Spearman rank correlations between the full and reduced models indicated minimal changes in rankings of individuals when using the AD and AI models (r = 0.96–1.00) and small changes in the A model compared to ADI model (r = 0.91–0.96) for genetic values and breeding values (Additional File 2).

Single year analysis

Variance components estimated with a single year of data varied substantially across years for all focus traits (Fig. 3). For all traits except harvest date and harvest time, the percentages of additive variance differed by 10% or more across years. Additive variance for harvest date and harvest time varied the least among the focus traits, 37 to 44% and 37 to 47%, respectively. Dominance variance components for SSC and TA were close to zero (<0.0001%) across all years, while at the other extreme, dominance variation

for fruit dimensions was always greater than 20%. Epistatic variance consistently composed a large percentage of genetic variance for firmness (>32%) and powdery mildew incidence (>49%). Genotype-by-year effects were greatest for TA (18%), bloom date (18%), and bloom time (12%).

Rankings of individuals by genetic values estimated from each a single year of data significantly differed from the multi-year genetic rankings in Spearman rank correlation tests ($p < 0.001$, Additional File 2). Rank correlations between the 2010-derived predictions and the multi-year predictions were lower than the subsequent years (2010: 0.35–0.63; 2011: 0.58–0.92; 2012: 0.85–0.97). However, correlations between single-year breeding values and phenotype implied a better fit for all years and traits than the single-year breeding values with their multi-year counterparts ($\rho = 0.64–1.00$) (Additional File 2).

Correlations among trait genetic values

The genetic values of the focus traits had weak to moderate positive correlations with each when considering only unreleased offspring and selection, with some exceptions (Table 3). Fruit weight and fruit dimensions, harvest date and harvest time, and bloom date and bloom time were all highly correlated pairs of traits ($r > 0.90$, Table 3). SSC was negatively correlated with all focus traits except TA. Titratable acidity was also negatively correlated with fruit dimensions, fruit weight and powdery mildew incidence. In a biplot of the correlation matrix of the named cultivars using eight independent traits, the first two principal components summarized 55% of the variance (Fig. 4). All variables but SSC and TA skewed to the left, corresponding to the negative correlations between SSC and all variables except TA. Wild ancestors and wild offspring were on the right side of the biplot corresponding to their high SSC, low powdery mildew incidence, and low fruit weight. Additional figure 6 further separates the sweet cherry founders and derived cultivars by fruit weight and SSC content.

Table 3: Pairwise trait correlations and covariances between genetic values for sweet cherry selections and unselected offspring.

	Bloom Date	Bloom Time	Harvest Date	Harvest Time	PFRF	Fruit Dimensions	Fruit Weight	Firmness	SSC	TA	Powdery Mildew
Bloom Date	**3.507**	0.897***	0.317***	0.314***	0.301***	0.136*	0.196***	0.223***	-0.101	0.133*	0.184**
Bloom Time	18.17	**117.1**	0.213***	0.208***	0.198***	0.087	0.130*	0.134‡	-0.071	0.064	0.218***
Harvest Date	3.446	13.38	**33.72**	0.998***	0.255***	0.340***	0.346***	0.547***	-0.364***	0.107	0.220***
Harvest Time	50.93	195.4	502.2	**7508**	0.251***	0.334***	0.341***	0.549***	-0.355***	0.106	0.225***
PFRF	74.88	283.9	196.2	2883	**17630**	0.566***	0.603***	0.465***	-0.161***	0.173‡	0.185***
Fruit Dimensions	0.3082	1.142	2.394	35.10	91.09	**1.468**	0.946***	0.511***	-0.507***	-0.210***	0.462***
Fruit Weight	0.9209	3.695	5.042	74.33	201.1	2.880	**6.311**	0.514***	-0.435***	-0.208***	0.505***
Firmness	17.26	59.91	131.1	1964	2546	25.55	53.26	**1702**	-0.392***	0.065	0.387***
SSC	-0.3403	-1.387	-3.793	-55.10	-38.44	-1.101	-1.958	-28.99	**3.214**	0.267***	-0.340***
TA	0.02486	0.06856	0.06204	0.9132	2.295	-0.02541	-0.05219	0.2668	0.04778	**0.009934**	-0.215**
Powdery Mildew	0.3509	2.409	1.304	19.91	25.11	0.5707	1.295	16.28	-0.6215	-0.02181	**1.041**

Correlations and covariances are given in the upper triangle and lower triangle, respectively, and trait variances are bolded on the diagonal.
Statistical significance is labeled as ‡ = p < 0.10, * = p < 0.05, ** = p < 0.01, *** = p < 0.001, signaling if the correlations are different from zero.

374

375 **DISCUSSION**

376 Results indicated high broad-sense heritability for all of the focus traits and also illuminated the
377 importance of non-additive variation in the sweet cherry traits studied. A poorly-fitting genetic prediction
378 model can mispresent the genetic variances of traits and the potential for genetic gain.

379 *Importance of model fit and consequences for predictive ability*

380 This study demonstrated that for most traits, non-additive sources of variation comprised an equal or
381 larger portion of the genetic variance than additive variance. A genetic model including additive, epistatic,
382 additive-by-year and epistasis-by-year effects was usually the most parsimonious approach for capturing
383 major sources of variation. Exceptions were fruit dimensions and fruit weight, which instead were best
384 described by a model with additive, dominance and additive-by-year effects, and harvest date, best
385 described by a main effects-only model.

386 Using an incorrect model to determine genome-wide breeding values can provide misleading
387 information for making breeding decisions. Table 4 illustrates the consequences of using a poorly-fitting
388 reduced model for estimating breeding values. Breeding values were often larger in relative magnitude in
389 the reduced models compared to the full model, which can exaggerate genetic gains possible in the
390 population. For example, days to maturity in an Ambrunes/Sweetheart cross would be overestimated by
391 twice as many days in the additive-only model compared to the ADI model. Likewise, crosses with the
392 wild accession MIM 23 were predicted to result in midparent values of fruit size twice as small in the A
393 model compared to the ADI model (Table 4). The inflation of additive variance when non-additive sources
394 are omitted has been documented in several other species including apple, loblolly pine, white spruce
395 cassava, cattle, pigs, Coho salmon, and rainbow trout [27,45,60,64–68].

396

Table 4: Breeding values and midparent values under different genetic models demonstrated with several individuals and traits.

Trait	Model	Parental values			Midparent values		
		Ambrunes	Sweetheart	MIM 23	Ambrunes/ Sweetheart	Ambrunes/ MIM 23	Sweetheart/ MIM 23
Harvest Date (-15.82, 16.35)	A	14.64	8.43	-9.93	11.53	2.36	-0.75
	AD	8.79	5.57	-6.34	7.18	1.23	-0.38
	AI	7.36	4.76	-6.60	6.06	0.38	-0.92
	ADI	6.48	4.05	-5.73	5.27	0.38	-0.84
Fruit Weight (-11.45, 5.06)	A	-1.86	1.64	-10.67	-0.11	-6.26	-4.51
	AD	-0.87	0.95	-4.58	0.04	-2.73	-1.82
	AI	-2.55	1.98	-8.80	-0.28	-5.67	-3.41
	ADI	-1.11	1.06	-4.72	-0.03	-2.91	-1.83
SSC (-3.77, 5.61)	A	-1.07	-1.98	3.38	-1.53	1.15	0.70
	AD	-0.84	-2.00	2.93	-1.42	1.05	0.47
	AI	0.10	-1.81	2.53	-0.86	1.31	0.36
	ADI	0.10	-1.83	2.48	-0.86	1.29	0.33
Powdery Mildew Incidence (-2.72, 1.99)	A	0.13	1.28	-2.38	0.70	-1.13	-0.55
	AD	0.39	0.83	-1.67	0.61	-0.64	-0.42
	AI	-0.31	0.89	-1.78	0.29	-1.05	-0.45
	ADI	-0.28	0.87	-1.74	0.29	-1.01	-0.44

Intervals given below each trait are the range of values in the additive-only model observed across all individuals. In the column “Model”, A, D, and I refer to additive, dominance, and epistatic effects, respectively, and their accompanying genotype-by-year interactions.

