

RESEARCH ARTICLE

Open Access



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

Julia Piaskowski^{1*} , Craig Hardner², Lichun Cai³, Yunyang Zhao⁴, Amy Iezzoni³ and Cameron Peace¹

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 0.76 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 model 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 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 kt 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*

* Correspondence: jpiaskowski@wsu.edu

¹Department of Horticulture, Washington State University, Pullman, WA 99164-6414, USA

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



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

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

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

This study addresses a deficiency of published information on genetic parameters for sweet cherry breeding-

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

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 included 77 F1 offspring derived from a wild parent and 359 F1 offspring derived from existing cultivars. All individuals are connected through pedigree or historic relationships across six generations. 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].

Gisela 6 rootstock can impact crop performance by decreasing tree size and increasing crop load, among other effects [49–51]. Because only released cultivars and wild germplasm were grown on rootstocks while all seedlings were grown on their own roots, any effect of rootstock was confounded with the effect of seedling versus non-seedling in this study. However, in this study, understanding the relative performance among seedlings or among cultivars was the primary purpose in order to identify possible parents and candidate cultivars.

Phenotypic data

This study used the sweet cherry phenotypic data set previously described in Chavoshi et al. [52] obtained in the RosBREED project. This data set consisted of 32 traits evaluated in 2010, 2011, and 2012. Standardized phenotyping protocols for sweet cherry [52] 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, respectively. Foliar powdery mildew incidence was scored in August of each year, immediately after the fruiting season, on a 0–5 scale, where 0 is no infection and 5 is highly infected leaves. These nine traits are referred to as “focus traits” for the rest of the study. All trait data were measured over three years except for powdery mildew incidence, which was not assayed in 2010. Results from the other traits are given in the supplementary material, but not discussed.

Several transformations of the trait data were performed for the focus traits. “Fruit dimensions” was determined newly here as the first component from a principal component analysis between fruit length and fruit width, which are both end-to-end fruit measurements in millimeters. The first principal component summarized 95.4% of total phenotypic variation for fruit length and width. Growing degree days was calculated for an alternative measure of phenological traits. Climatic data was obtained from Washington State University’s AgWeatherNet using the “Roza” station [53], using a base temperature of 4.5 °C and maximum of 30 °C. Daily maximum temperatures above 30 °C were reduced to 30 °C, and negative temperatures were set to zero, following McMaster and Wilhelm [54]. Erroneous data points, defined as those larger than twice the next largest value or less than one-half of the next smallest value and having a studentized residual with an absolute value greater than 5, were removed. Such data were assumed to be data entry errors. There were 97 individuals with no phenotypic data: 13 selections and 84 unselected

progeny. These individuals were used in the model-building and prediction steps for all models except for cross validation.

SNP data

The SNP data were obtained from the RosBREED project using the RosBREED cherry 6 K SNP array v1 (an Illumina Infinium® II array) [55]. The SNP curation pipeline was described in Cai et al. [56]. Missing data were imputed with Beagle as implemented in SynBreed [57, 58] using the hidden Markov model and a minor allele frequency of 0.05. Individuals or SNPs missing more than 25% data were removed from analysis. In total, a genome-wide set of 1615 SNPs was used.

Statistical modeling

Variance components were estimated with R-ASReml 3.0 [59], and additional statistical analyses were conducted in R v3.4 [60]. The following model was used for initial estimates of genetic effects for a single trait, **Y**:

$$Y = Xb + Z_1a + Z_2d + Z_3i + Z_4a_Y + Z_5d_Y + Z_6i_Y + e$$

where **a**, **d**, **i**, **a_Y**, **d_Y** and **i_Y** are the random variables for additive effects, dominance effects, effects from additive-by-additive epistatic, additive-by-year effects, dominance-by-year effects, and epistasis-by-year effects, respectively. Variables **Z₁₋₃** and **Z₄₋₆** are design matrices for main effects and interaction terms, respectively. Dimensions of **Z₁₋₃** are $nY \times Y$ and **Z₄₋₆** are $nY \times nY$, where n is the number of individuals and Y is the number of years with trait data for an individual. Year was treated as a fixed effect, where **X** is the design matrix relating observations to years and **b** is a vector of fixed effects due to year. In a preliminary analysis, the effect of location was evaluated as a fixed effect using a Wald test. Location did not have a significant effect on the focus traits (p -value > 0.10) and was omitted from the model. Random variables were assumed to follow a normal distribution:

$$\begin{aligned} a &\sim N(0, G_a \sigma_a^2), d \sim N(0, D \sigma_d^2), i \sim N(0, G_{aa} \sigma_{aa}^2), \\ a_Y &\sim N(0, I_Y \otimes G_a \sigma_{aY}^2), d_Y \sim N(0, I_Y \otimes D \sigma_{dY}^2), \\ i_Y &\sim N(0, I_Y \otimes G_{aa} \sigma_{aaY}^2) e \sim N(0, R) \end{aligned}$$

The covariance structure for year was modeled as a repeated measure: $R = I_{Individual} \otimes e_Y$ where $I_{Individual}$ is an identity matrix of individuals included in the study and e_Y is a 3×3 matrix of year error terms using a general correlation structure implemented in ASReml. The genomic additive relationship matrix was computed with R/rBLUP [61] using the VanRaden method [62]:

$$G_a = \frac{HH^T}{2 \sum_j p_j(1-p_j)}$$

where p_j is frequency of the positive allele for a single marker column, and **H** was computed as equal to centered marker data, $\{H\}_{ij} = \{M\}_{ij} - 2(p_j - 0.5)$. **M** is an $n \times m$ marker matrix with n individuals and m markers expressed as $(-1, 0, 1)$ frequency. The dominance relationship matrix was computed using normalized matrices described by Su et al. [63] and implemented using a custom R program [64]:

$$D = \frac{ZZ^T}{\sum_j 2p_j(1-p_j)(1-2p_j(1-p_j))}$$

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

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

The epistatic relationship matrix for additive-by-additive effects was computed by taking the Hadamard product between G_a , the additive genomic relationship matrix, and itself: $G_{aa} = G_a \cdot G_a$.