If genetic values are used to select individuals to be clonally propagated for further trialing or cultivar release, then the genetic model has a lower, perhaps negligible, influence on prediction of total genetic performance. Ceballos et al [69] argued that using total genetic values from additive and non-additive variance components provides greater potential for genetic gain under clonal selection. However, our results showed that the estimated broad-sense heritability and the genetic values of sweet cherry individuals are largely unchanged across the different genetic models. This demonstrates that there is effectively no change in genetic gain if a more complex model is used for identifying high-performing individuals (Figure 2, Additional File 2).

Including year as a main effect was warranted in this study, given the statistically significant effect of year on all traits. However, the effect of including genotype-by-year interactions varied by the trait and genetic variance component. Genotype-by-year interactions were generally of much smaller magnitude than the main genetics effects and largely absent for dominance effects (Table 1, Fig. 3). Nevertheless, year had a major effect on genetic effects estimates and was included as a fixed variable to obtain robust predictions across years. Year often has a statistically significant effect on the traits of sweet cherry and

417 other rosaceous crops, including sweet cherry pedicel-fruit retention force [70], apple fruit texture [71],
418 sugar content in peach and nectarine [72], and several phenological and fruit quality traits in strawberry
419 [73].

420 This study demonstrated the need for a training population to fully capture variation of the target
421 population in order to maximize prediction accuracy. The single year analysis showed that although a
422 model built using a single year of data could be used accurately to predict individuals evaluated in that
423 year, it could not be easily extrapolated to individuals whose genetic values lie outside the distribution of
424 the training data (Table 2, Additional File 2). The GBLUP approach relies on information from relatives to
425 improve the accuracy of the estimates [74]. Because there were often sparse observations for a single
426 year, sampling error biased the single-year estimates and resulted in models that fit the data within each
427 year, but not across years. These effects were likely exacerbated with wild accession, distantly related
428 cultivars and derivatives from both groups. However, the true pairwise genetic covariance between the
429 distantly related germplasm is estimated with less reliability with the realized relationship matrix than more
430 closely related germplasm [75].

431 *Genetic architecture of focus traits in sweet cherry*

432 This study confirmed the extensive opportunity in North American sweet cherry germplasm for genetic
433 improvement of the phenological traits of harvest timing and, to lesser extent, bloom timing. Previous QTL
434 studies for fruit maturity date across several *Prunus* species determined bloom timing and harvest timing
435 to be highly heritable with a large-effect QTL on LG4 [76]. Our findings also demonstrate the large broad-
436 sense heritability for these traits – reaching a ceiling of 0.83 for harvest time and 0.65 for bloom date (Fig.
437 2). There appears to be little advantage to using GDD to Julian days, since pairs of phenological traits for
438 bloom and harvest timing displayed highly similar genetic architecture and predictive accuracy. The data
439 were all gathered from a single location, in which GDD did not vary dramatically during the years of
440 evaluation. This may explain why GDD did not improve the model predictive ability over Julian days (Fig
441 2, Table 2). Bloom timing has become increasingly important as a trait relevant to productivity, since
442 variable climatic patterns in temperate regions can result in earlier flowering and an increased risk of floral
443 freeze damage [77]. Furthermore, since sweet cherries are a fresh market product that is subject to rapid
444 postharvest deterioration, it is crucial to for sweet cherry breeders and producers to understand the

445 expected time frame for fruit maturation [76]. These results may help sweet cherry breeders identify the
446 best parents in order to target a harvest timing window.

447 Moderate prospects were observed for genetic improvement of pedicel-fruit retention force (h^2 =
448 0.20, H^2 = 0.46, Table 2), where a low PFRF value is sought for mechanical harvest systems. Positive
449 correlations observed between PFRF and fruit dimensions, fruit weight, and firmness (Table 3) contrasted
450 with findings by Zhao et al. [70], in which PFRF was largely uncorrelated with firmness, fruit diameter, or
451 fruit length. However, that study was smaller in scope, using only 30 named cultivars and 26 unselected
452 F1 progeny.

453 The potential for genetic gain in fruit dimensions and fruit weight, two highly correlated
454 measurements of fruit size, was perhaps the highest among all focus traits due to large additive and
455 dominance effects (Table 2). These results are consistent with previous sweet cherry studies that showed
456 high correlations between fruit size measurements and high H^2 [18,78–80]. In those studies, six putative
457 QTLs influencing fruit size in cherry were identified and together accounted for 76–88% of the phenotypic
458 variance. Because fruit weight was highly correlated with fruit dimensions in the present study (Table 3,
459 Fig. 4) and can be evaluated rapidly, we considered it an effective proxy for fruit dimensions and general
460 fruit size.

461 The high broad-sense heritability for firmness (0.77) (Table 2) was consistent with estimates from
462 a study conducted on a biparental population in which H^2 was estimated at 0.78 to 0.85 [80]. In our study,
463 the moderate positive correlations (r = 0.51) between fruit firmness and fruit dimensions among the
464 unreleased progeny suggests genetic linkage among loci influencing these traits. This outcome was in
465 contrast to that of a multi-year QTL study, in which the Pearson correlations between fruit firmness and
466 fruit weight ranged from -0.64 to -0.67 for Regina × Lapins and -0.40 to -0.15 for Regina × Garnet F1
467 families [80]. Those correlations are likely due to unique linkage in Regina, Garnet and Lapins. The
468 correlations reported here may have also been influenced by the 77 progeny derived from the three wild
469 parents: MIM 17, MIM 23, and NY54. These individuals all had high SSC, small fruit size, and low fruit
470 firmness in their estimated genetic values relative to the population mean (Additional File 7).

471 Expectations for genetic improvement in SSC were moderately positive. Narrow-sense heritability
472 was estimated at 0.22, typical of the other focus traits in this study, where h^2 was most often between 0.2

and 0.3 (Table 2). Broad-sense heritability of SSC (H^2 = 0.48) was similar to that of other stone fruit: approximately 0.50 for apricot [81], 0.72 for peach [82], and 0.49 to 0.55 for apple [33]. Previous results confirm SSC had moderately negative correlation with fruit dimensions and fruit weight (-0.55 and -0.48, respectively). Our results are consistent with previous research, suggesting that SSC is directly related to photoassimilate partitioning and hence inversely correlated with fruit size [83,84]. Titratable acidity, the second most important contributor to fruit flavor after SSC, had similar variance component proportions and predictive accuracy to SSC. Major QTLs for TA have been detected on linkage groups 1, 5, and 6, explaining 99% of phenotypic variation in an F1 biparental peach population that was segregating for a large-effect locus [85]. These QTLs have not been reported in cherry. The broad-sense heritability of sweet cherry TA was lower in this study at H^2 = 0.60 and h^2 = 0.27. However, the population used in Dirlewanger et al [85] was created expressly to detect QTLs associated with TA, which might explain its very high H^2.

The large H^2 and h^2 estimated for foliar powdery mildew incidence indicated excellent potential for genetic improvement, but the lack of genome-wide dominance effects was surprising (Table 2). Powdery mildew resistance in U.S. sweet cherry germplasm was first traced to a single dominant allele in the ancestor PMR-1 [86,87]. There may be evidence for other sources of powdery mildew resistance among Pacific Northwest-adapted germplasm (Zhao et al, *In Prep*). Haploblock analysis might be required to detect dominance effects, which appeared to be absorbed by the other relationship matrices. The large epistatic component (42%) determined for this trait in sweet cherry was consistent with resistance to other plant diseases such as soybean to sudden death disease (causative agent *Fusarium virguliforme*) and rice to rice blast disease (*Pyricularia oryzae*) [88–90].

Implications for sweet cherry genetic improvement

The improvement in prediction accuracy when incorporating epistasis into the genetic model is consistent with studies on apple, Eucalyptus, wheat, cassava and maize [45,68,91–96]. Additive by additive epistasis is difficult to untangle from additive main effects due to selection, assortative mating and nongenetic covariances [44], all common facets of many breeding programs. The genomic relationship matrix for epistasis used here is considered to be an approximation since the assumption of random mating is not met [60,97]. The additive and dominance genomic relationship matrices used in this

study were not necessarily orthogonal due to linkage disequilibrium between SNPs [27], and the modest correlations between the additive dominance, and epistatic values were evidence of covariance between the different genetic effects.

Epistasis has not typically been targeted for parental selection in genetic improvement programs, although it can be captured indirectly with additive effects if epistatic alleles are fixed through inbreeding or drift [68,98]. Allele fixation is challenging in predominantly heterozygous crop such as sweet cherry whose high heterozygosity is maintained by a self-incompatibility mechanism [99]. However, knowledge of allele phasing, a feature of the RosBREED sweet cherry Crop Reference Set, could enable the capture of valuable epistatic interactions through known allelic interactions for both clonal performance and breeding parent utility.

Distributions of genome-estimated breeding values from the ADI model (Additional File 7) reveals a broad base of genetic diversity and opportunity for cherry improvement. This study confirmed that the cultivar Moreau has lowest breeding values for harvest date, denoting earliness. Early Burlat and several unreleased offspring mature several days after Moreau. The highest breeding values for harvest date, designating late-season maturation, included many unreleased offspring with higher breeding values than the highest-value cultivar (Ambrunes), particularly among the families Fam35 and Fam30 that might be useful as parents. There are also many unreleased offspring with desirable breeding values for certain traits. Families Fam1 and Fam21 have high breeding values for SSC and TA. Families Fam35 and Fam16have high fruit weight and firmness breeding values, in addition to the cultivars Cowiche, Sweetheart and Selah. The breeding values reported here will enable breeders to identify valuable parents earlier in the breeding program than through phenotyping alone. Identification of parents earlier in a breeding program is a major application of genomic selection [100] and has been widely used for many crops including long-lived perennial trees [40,45,67,101–103].

Using genomic selection to skip a breeding phase has also been proposed or implemented in several crops including apple, loblolly pine, Eucalyptus and several self-pollinated and hybrid crops [29,102,104–107]. The genetic values among unreleased progeny and selections described here revealed several promising individuals with commercial potential (Additional File 7, results not shown for selections). Because sweet cherry maintains the same genetic composition and genetic potential through

529 the breeding phases, genetic values obtained early in the breeding process will not change due to
530 recombination. Knowing the genetic potential of an individual will help cherry breeders discard low-
531 performing individuals and advance selections to the next phase with strong evidence. Knowledge of the
532 genetic potential of a candidate selection may enable breeders to skip a cycle of field evaluation, thus
533 increasing the pace of cultivar release and saving resources that can be diverted elsewhere. Given the
534 lengthy time period for developing a sweet cherry cultivar, shortening this process may represent
535 considerable savings.

536 **CONCLUSIONS**

537 The genetic values and the improved understanding of the genetic architecture of important traits
538 in sweet cherry obtained from this multi-year data set of a large pedigree-connected population represent
539 a clear opportunity for genetic improvement. This application – estimating genetic variance components
540 and genome-enabled genetic values – extended the original purpose of the RosBREED sweet cherry
541 Crop Reference Set: QTL detection and validation. We plan to update the genetic models by
542 incorporating new phenotypic data on the existing germplasm, adding new individuals and expanding the
543 genome-wide SNP set for denser genome coverage. Further research is needed to validate the accuracy
544 of the genetic predictions on an independent data set and to understand the extent of genotype-by-
545 environment effects for the obtained breeding values and genetic values.

546 **DECLARATIONS**

547 *Ethics approval and consent to participate*

548 Not applicable.

549 *Consent for publication*

550 Not applicable.

551 *Availability of data and material*

552 All data used in the analyses are available at the Genome Database for Rosaceae at www.rosaceae.org
553 [108] (persistent web link available upon article acceptance).

554 *Competing interests*

555 The authors declare no competing interests.

556 *Funding*

RosBREED: Combining disease resistance with horticultural quality in new rosaceous cultivars, provided by the USDA-NIFA Specialty Crop Research Initiative (2014-51181-22378)

Authors' contributions

Julia Piaskowski was responsible for the bulk of the data analysis and interpretation. Craig Hardner advised on model building interpretation. Lichun Cai prepared the SNP data and Yungyang Zhao gathered phenotypic data. Amy Iezzoni and Cameron Peace contributed to study design and interpretation. All authors contributed to manuscript preparation.

Acknowledgements

We gratefully acknowledge our funding source, RosBREED: Combining disease resistance with horticultural quality in new rosaceous cultivars, provided by the USDA-NIFA Specialty Crop Research Initiative (2014-51181-22378). We also gratefully acknowledge the contributions of an anonymous reviewer via the Peerage of Science review process, and Laura Giradeau for manuscript editing.

ADDITIONAL FILES

Seven supplementary materials accompany this manuscript: Additional File 1 includes germplasm information on all individuals included in the study; Additional File 2 lists Spearman rank correlations by year and model and the number of observations for each year; Additional File 3 is the additive relationship matrix diagonals and off-diagonals, labeled by relationship to wild germplasm; Additional File 4 lists all variances and standard errors for all traits whose models converged; Additional File 5 lists percent variance for all components and heritability for all traits whose models converged; Additional File 6 is a biplot of the cultivars and ancestors; Additional File 7 lists breeding values, dominance values and genetic values for all individuals except selections and all traits whose models converged. An online app for exploring the breeding and genetic values presented in Additional File 7 is available at <https://jpiaskowski.shinyapps.io/cherry_gebv_xplorr/>.

REFERENCES

1. FAOSTAT Data [Internet]. FAOSTAT. [cited 2017 Sep 10]. Available from: http://www.fao.org/faostat/en/#data

2. National Statistics for Cherry [Internet]. NASS. [cited 2017 Sep 10]. Available from: https://quickstats.nass.usda.gov/results/A8988197-374E-3950-BA97-9CBECA511544?pivot=short_desc

3. Sweet Cherry Production Up 36 Percent [Internet]. NASS. [cited 2017 Sep 10]. Available from: https://www.nass.usda.gov/Statistics_by_State/Washington/Publications/Fruit/2017/CH06.pdf

4. Yue C, Gallardo RK, Luby JJ, Rihn AL, McFerson JR, McCracken V, et al. An evaluation of U.S. tart and sweet cherry producers trait prioritization: evidence from audience surveys. HortScience. 2014;49:931–7.

5. Zheng X, Yue C, Gallardo K, McCracken V, Luby J, McFerson J. What attributes are consumers looking for in sweet cherries? Evidence from choice experiments. Agric. Resour. Econ. Rev. 2016;45:124–42.