When a relationship matrix was not positive definite, a 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 $0.002 \times$ 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:

$$WC^{-1}W^T$$

$$\text{Where } C = W^T R^{-1} W + \begin{pmatrix} 0 & 0 \\ 0 & G^{-1} \end{pmatrix} \text{ and } W = [X \ Z].$$

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 *asreml-Plus* [65] 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_a^2 for narrow-sense heritability (h^2) and as $\sigma_a^2 + \sigma_d^2 + \sigma_{aa}^2$ 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 **a_y**, **d_y**, and **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 predictions. 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 [66], where the rotations for plotting the variables were scaled by the first eigenvalue.

Results

Distribution of phenotypic data

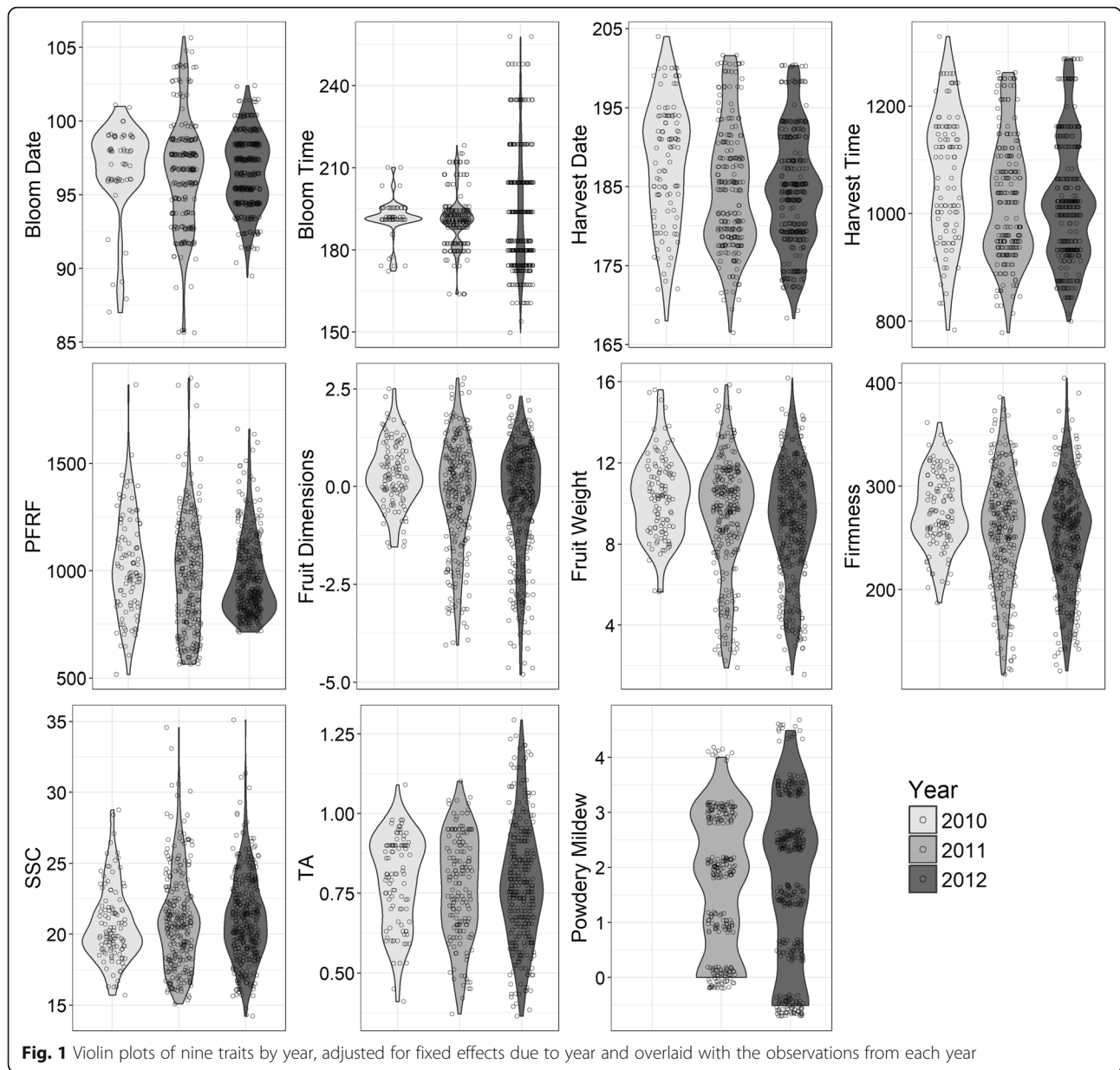
All trait distributions (consisting of 600–755 data points for each trait) were influenced by the year of data collection (Fig. 1). Wald test results for year were consistently highly significant for all focus traits across all models ($p < 0.001$ in all cases).