6. Miller D, Casavant K, Buteau J. An analysis of Japanese consumer preferences for Pacific Northwest and Japanese sweet cherries. 1986. Report No.: XB0974.

7. Crisosto CH, Crisosto GM, Metheney P. Consumer acceptance of 'Brooks' and 'Bing' cherries is mainly dependent on fruit SSC and visual skin color. Postharvest Biol. Technol. 2003;28:159–67.

8. Hu Y. Sensory influences on consumers' willingness to pay: the apple and cherry markets [Doctoral Dissertation]. [Pullman, WA]: Washington State University; 2007.

9. Gallardo RK, Li H, McCracken V, Yue C, Luby J, McFerson JR. Market intermediaries' willingness to pay for apple, peach, cherry, and strawberry quality attributes. Agribusiness. 2015;31:259–80.

10. Gallardo RK, Li H, Yue C, Luby J, McFerson JR, McCracken V. Market intermediaries' ratings of importance for Rosaceous fruits' quality attributes. Int. Food Agribus. Manag. Rev. 2015;18:121–54.

11. Sweet Cherries Grades and Standards [Internet]. USDA Agric. Mark. Serv. [cited 2017 Feb 1]. Available from: https://www.ams.usda.gov/grades-standards/sweet-cherries-grades-and-standards

12. Olmstead JW, Ophardt DR, Lang GA. Sweet cherry breeding at Washington State University. Acta Hortic. 2000;103–10.

13. Oraguzie NC, Watkins CS, Chavoshi MS, Peace C. Emergence of the Pacific Northwest sweet cherry breeding program. Acta Hortic. 2017;73–8.

14. Haldar S, Haendiges S, Edge-Garza D, Oraguzie N, Olmstead J, Peace C. Applying genetic markers for self-compatibility in the WSU sweet cherry breeding program. ISHS Acta Hortic. [Internet]. 2009;859. Available from: http://www.actahort.org/books/859/859_45.htm

15. Sandefur P, Oraguzie N, Peace C. A DNA test for routine prediction in breeding of sweet cherry fruit color, Pav-R f -SSR. New Strateg. Plant Improv. 2016;36:1–11.

16. Quero-García J, Campoy JA, Castède S, Pitiot C, Barreneche T, Lerigoleur-Balsemin E, et al. Breeding sweet cherries at INRA-Bordeaux: from conventional techniques to marker-assisted selection. Acta Hortic. 2017;1–14.

618 17. Iezzoni A. Variance components and sampling procedures for fruit size and quality in sour cherry.
619 HortScience. 1986;21:1040–2.

620 18. Rosyara U, Bink MAM, van de Weg E, Zhang G, Wang D, Sebolt A, et al. Fruit size QTL identification
621 and the prediction of parental QTL genotypes and breeding values in multiple pedigreed populations of
622 sweet cherry. Mol. Breed. 2013;32:875–87.

623 19. Srivastava K, Verma M, Ahmad N, Ravi S, Ahmad S. Genetic variability and divergence analysis in
624 sweet cherry (*Prunus avium* L.). Indian J. Hortic. 2014;71:156–61.

625 20. Besford RT, Hand P, Peppitt SD, Richardson CM, Thomas B. Phase Change in *Prunus avium*:
626 differences between juvenile and mature shoots Identified by 2-dimensional protein separation and *in*
627 *vitro* translation of mRNA. J. Plant Physiol. 1996;147:534–8.

628 21. Henderson CR. Sire evaluation and genetic trends. J. Anim. Sci. 1973;1973:10–41.

629 22. Lynch M, Walsh B. Genetic and Analysis of Quantitative Traits. Sunderland, MA: Sinauer Associates,
630 Inc; 1998.

631 23. E Silva FF, Viana JMS, Faria VR, de Resende MDV. Bayesian inference of mixed models in
632 quantitative genetics of crop species. 2013;126:1749–61.

633 24. Henderson CR. Use of relationships among sires to increase accuracy of sire evaluation. J. Dairy Sci.
634 1975;58:1731–8.

635 25. Hayes B, Visscher P, Goddard M. Increased accuracy of artificial selection by using the realized
636 relationship matrix. Genet. Res. [Internet]. 2009;91. Available from:
637 http://dx.doi.org/10.1017/S0016672308009981

638 26. Vitezica ZG, Varona L, Legarra A. On the additive and dominant variance and covariance of
639 individuals within the genomic selection scope. Genetics. 2013;195:1223–30.

640 27. Muñoz PR, Resende MFR, Gezan SA, Resende MDV, de los Campos G, Kirst M, et al. Unraveling
641 additive from non-additive effects using genomic relationship matrices. Genetics [Internet]. 2014;
642 Available from: http://www.genetics.org/content/early/2014/10/15/genetics.114.171322.abstract

643 28. Habier D, Fernando RL, Dekkers JCM. The impact of genetic relationship Information on genome-
644 assisted breeding values. Genetics. 2007;177:2389–97.

645 29. Vela-Avitua S, Meuwissen THE, Luan T, Odegard J. Accuracy of genomic selection for a sib-
646 evaluated trait using identity-by-state and identity-by-descent relationships. Genet. Sel. Evol. GSE.
647 2015;47:9.

648 30. Junqueira VS, Cardoso FF, Oliveira MM, Sollero BP, Silva FF, Lopes PS. Use of molecular markers
649 to improve relationship information in the genetic evaluation of beef cattle tick resistance under pedigree-
650 based models. J. Anim. Breed. Genet. 2017;134:14–26.

651 31. Durel CE, Laurens F, Fouillet A, Lespinasse Y. Utilization of pedigree information to estimate genetic
652 parameters from large unbalanced data sets in apple. Theor. Appl. Genet. 1998;96:1077–85.

653 32. de Souza VAB, Byrne DH, Taylor JF. Predicted breeding values for nine plant and fruit characteristics
654 of 28 peach genotypes. J. Am. Soc. Hortic. Sci. 2000;125:460–5.

655 33. Kouassi A, Durel C-E, Costa F, Tartarini S, van de Weg E, Evans K, et al. Estimation of genetic
656 parameters and prediction of breeding values for apple fruit-quality traits using pedigreed plant material in
657 Europe. Tree Genet. Genomes. 2009;5:659–72.

658 34. Stephens MJ, Alspach PA, Beatson RA, Winefield C, Buck EJ. Genetic parameters and breeding for
659 yield in red raspberry. J. Am. Soc. Hortic. Sci. 2012;137:229–35.

660 35. Whitaker VM, Osorio LF, Hasing T, Gezan S. Estimation of genetic parameters for 12 fruit and
661 vegetative traits in the University of Florida strawberry breeding population. J. Am. Soc. Hortic. Sci.
662 2012;137:316–24.

663 36. Fresnedo-Ramírez J, Crisosto CH, Gradziel TM, Famula TR. Pedigree correction and estimation of
664 breeding values for peach genetic improvement. Acta Hortic. 2015;249–56.

665 37. Gezan SA, Osorio LF, Verma S, Whitaker VM. An experimental validation of genomic selection in
666 octoploid strawberry. Hortic. Res. 2017;4:16070.

667 38. Tancred SJ, Zeppa AG, Cooper M, Stringer JK. Heritability and patterns of inheritance of the ripening
668 date of apples. HortScience. 1995;30:325–8.

669 39. Hardner CM, Kumar S, Peace CM, Luby J, Evans KM. Reconstructing relationship matrices from
670 dense SNP arrays for the prediction of genetic potential in unreplicated multilocation plantings of apple
671 progeny. Acta Hortic. 2016;275–82.