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 likely 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 the full model (p -values < 0.05) for all focus traits, except for powdery mildew incidence and SSC, which had notable p -values defined as less than 0.10. Reduced models consisting of single main effects (additive, dominance or epistatic) or single main effects plus their year interaction term (e.g., additive and additive-by-year) were highly significant for all traits. This demonstrates that the reduced models did not adequately capture variation compared to the full model. For most focus traits, genetic models that included additive, epistatic, additive-by-year and epistasis-by-year effects were not statistically different from the full model. Thus, dominance and dominance-by-year could be dropped from their genetic models without significant loss

of information. Traits that were exceptions to the above were fruit weight, fruit dimensions, and bloom date, for which optimal fit was obtained by including dominance in the model. For all traits, dominance-by-year and epistasis-by-year effects could be removed from the model without much loss of information. Additive-by-year effects had a statistically significant effect on bloom date, bloom time, and PFRF ($p < 0.01$).

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).

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

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 (a_Y = 11%) and TA (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

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

and traits, and variance percentages, are provided in Additional files 4 and 5, respectively.

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 breeding values estimated from a single year and their respective phenotypes implied a better fit within a year for each year and trait than the single-year breeding values with their multi-year counterparts ($\rho = 0.64$ – 1.00) (Additional file 2).

Correlations among trait genetic values

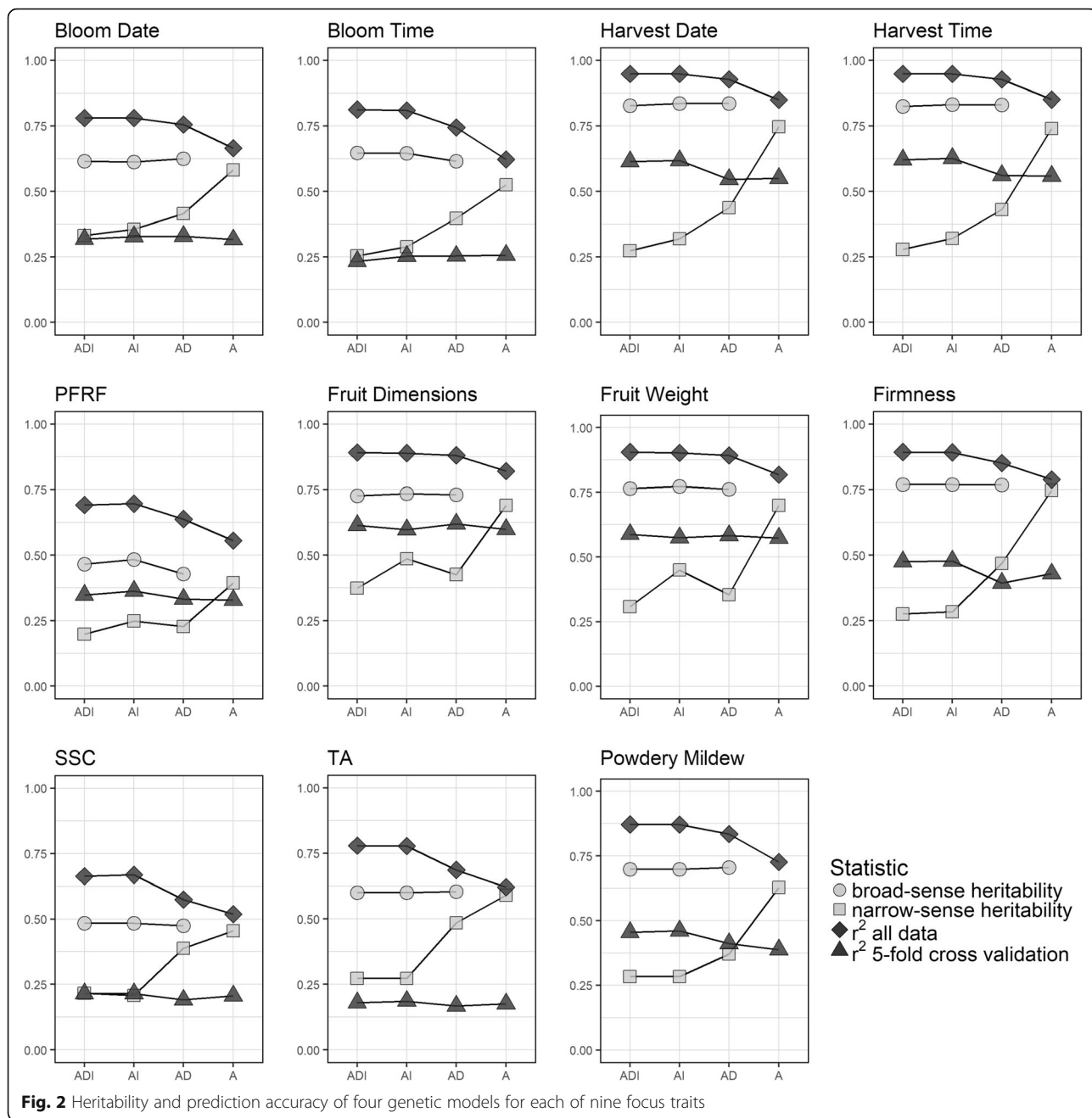
The genetic values of the focus traits had weak to moderate positive correlations with each when considering only unreleased offspring and selections, with some exceptions (Table 3). Fruit weight and fruit dimensions, harvest date and harvest time, and bloom date and bloom time were all highly correlated 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 file 6 further separates the sweet cherry founders and derived cultivars by fruit weight and SSC content.

Discussion

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

Importance of model fit and consequences for predictive ability

This study demonstrated that for most traits, non-additive sources of variation comprised an equal or larger portion of the genetic variance than additive variance. A genetic



model including additive, epistatic, additive-by-year and epistasis-by-year effects was usually the most parsimonious approach for capturing major sources of variation. Exceptions were fruit dimensions and fruit weight, which instead were best described by a model with additive, dominance and additive-by-year effects, and harvest date, best described by a main effects-only model.

Using an incorrect model to determine genome-wide breeding values can provide misleading information for making breeding decisions. Table 4 illustrates the consequences of using a poorly-fitting reduced model for

estimating breeding values. Breeding values were often larger in relative magnitude in the reduced models compared to the full model, which can exaggerate genetic gains possible in the population. For example, days to maturity in an Ambrunes/Sweetheart cross would be overestimated by twice as many days in the additive-only model compared to the ADI model. Likewise, crosses with the wild accession MIM 23 were predicted to result in midparent values of fruit size twice as small in the A model compared to the ADI model (Table 4). The inflation of additive variance when non-additive sources are omitted has been

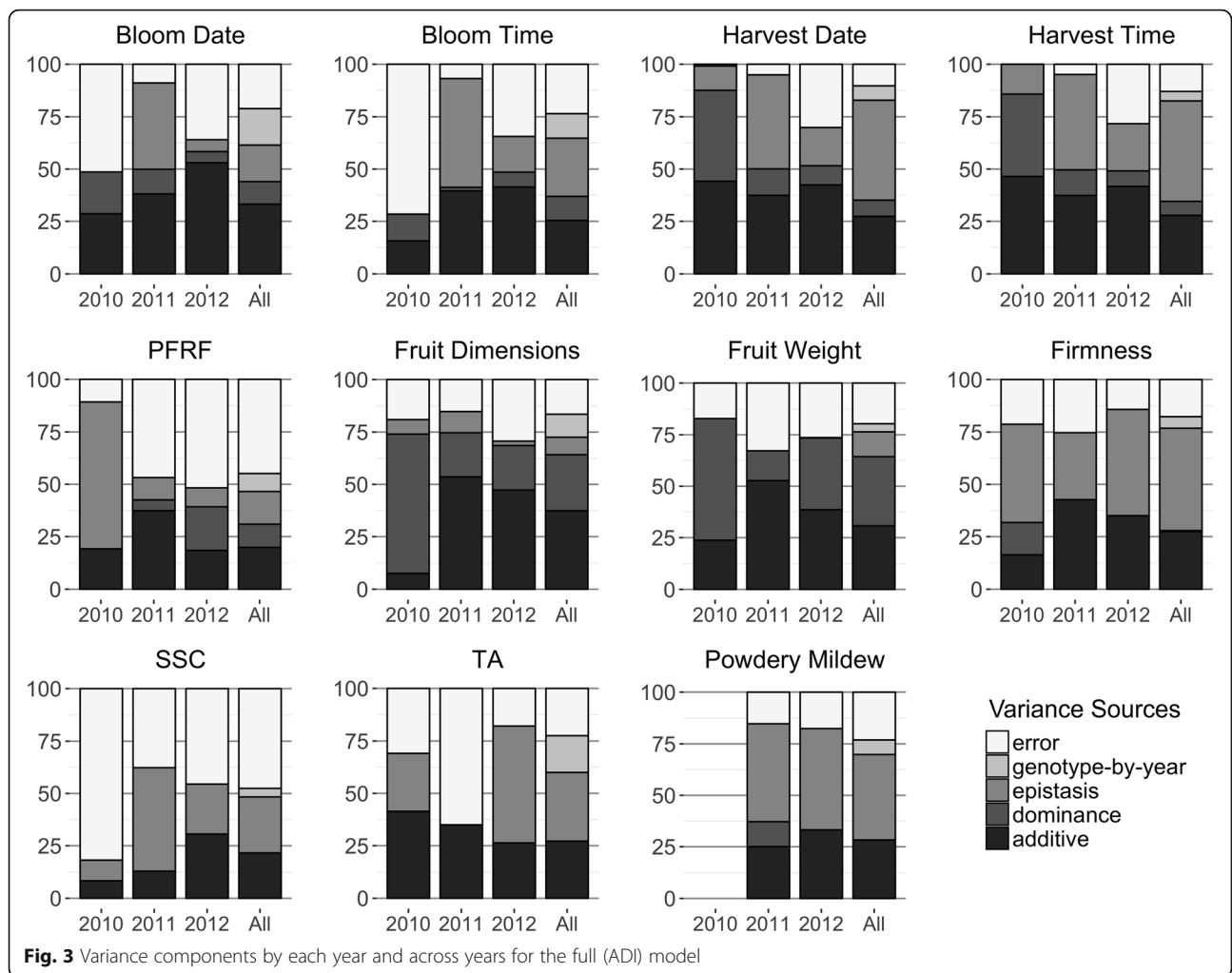
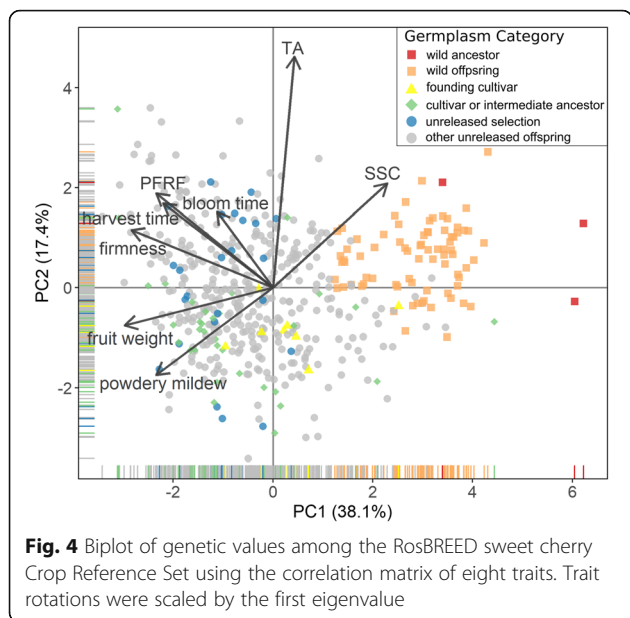