672 40. Furlani RCM, Moraes MLT de, Resende MDV de, Furlani Junior E, Gonçalves P de S, Valério Filho
673 WV, et al. Estimation of variance components and prediction of breeding values in rubber tree breeding
674 using the REML/BLUP procedure. Genet. Mol. Biol. 2005;28:271–6.

675 41. Hardner CM, Healey AL, Downes G, Herberling M, Gore PL. Improving prediction accuracy and
676 selection of open-pollinated seed-lots in *Eucalyptus dunnii* Maiden using a multivariate mixed model
677 approach. Ann. For. Sci. 2016;73:1035–46.

678 42. Imai A, Kuniga T, Yoshioka T, Nonaka K, Mitani N, Fukamachi H, et al. Evaluation of the best linear
679 unbiased prediction method for breeding values of fruit-quality traits in citrus. Tree Genet. Genomes.
680 2016;12:119.

681 43. Minamikawa MF, Nonaka K, Kaminuma E, Kajiya-Kanegae H, Onogi A, Goto S, et al. Genome-wide
682 association study and genomic prediction in citrus: Potential of genomics-assisted breeding for fruit
683 quality traits. Sci. Rep. 2017;7:4721.

684 44. Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for
685 complex traits. PLOS Genet. 2008;4:e1000008.

686 45. Kumar S, Molloy C, Muñoz P, Daetwyler H, Chagne D, Volz R. Genome-enabled estimates of additive
687 and nonadditive genetic variances and prediction of apple phenotypes across environments. G3.
688 2015;5:2711–8.

689 46. Iezzoni A, Weebadde C, Luby J, Yue C, van de Weg E, Fazio G, et al. RosBREED: enabling marker-
690 assisted breeding in Rosaceae. Acta Hortic. 2010;859:389–94.

691 47. Long L, Kaiser C. Sweet cherry rootstocks for the Pacific Northwest. Corvallis: Oregon State
692 University; 2010 p. 1–8. Report No.: PNW 619.

693 48. Peace C, Luby J, Van de Weg WE, Bink M, Iezzoni A. A strategy for developing representative
694 germplasm sets for systematic QTL validation, demonstrated for apple, peach, and sweet cherry. Tree
695 Genet. Genomes. 2014;10:1679–94.

696 49. Chavoshi M, Watkins C, Oraguzie B, Zhao Y, Iezzoni A, Oraguzie N. Phenotyping protocol for sweet
697 cherry (*Prunus avium* L.) to facilitate an understanding of trait inheritance. Am. Pomol. Soc. 2014;68:125–
698 34.

699 50. Washington State University. AgWeatherNet Roza Station Data [Internet]. AgWeatherNet. [cited 2017
700 Jul 11]. Available from: www.weather.wsu.edu

701 51. McMaster G, Wilhelm W. Growing degree-days: one equation, two interpretations. Agric. For.
702 Meteorol. 1997;87:291–300.

703 52. Peace C, Bassil N, Main D, Ficklin S, Rosyara UR, Stegmeir T, et al. Development and evaluation of
704 a genome-wide 6K SNP Array for diploid sweet cherry and tetraploid sour cherry. PLoS ONE.
705 2012;7:e48305.

706 53. Cai L, Voorrips RE, van de Weg E, Peace C, Iezzoni A. Genetic structure of a QTL hotspot on
707 chromosome 2 in sweet cherry indicates positive selection for favorable haplotypes. Mol. Breed.
708 2017;37:85.

709 54. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase
710 inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 2009;84:210–23.

711 55. Wimmer V, Albrecht T, Auinger H-J, Schoen C-C. Synbreed: a framework for the analysis of genomic
712 prediction data using R. Bioinformatics. 2012;28:2086–7.

713 56. Butler D, Cullis B, Gilmour A, Gogel B. Analysis of mixed models for S language environments:
714 ASReml-R Reference Manual (version 3). The State of Queensland, Department of Primary Industries
715 and Fisheries; 2009.

716 57. R Development Core Team. R: A language and environment for statistical computing [Internet]. 2011.
717 Available from: http://www.R-project.org/

718 58. Endelman JB. Ridge regression and other kernels for genomic selection with R package rrBLUP.
719 Plant Genome. 2011;4:25–255.

720 59. VanRaden PM. Efficient methods to compute genomic predictions. J. Dairy Sci. 2008;91:4414–23.

721 60. Su G, Christensen OF, Ostersen T, Henryon M, Lund MS. Estimating additive and non-additive
722 genetic variances and predicting genetic merits using genome-wide dense single nucleotide
723 polymorphism markers. PLOS ONE. 2012;7:e45293.

724 61. Piaskowski J. Genomic Dominance Relationship Matrix [Internet]. 2017. Available from:
725 https://github.com/jpiaskowski/Genomic-Dominance-Relationship-Matrix

726 62. Brien C. asremlPlus [Internet]. 2016. Available from: https://cran.r-
727 project.org/web/packages/asremlPlus/asremlPlus.pdf

728 63. Gabriel K. The biplot graphic display of matrices with application to principal component analysis.
729 Biometrika. 1971;58:453–67.

730 64. Rodriguez-Almeida FA, Van Vleck LD, Willham RL, Northcutt SL. Estimation of non-additive genetic
731 variances in three synthetic lines of beef cattle using an animal model. J. Anim. Sci. 1995;73:1002–11.

732 65. Pante M, Gjerde B, McMillan I, Misztal I. Estimation of additive and dominance genetic variances for
733 body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. Aquaculture. 2002;204:383–92.

734 66. Gallardo JA, Lhorente JP, Neira R. The consequences of including non-additive effects on the genetic
735 evaluation of harvest body weight in Coho salmon (*Oncorhynchus kisutch*). Genet. Sel. Evol. 2010;42:19.

736 67. Gamal El-Dien O, Ratcliffe B, Klapste J, Porth I, Chen C, El-Kassaby YA. Implementation of the
737 realized genomic relationship matrix to open-pollinated white spruce family testing for disentangling
738 additive from nonadditive genetic effects. G3. 2016;6:743–53.

739 68. Wolfe MD, Kulakow P, Rabbi IY, Jannink J-L. Marker-based estimates reveal significant non-additive
740 effects in clonally propagated cassava (*Manihot esculenta*): implications for the prediction of total genetic
741 value and the selection of varieties. G3. 2016;

742 69. Ceballos H, Kawuki RS, Gracen VE, Yencho GC, Hershey CH. Conventional breeding, marker-
743 assisted selection, genomic selection and inbreeding in clonally propagated crops: a case study for
744 cassava. Theor. Appl. Genet. 2015;128:1647–67.

745 70. Zhao Y, Athanson B, Whiting M, Oraguzie N. Pedicel-fruit retention force in sweet cherry (*Prunus*
746 *avium* L.) varies with genotype and year. Sci. Hortic. 2013;150:135–41.

747 71. Schmitz CA, Clark MD, Luby JJ, Bradeen JM, Guan Y, Evans K, et al. Fruit texture phenotypes of the
748 RosBREED U.S. apple reference germplasm set. HortScience. 2013;48:296–303.

749 72. Cantín CM, Gogorcena Y, Moreno MÁ. Analysis of phenotypic variation of sugar profile in different
750 peach and nectarine [*Prunus persica* (L.) Batsch] breeding progenies. J. Sci. Food Agric. 2009;89:1909–
751 17.

752 73. Mathey MM, Mookerjee S, Mahoney LL, Gündüz K, Rosyara U, Hancock JF, et al. Genotype by
753 environment interactions and combining ability for strawberry families grown in diverse environments.
754 Euphytica. 2017;213:112.

755 74. Clark S, van de Werf J. Genomic best unbiased linear prediction (gBLUP) for the estimation of
756 genomic breeding values. In: Gondro C, van de Werf J, Hayes B, editors. Genome-Wide Assoc. Stud.
757 Genomic Predict. Springer; 2013. p. 321–30.

758 75. Heslot N, Jannink J-L, Sorrells ME. Perspectives for genomic selection applications and research in
759 plants. Crop Sci. 2015;55:1–12.

760 76. Dirlewanger E, Quero-Garcia J, Le Dantec L, Lambert P, Ruiz D, Dondini L, et al. Comparison of the
761 genetic determinism of two key phenological traits, flowering and maturity dates, in three *Prunus* species:
762 peach, apricot and sweet cherry. Heredity. 2012;109:280–92.

763 77. Kim Y, Kimball JS, Didan K, Henebry GM. Response of vegetation growth and productivity to spring
764 climate indicators in the conterminous United States derived from satellite remote sensing data fusion.
765 Agric. For. Meteorol. 2014;194:132–43.

766 78. Zhang G, Sebolt A, Sooriyapathirana S, Wang D, Bink M, Olmstead J, et al. Fruit size QTL analysis of
767 an F1 population derived from a cross between a domesticated sweet cherry cultivar and a wild forest
768 sweet cherry. Tree Genet. Genomes. 2010;6:25–36.

769 79. De Franceschi P, Stegmeir T, Cabrera A, van der Knaap E, Rosyara UR, Sebolt AM, et al. Cell
770 number regulator genes in *Prunus* provide candidate genes for the control of fruit size in sweet and sour
771 cherry. Mol. Breed. 2013;32:311–26.

772 80. Campoy JA, Le Dantec L, Barreneche T, Dirlewanger E, Quero-García J. New insights into fruit
773 firmness and weight control in sweet cherry. Plant Mol. Biol. Report. 2015;33:783–96.

774 81. Bassi D, Bartolozzi F, Muzzi E. Patterns and heritability of carboxylic acids and soluble sugars in fruits
775 of apricot (*Prunus armeniaca* L.). Plant Breed. 1996;115:67–70.

776 82. Brooks SJ, Moore JN, Murphy JB. Quantitative and qualitative changes in sugar content of peach
777 genotypes [*Prunus persica* (L.) Batsch.]. J. Am. Soc. Hortic. Sci. 1993;118:97–100.

778 83. Genard M, Lescourret F, Gomez L, Habib R. Changes in fruit sugar concentrations in response to
779 assimilate supply, metabolism and dilution: a modeling approach applied to peach fruit (*Prunus persica*).
780 Tree Physiol. 2003;23:373–85.

781 84. Morandi B, Corelli Grappadelli L, Rieger M, Lo Bianco R. Carbohydrate availability affects growth and
782 metabolism in peach fruit. Physiol. Plant. 2008;133:229–41.

783 85. Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, et al. Mapping QTLs controlling
784 fruit quality in peach (*Prunus persica* (L.) Batsch). Theor. Appl. Genet. 1999;98:18–31.

785 86. Olmstead J, Lang G. A leaf disk assay for screening sweet cherry genotypes for susceptibility to
786 powdery mildew. HortScience. 2000;35:274–7.

787 87. Olmstead J, Lang G, Grove G. Inheritance of powdery mildew resistance in sweet cherry.
788 HortScience. 2001;36:337–40.

789 88. Wilfert L, Schmid-Hempel P. The genetic architecture of susceptibility to parasites. BMC Evol. Biol.
790 2008;8:187–187.

791 89. Divya B, Biswas A, Robin S, Rabindran R, Joel AJ. Gene interactions and genetics of blast resistance
792 and yield attributes in rice (*Oryza sativa* L.). J. Genet. 2014;93:415–24.

793 90. Zhang J, Singh A, Mueller DS, Singh AK. Genome-wide association and epistasis studies unravel the
794 genetic architecture of sudden death syndrome resistance in soybean. Plant J. Cell Mol. Biol.
795 2015;84:1124–36.

796 91. Cach NT, Perez JC, Lenis JI, Calle F, Morante N, Ceballos H. Epistasis in the expression of relevant
797 traits in cassava (*Manihot esculenta* Crantz) for subhumid conditions. J. Hered. 2005;96:586–92.

798 92. Oakey H, Verbyla A, Pitchford W, Cullis B, Kuchel H. Joint modeling of additive and non-additive
799 genetic line effects in single field trials. Theor. Appl. Genet. 2006;113:809–19.

800 93. Bai W, Zhang H, Zhang Z, Teng F, Wang L, Tao Y, et al. The evidence for non-additive effect as the
801 main genetic component of plant height and ear height in maize using introgression line populations.
802 Plant Breed. 2010;129:376–84.

803 94. Wang D, Salah El-Basyoni I, Stephen Baenziger P, Crossa J, Eskridge KM, Dweikat I. Prediction of
804 genetic values of quantitative traits with epistatic effects in plant breeding populations. Heredity.
805 2012;109:313–9.

806 95. Dudley JW, Johnson GR. Epistatic models and pre-selection of markers improve prediction of
807 performance in corn. Mol. Breed. 2013;32:585–93.

808 96. Nazarian A, Gezan SA. Integrating nonadditive genomic relationship matrices into the study of genetic
809 architecture of complex traits. J. Hered. 2016;107:153–62.

810 97. Cockerham CC. An extension of the concept of partitioning hereditary variance for analysis of
811 covariances among relatives when epistasis is present. Genetics. 1954;39:859–82.

812 98. Holland J. Epistasis and plant breeding. In: Janick J, editor. Plant Breed. Rev. Oxford, UK: Oxford,
813 UK: John Wiley & Sons, Inc.; 2010. p. 27–92.

814 99. Tao R, Iezzoni AF. The S-RNase-based gametophytic self-incompatibility system in *Prunus* exhibits
815 distinct genetic and molecular features. Sci. Hortic. 2010;124:423–33.

816 100. Heffner EL, Sorrells ME, Jannink J-L. Genomic selection for crop improvement. Crop Sci.
817 2009;49:1–12.

818 101. Ratcliffe B, El-Dien OG, Klapste J, Porth I, Chen C, Jaquish B, et al. A comparison of genomic
819 selection models across time in interior spruce (*Picea engelmannii* x *glauca*) using unordered SNP
820 imputation methods. Heredity. 2015;115:547–55.

821 102. Muranty H, Troggio M, Sadok IB, Rifaï MA, Auwerkerken A, Banchi E, et al. Accuracy and responses
822 of genomic selection on key traits in apple breeding. Hortic. Res. 2015;2:15060.

823 103. Biscarini F, Nazzicari N, Bink M, Arús P, Aranzana MJ, Verde I, et al. Genome-enabled predictions
824 for fruit weight and quality from repeated records in European peach progenies. BMC Genomics.
825 2017;18:432.

826 104. Garcia MR, Carbonell EA, Asíns MJ. QTL analysis of yield and seed number in Citrus. Theor. Appl.
827 Genet. 2000;101:487–93.

828 105. Resende MFR, Muñoz P, Acosta JJ, Peter GF, Davis JM, Grattapaglia D, et al. Accelerating the
829 domestication of trees using genomic selection: accuracy of prediction models across ages and
830 environments. New Phytol. 2012;193:617–24.

831 106. Resende RT, Resende MDV, Silva FF, Azevedo CF, Takahashi EK, Silva-Junior OB, et al.
832 Assessing the expected response to genomic selection of individuals and families in Eucalyptus breeding
833 with an additive-dominant model. Heredity. 2017;119:245–55.