Fig. 3 Variance components by each year and across years for the full (ADI) model

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	17,630	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***	-	0.505***
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



documented in several other species including apple, loblolly pine, white spruce cassava, cattle, pigs, Coho salmon, and rainbow trout [27, 45, 63, 67–71].

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. [72] 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 (Fig. 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 other rosaceous crops, including sweet cherry pedicel-fruit retention force [73], apple fruit texture [74], sugar content in peach and nectarine [75], and several phenological and fruit quality traits in strawberry [76].

This study also demonstrated the need for a training population to fully capture variation of the target population in order to maximize prediction accuracy. The single year analysis showed that although a model built using a single year of data could be used accurately to

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

predict individuals evaluated in that year, it could not be easily extrapolated to individuals whose genetic values were outside the distribution of the training data (Table 2, Additional file 2). The GBLUP approach relies on information from relatives to improve the accuracy of the estimates [77]. Because there were often sparse observations for a single year, sampling error biased the single-year estimates and resulted in models that fit the data within each year, but not across years. These effects were likely exacerbated with wild accessions, distantly related cultivars and derivatives from both groups. However, the true pairwise genetic covariance between the distantly related germplasm is estimated with less reliability with the realized relationship matrix than more closely related germplasm [78].

Genetic architecture of focus traits in sweet cherry

This study confirmed the extensive opportunity in North American sweet cherry germplasm for genetic improvement of the phenological traits of harvest timing and, to lesser extent, bloom timing. Previous QTL studies for fruit maturity date across several *Prunus* species determined bloom timing and harvest timing to be highly heritable with a large-effect QTL on LG4 [79]. Our findings also demonstrate the large broad-sense heritability for these traits – reaching a ceiling of 0.83 for harvest time and 0.65 for bloom date (Fig. 2). There appears to be little advantage to using GDD to Julian days, since pairs of phenological traits for bloom and harvest timing displayed highly similar genetic architecture and predictive accuracy. The data were all gathered from a single location, in which GDD did not vary dramatically during the years of evaluation. This may explain why GDD did not improve the model predictive ability over Julian days (Fig. 2, Table 2). Bloom timing has become increasingly important as a trait relevant to productivity, since variable climatic patterns in temperate regions can result in earlier flowering and an increased risk of floral freeze damage [80]. Furthermore, since sweet cherries are a fresh market product that is subject to rapid postharvest deterioration, it is crucial to for sweet cherry breeders and producers to understand the expected time frame for fruit maturation [79]. These results may help sweet cherry breeders identify the best parents in order to target a harvest timing window.

Moderate prospects were observed for genetic improvement of pedicel-fruit retention force ($h^2 = 0.20$, $H^2 = 0.46$, Table 2), where a low PFRF value is sought for mechanical harvest systems. Positive correlations observed between PFRF and fruit dimensions, fruit weight, and firmness (Table 3) contrasted with findings by Zhao et al. [73], in which PFRF was largely uncorrelated with firmness, fruit diameter, or fruit length. However,

that study was smaller in scope, using only 30 named cultivars and 26 unselected F1 progeny.

The potential for genetic gain in fruit dimensions and fruit weight, two highly correlated measurements of fruit size, was perhaps the highest among all focus traits due to large additive and dominance effects (Table 2). These results are consistent with previous sweet cherry studies that showed high correlations between fruit size measurements and high H^2 [18, 81–83]. In those studies, six putative QTLs influencing fruit size in cherry were identified and together accounted for 76–88% of the phenotypic variance. Because fruit weight was highly correlated with fruit dimensions in the present study (Table 3, Fig. 4) and can be evaluated rapidly, we considered it an effective proxy for fruit dimensions and general fruit size.

The high broad-sense heritability for firmness (0.77) (Table 2) was consistent with estimates from a study conducted on a biparental population in which H^2 was estimated at 0.78 to 0.85 [83]. In our study, the moderate positive correlations ($r = 0.51$) between fruit firmness and fruit dimensions among the unreleased progeny suggests genetic linkage among loci influencing these traits. This outcome was in contrast to that of a multi-year QTL study, in which the Pearson correlations between fruit firmness and fruit weight ranged from -0.64 to -0.67 for Regina \times Lapins and -0.40 to -0.15 for Regina \times Garnet F1 families [83]. Those correlations are likely due to unique genetic linkage in Regina, Garnet and Lapins. The correlations reported here may have also been influenced by the 77 progeny derived from the three wild parents: MIM 17, MIM 23, and NY54. These individuals all had high SSC, small fruit size, and low fruit firmness in their estimated genetic values relative to the population mean (Additional file 7).

Expectations for genetic improvement in SSC were moderately positive. Narrow-sense heritability 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 [84], 0.72 for peach [85], 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 [86, 87]. 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 [88]. 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. [88] 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 [89, 90]. 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*) [91–93].

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, 71, 94–99]. 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 [63, 100]. 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 [71, 101]. Allele fixation is challenging in predominantly heterozygous crop such as sweet cherry whose high heterozygosity is maintained by a self-incompatibility mechanism [102]. 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 Fam16 have 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 [103] and has been widely used for many crops including long-lived perennial trees [40, 45, 70, 104–106].

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, 105, 107–110]. 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 the breeding phases, genetic values obtained early in the breeding process will not change due to recombination. Knowing the genetic potential of an individual will help cherry breeders discard low-performing individuals and advance selections to the next phase with strong evidence. Knowledge of the genetic potential of a candidate selection may enable breeders to skip a cycle of field evaluation, thus increasing the pace of cultivar release and saving resources that can be diverted elsewhere. Given the lengthy time period for developing a sweet cherry cultivar, shortening this process can represent considerable savings.

Conclusions

The genetic values and the improved understanding of the genetic architecture of important traits in sweet cherry obtained from this multi-year data set of a large pedigree-connected population represent a clear opportunity for genetic improvement. This application – estimating genetic variance components and genome-estimated genetic values – extended the original purpose

of the RosBREED sweet cherry Crop Reference Set: QTL detection and validation. We plan to update the genetic models by incorporating new phenotypic data on existing germplasm, adding new individuals, and expanding the genome-wide SNP set for denser genome coverage. Further research is needed to validate the accuracy of genetic predictions on an independent data set and to understand the extent of genotype-by-environment effects for obtained breeding values and genetic values.

Additional files

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. (XLSX 17 kb)

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). (XLSX 11 kb)

Additional file 3: Histogram of the diagonals and off-diagonals from the additive relationship matrix. (PNG 268 kb)

Additional file 4: Variance component estimates and standard errors for all RosBREED sweet cherry traits. (CSV 5 kb)

Additional file 5: Variance component percentages for all RosBREED sweet cherry traits. (CSV 3 kb)

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. (PNG 9522 kb)

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. An online app for exploring the breeding and genetic values presented in Additional file 7 is available at: <https://www.rosbreed.org/breeding/cherry/gebv>. (XLSX 791 kb)

Acknowledgements

We gratefully acknowledge contributions of an anonymous reviewer via the Peerage of Science review process, and Laura Giradeau for manuscript editing.

Funding

This work was supported by the USDA National Institute of Food and Agriculture (NIFA) – Specialty Crop Research Initiative project, “RosBREED: Combining disease resistance with horticultural quality in new rosaceous cultivars” (grant number 2014–51181–22378) and USDA NIFA Hatch project 1014919. The funding agency had no role in the design of the study, the collection, analysis, and interpretation of data, or in writing the manuscript.

Availability of data and materials

All data used are available in the Genome Database for Rosaceae [111]. Phenotypic data can be found by filtering for “Sweet Cherry (RosBREED)” at https://www.rosaceae.org/search/quantitative_traits. Genotypic data can be found by filtering for “Sweet_Cherry_CRS_SNP_genotyping” at https://www.rosaceae.org/search/snp_genotype.

Authors’ contributions

JP conducted the bulk of the data analysis and results interpretation. CH advised on model building and interpretation. LC, JP, and CP curated SNP data and YZ gathered phenotypic data. AI and CP contributed to study design and interpretation. All authors contributed to manuscript preparation and read and approved the final version.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Horticulture, Washington State University, Pullman, WA 99164-6414, USA. ²Centre for Horticultural Science, Queensland Alliance for Agriculture and Food Innovation University of Queensland, Brisbane, Australia. ³Department of Horticulture, Michigan State University, East Lansing, MI 48824-1325, USA. ⁴Plants for Human Health Institute, North Carolina State University, Kannapolis, NC 28081, USA.

Received: 13 December 2017 Accepted: 22 March 2018

Published online: 10 April 2018

References

1. FAOSTAT Data [Internet]. FAOSTAT. [cited 2018 Mar 27]. Available from: <http://www.fao.org/faostat/en/#data>.
2. National Statistics for Cherry [Internet]. NASS. [cited 2018 Mar 27]. Available from: https://quickstats.nass.usda.gov/results/A8988197-374E-3950-BA97-9CBCEA511544?pivot=short_desc.
3. Sweet Cherry Production Up 36 Percent [Internet]. NASS. [cited 2018 Mar 27]. 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. *J Agric Resour Econ*. 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, Metheny 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 2018 Mar 27]. 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*. 2009:859.
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.