834 107. Marulanda JJ, Mi X, Melchinger AE, Xu J-L, Würschum T, Longin CFH. Optimum breeding strategies
835 using genomic selection for hybrid breeding in wheat, maize, rye, barley, rice and triticale. Theor. Appl.
836 Genet. 2016;129:1901–13.

837 108. Jung S, Ficklin SP, Lee T, Cheng C-H, Blenda A, Zheng P, et al. The Genome Database for
838 Rosaceae (GDR): year 10 update. Nucleic Acids Res. 2014;42:D1237-44.

839

840

FIGURE CAPTIONS

Figure 1: Violin plots of nine traits by year, adjusted for fixed effects due to year and overlaid with the observations from each year.

Figure 2: Heritability and predictive ability of four genetic models for each of nine focus traits.

Figure 3: Variance components by each year and across years for the full (ADI) model.

Figure 4: Biplot of genetic values among the RosBREED sweet cherry Crop Reference Set using the correlation matrix of eight traits. Trait rotations were scaled by the first eigenvalue.

ADDITIONAL FILE CAPTIONS

Additional File 1: Individuals from the RosBREED sweet cherry Crop Reference set used in this study. "Self" refers to individuals derived from self-pollination, and "Unk" means that at least one parent is not known (file extension is .csv).

Additional File 2: Number of observations (N) and Spearman rank correlations (ρ) between genetic values derived from a single year and the multi-year genetic values using the ADI model (Panel A) or the phenotypic data (Panel B), genetic values derived from the reduced models and the full model (Panel C), and breeding values derived from the reduced models and the full model (Panel D) (file extension is .xlsx).

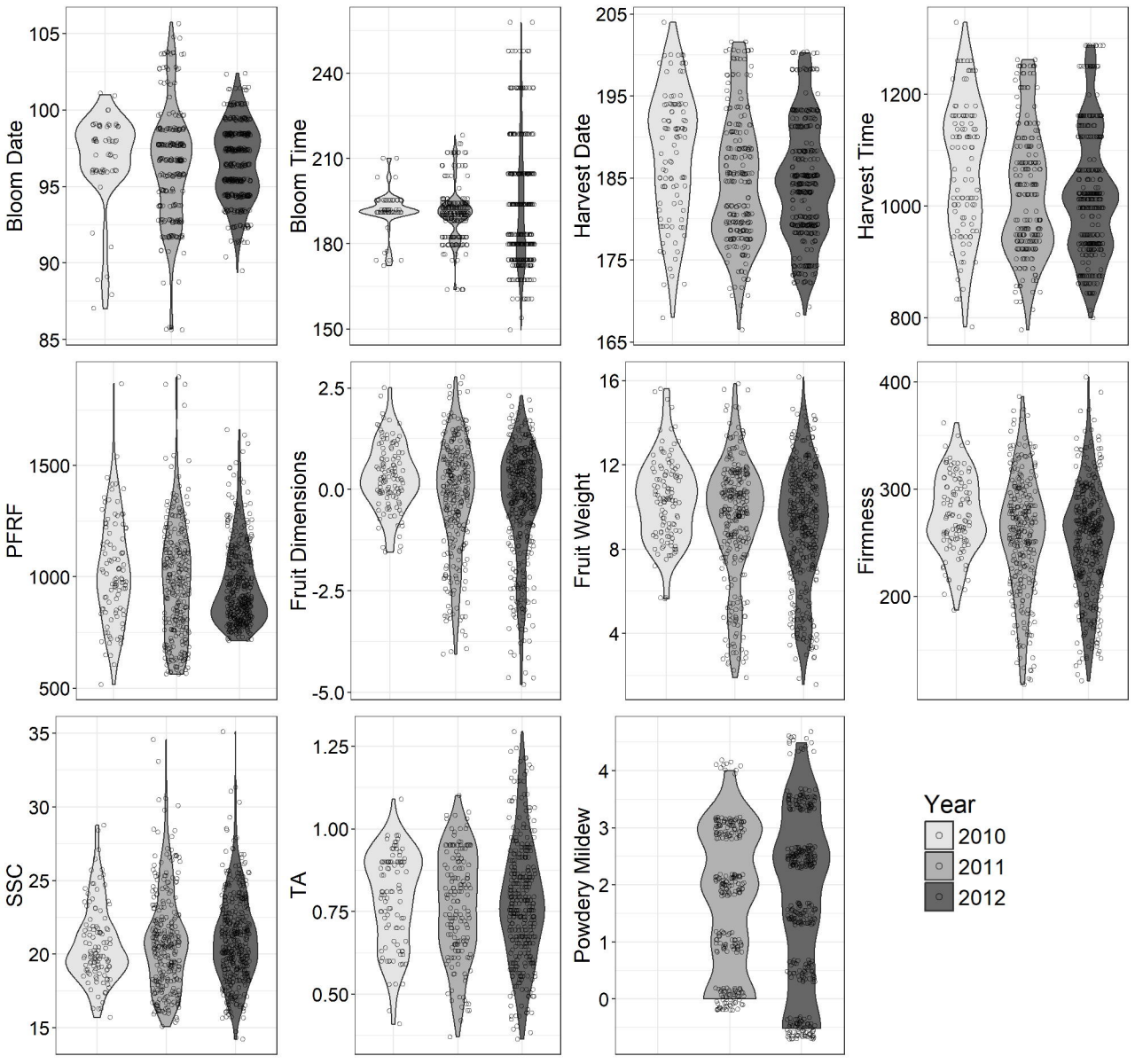
Additional File 3: Histogram of the diagonals and off-diagonal from the additive relationship matrix (file extension is .png).

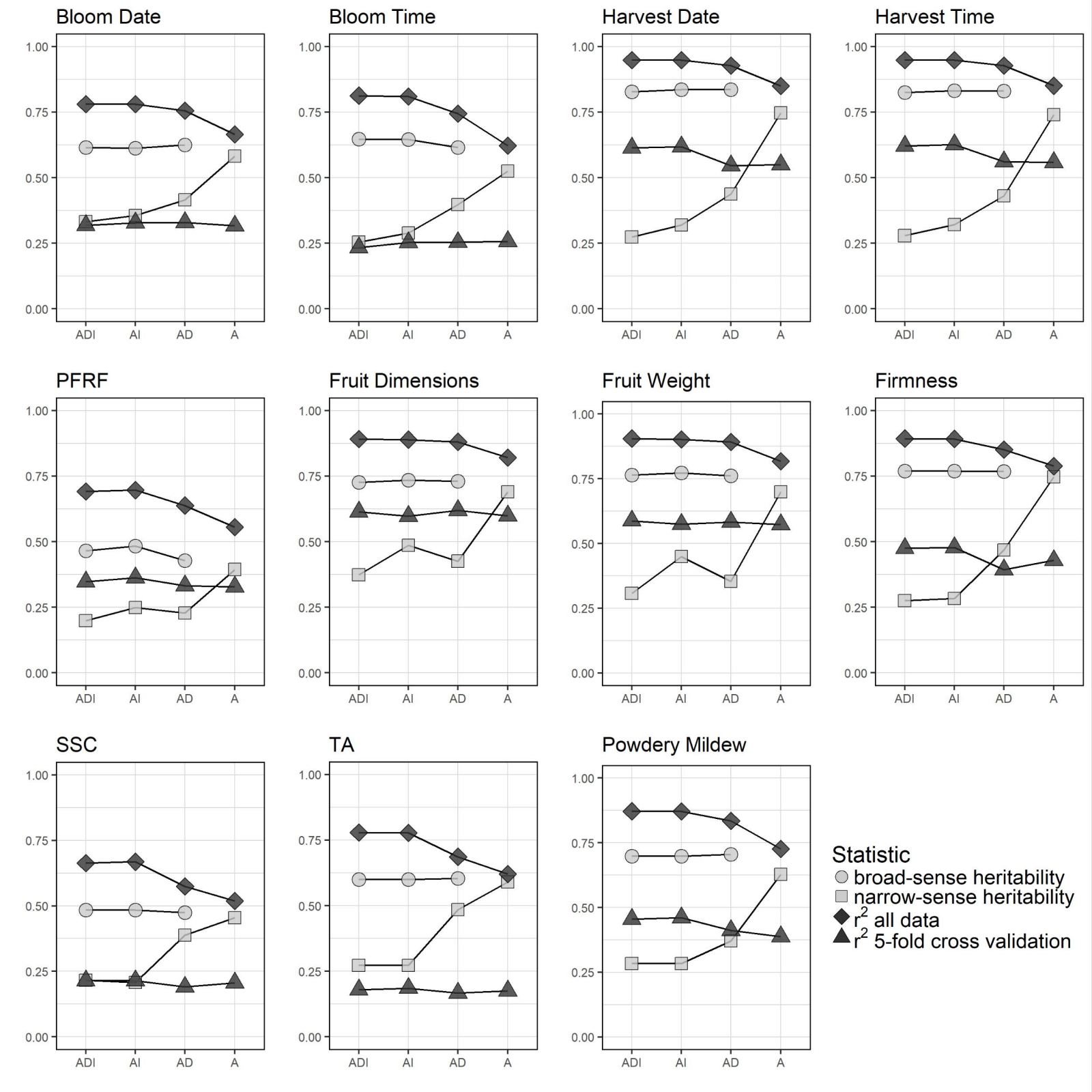
Additional File 4: Variance component estimates and standard errors for all RosBREED sweet cherry traits (file extension is .csv).

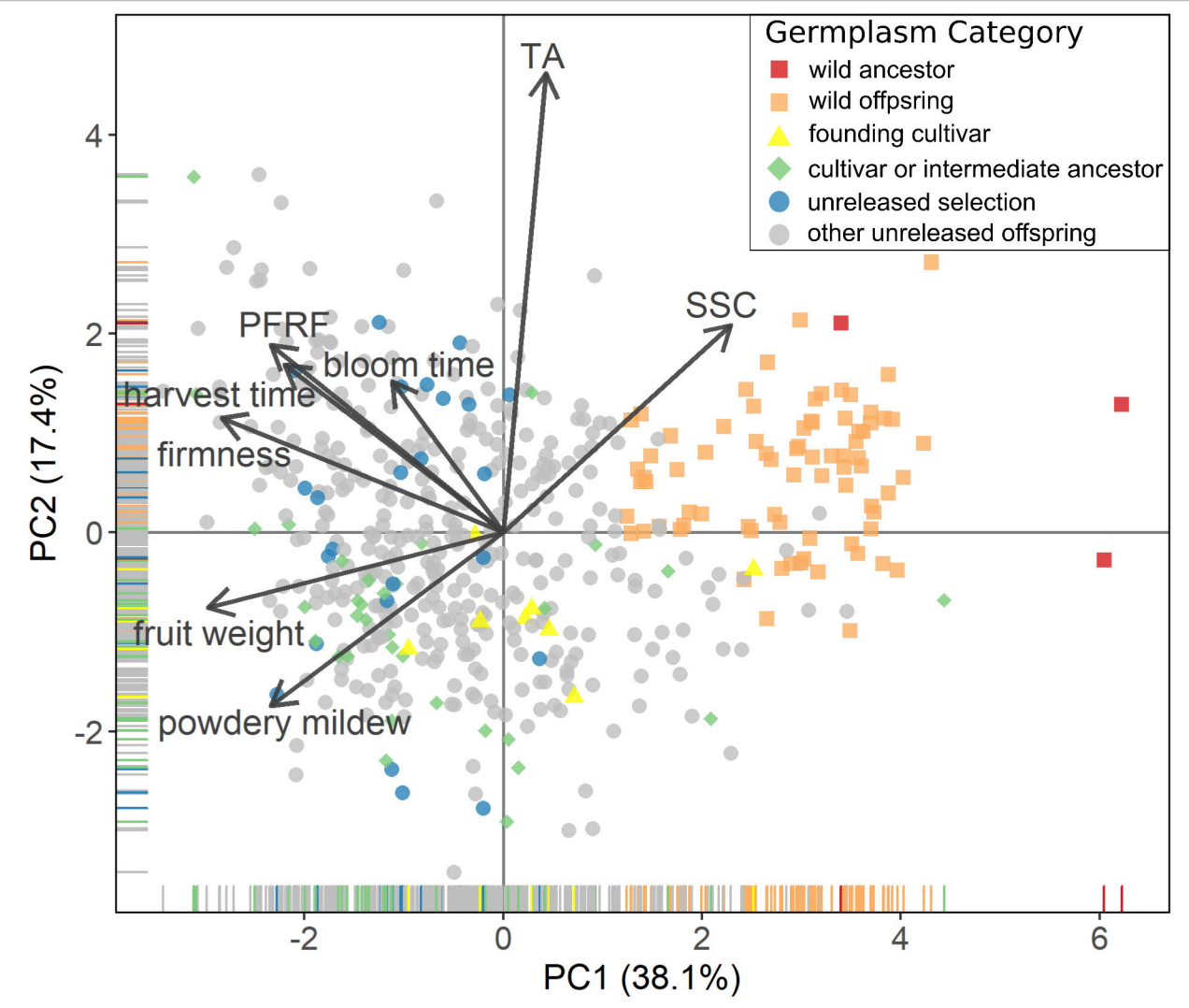
Additional File 5: Variance component percentages for all RosBREED sweet cherry traits (file extension is .csv).

Additional File 6: Biplot of genetic values among sweet cherry cultivars and their ancestors using the correlation matrix of eight traits. Trait rotations were scaled by the first eigenvalue (file extension is .png).

Additional File 7: Breeding values, dominance values, epistatic values and genetic values of all individuals for all traits in the RosBREED sweet cherry Crop Reference Set (file extension is .xlsx).


Bloom Date
105
100
95
90
85
Bloom Time
240
210
180
150
Harvest Date
205
195
185
175
165
Harvest Time
1200
1000
800
PFRF
1500
1000
500
Fruit Dimensions
2.5
0.0
-2.5
-5.0
Fruit Weight
16
12
8
4
Firmness
400
300
200
SSC
35
30
25
20
15
TA
1.25
1.00
0.75
0.50
Powdery Mildew
4
3
2
1
0
Year
2010
2011
2012


Bloom Date
Bloom Time
Harvest Date
Harvest Time
PFRF
Fruit Dimensions
Fruit Weight
Firmness
SSC
TA
Powdery Mildew
1.00
0.75
0.50
0.25
0.00
ADI
AI
AD
A
Statistic
broad-sense heritability
narrow-sense heritability
r2 all data
r2 5-fold cross validation


Germplasm Category
wild ancestor
wild offpsring
founding cultivar
cultivar or intermediate ancestor
unreleased selection
other unreleased offspring
TA
SSC
PFRF
bloom time
harvest time
firmness
fruit weight
powdery mildew
PC2 (17.4%)
PC1 (38.1%)
4
2
0
-2
-2
0
2
4
6