17. Iezzoni A. Variance components and sampling procedures for fruit size and quality in sour cherry. *Hortscience*. 1986;21:1040–2.
18. Rosyara U, Bink MAM, van de Weg E, Zhang G, Wang D, Sebolt A, et al. Fruit size QTL identification and the prediction of parental QTL genotypes and breeding values in multiple pedigreed populations of sweet cherry. *Mol Breed*. 2013;32:875–87.
19. Srivastava K, Verma M, Ahmad N, Ravi S, Ahmad S. Genetic variability and divergence analysis in sweet cherry (*Prunus avium* L.). *Indian J Hort*. 2014;71:156–61.
20. Besford RT, Hand P, Peppitt SD, Richardson CM, Thomas B. Phase change in *Prunus avium*: differences between juvenile and mature shoots identified by 2-dimensional protein separation and *in vitro* translation of mRNA. *J Plant Physiol*. 1996;147:534–8.
21. Henderson CR. Sire evaluation and genetic trends. *J Anim Sci*. 1973;1973:10–41.
22. Lynch M, Walsh B. Genetic and analysis of quantitative traits. Sunderland, MA: Sinauer Associates, Inc; 1998.
23. E, Silva FF, Viana JMS, Faria VR, de Resende MDV. Bayesian inference of mixed models in quantitative genetics of crop species. *Theor Appl Genet*. 2013;126:1749–61. <https://doi.org/10.1007/s00122-013-2089-6>.
24. Henderson CR. Use of relationships among sires to increase accuracy of sire evaluation. *J Dairy Sci*. 1975;58:1731–8.
25. Hayes B, Visscher P, Goddard M. Increased accuracy of artificial selection by using the realized relationship matrix. *Genet Res*. 2009;91.
26. Vitezica ZG, Varona L, Legarra A. On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics*. 2013;195:1223–30.
27. Muñoz PR, Resende MFR, Gezan SA, Resende MDV, de los Campos G, Kirt M, et al. Unraveling additive from non-additive effects using genomic relationship matrices. *Genetics*. 2014.
28. Habier D, Fernando RL, Dekkers JCM. The impact of genetic relationship information on genome-assisted breeding values. *Genetics*. 2007;177.
29. Vela-Avitua S, Meuwissen THE, Luan T, Odegard J. Accuracy of genomic selection for a sib-evaluated trait using identity-by-state and identity-by-descent relationships. *Genet Sel Evol GSE*. 2015;47:9.
30. Junqueira VS, Cardoso FF, Oliveira MM, Sollero BP, Silva FF, Lopes PS. Use of molecular markers to improve relationship information in the genetic evaluation of beef cattle tick resistance under pedigree-based models. *J Anim Breed Genet*. 2017;134:14–26.
31. Durel CE, Laurens F, Fouillet A, Lespinasse Y. Utilization of pedigree information to estimate genetic parameters from large unbalanced data sets in apple. *Theor Appl Genet*. 1998;96:1077–85.
32. de Souza VAB, Byrne DH, Taylor JF. Predicted breeding values for nine plant and fruit characteristics of 28 peach genotypes. *J Am Soc Hort Sci*. 2000;125:460–5.
33. Kouassi A, Durel C-E, Costa F, Tartarini S, van de Weg E, Evans K, et al. Estimation of genetic parameters and prediction of breeding values for apple fruit-quality traits using pedigreed plant material in Europe. *Tree Genet Genomes*. 2009;5:659–72.
34. Stephens MJ, Alspach PA, Beatson RA, Winefield C, Buck EJ. Genetic parameters and breeding for yield in red raspberry. *J Am Soc Hort Sci*. 2012;137:229–35.
35. Whitaker VM, Osorio LF, Hasing T, Gezan S. Estimation of genetic parameters for 12 fruit and vegetative traits in the University of Florida strawberry breeding population. *J Am Soc Hort Sci*. 2012;137:316–24.
36. Fresnedo-Ramírez J, Crisosto CH, Gradziel TM, Famula TR. Pedigree correction and estimation of breeding values for peach genetic improvement. *Acta Hort*. 2015:249–56.
37. Gezan SA, Osorio LF, Verma S, Whitaker VM. An experimental validation of genomic selection in octoploid strawberry. *Hortic Res*. 2017;4:16070.
38. Tancred SJ, Zeppa AG, Cooper M, Stringer JK. Heritability and patterns of inheritance of the ripening date of apples. *Hortscience*. 1995;30:325–8.
39. Hardner CM, Kumar S, Peace CM, Luby J, Evans KM. Reconstructing relationship matrices from dense SNP arrays for the prediction of genetic potential in unreplicated multilocation plantings of apple progeny. *Acta Hort*. 2016:275–82.
40. Furlani RCM, de MMLT, de RMDV, Furlani Junior E, Gonçalves P de S, Valério Filho WV, et al. Estimation of variance components and prediction of breeding values in rubber tree breeding using the REML/BLUP procedure. *Genet Mol Biol*. 2005;28:271–6.
41. Hardner CM, Healey AL, Downes G, Herberling M, Gore PL. Improving prediction accuracy and selection of open-pollinated seed-lots in *Eucalyptus dunnii* Maiden using a multivariate mixed model approach. *Ann For Sci*. 2016;73:1035–46.
42. Imai A, Kuniga T, Yoshioka T, Nonaka K, Mitani N, Fukamachi H, et al. Evaluation of the best linear unbiased prediction method for breeding values of fruit-quality traits in citrus. *Tree Genet Genomes*. 2016;12:119.
43. Minamikawa MF, Nonaka K, Kaminuma E, Kajiya-Kanegae H, Onogi A, Goto S, et al. Genome-wide association study and genomic prediction in citrus: potential of genomics-assisted breeding for fruit quality traits. *Sci Rep*. 2017;7:4721.
44. Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet*. 2008;4:e1000008.
45. Kumar S, Molloy C, Muñoz P, Daetwyler H, Chagne D, Volz R. Genome-enabled estimates of additive and nonadditive genetic variances and prediction of apple phenotypes across environments. *G3*. 2015;5:2711–8.
46. Iezzoni A, Weebadde C, Luby J, Yue C, van de Weg E, Fazio G, et al. RosBREED: enabling marker-assisted breeding in Rosaceae. *Acta Hort*. 2010;859:389–94.
47. Long L, Kaiser C. Sweet cherry rootstocks for the Pacific Northwest. Corvallis: Oregon State University; 2010 p. 1–8. Report No.: PNW 619.
48. Peace C, Luby J, Van de Weg WE, Bink M, Iezzoni A. A strategy for developing representative germplasm sets for systematic QTL validation, demonstrated for apple, peach, and sweet cherry. *Tree Genet Genomes*. 2014;10:1679–94.
49. Webster AD. Rootstock and interstock effects on deciduous fruit tree vigour, precocity, and yield productivity. *N Z J Crop Hort Sci*. 1995;23:373–82.
50. Usenik V, Fajt N, Atampar F. Effects of rootstocks and training system on growth, precocity and productivity of sweet cherry. *J Hort Sci Biotechnol*. 2006;81:153–7.
51. Whiting MD, Lang G, Ophardt D. Rootstock and training system affect sweet cherry growth, yield, and fruit quality. *Hortscience*. 2005;40:582–6.
52. Chavoshi M, Watkins C, Oraguzie B, Zhao Y, Iezzoni A, Oraguzie N. Phenotyping protocol for sweet cherry (*Prunus avium* L.) to facilitate an understanding of trait inheritance. *Am Pomol Soc*. 2014;68:125–34.
53. Washington State University. AgWeatherNet Roza Station Data [Internet]. AgWeatherNet. [cited 2018 Mar 27]. Available from: www.weather.wsu.edu.
54. McMaster G, Wilhelm W. Growing degree-days: one equation, two interpretations. *Agric For Meteorol*. 1997;87:291–300.
55. Peace C, Bassil N, Main D, Ficklin S, Rosyara UR, Stegmeir T, et al. Development and evaluation of a genome-wide 6K SNP Array for diploid sweet cherry and tetraploid sour cherry. *PLoS One*. 2012;7:e48305.
56. Cai L, Voorrips RE, van de Weg E, Peace C, Iezzoni A. Genetic structure of a QTL hotspot on chromosome 2 in sweet cherry indicates positive selection for favorable haplotypes. *Mol Breed*. 2017;37:85.
57. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet*. 2009;84:210–23.
58. Wimmer V, Albrecht T, Auinger H-J, Schoen C-C. Synbreed: a framework for the analysis of genomic prediction data using R. *Bioinformatics*. 2012;28:2086–7.
59. Butler D, Cullis B, Gilmour A, Gogel B. Analysis of mixed models for S language environments: ASReml-R reference manual (version 3). The State of Queensland, Department of Primary Industries and Fisheries. 2009.
60. R Development Core Team. R: A language and environment for statistical computing [Internet]. 2011. Available from: <http://www.R-project.org/>.
61. Endelman JB. Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome*. 2011;4:25–255.
62. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci*. 2008;91:4414–23.
63. Su G, Christensen OF, Ostensen T, Henryon M, Lund MS. Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. *PLoS One*. 2012;7:e45293.
64. Piaskowski J. Genomic dominance relationship matrix [internet]. 2017. Available from: <https://github.com/jpiaskowski/Genomic-Dominance-Relationship-Matrix>.
65. Brien C. asremIPlus [Internet]. 2016. Available from: <https://cran.r-project.org/web/packages/asremIPlus/asremIPlus.pdf>.
66. Gabriel K. The biplot graphic display of matrices with application to principal component analysis. *Biometrika*. 1971;58:453–67.
67. Rodriguez-Almeida FA, Van Vleck LD, Willham RL, Northcutt SL. Estimation of non-additive genetic variances in three synthetic lines of beef cattle using an animal model. *J Anim Sci*. 1995;73:1002–11.

68. Pante M, Gjerde B, McMillan I, Misztal I. Estimation of additive and dominance genetic variances for body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. *Aquaculture*. 2002;204:383–92.
69. Gallardo JA, Lhorente JP, Neira R. The consequences of including non-additive effects on the genetic evaluation of harvest body weight in Coho salmon (*Oncorhynchus kisutch*). *Genet Sel Evol*. 2010;42:19.
70. Gamal El-Dien O, Ratcliffe B, Klapste J, Porth I, Chen C, El-Kassaby YA. Implementation of the realized genomic relationship matrix to open-pollinated white spruce family testing for disentangling additive from nonadditive genetic effects. *G3*. 2016;6:743–53.
71. Wolfe MD, Kulakow P, Rabbi IY, Jannink J-L. Marker-based estimates reveal significant non-additive effects in clonally propagated cassava (*Manihot esculenta*): implications for the prediction of total genetic value and the selection of varieties. *G3*. 2016.
72. Ceballos H, Kawuki RS, Gracen VE, Yencho GC, Hershey CH. Conventional breeding, marker-assisted selection, genomic selection and inbreeding in clonally propagated crops: a case study for cassava. *Theor Appl Genet*. 2015;128:1647–67.
73. Zhao Y, Athanson B, Whiting M, Oraguzie N. Pedicel-fruit retention force in sweet cherry (*Prunus avium* L.) varies with genotype and year. *Sci Hortic*. 2013;150:135–41.
74. Schmitz CA, Clark MD, Luby JJ, Bradeen JM, Guan Y, Evans K, et al. Fruit texture phenotypes of the RosBREED U.S. apple reference germplasm set. *Hortscience*. 2013;48:296–303.
75. Cantin CM, Gogorcena Y, Moreno MÁ. Analysis of phenotypic variation of sugar profile in different peach and nectarine [*Prunus persica* (L.) Batsch] breeding progenies. *J Sci Food Agric*. 2009;89:1909–17.
76. Mathey MM, Mookerjee S, Mahoney LL, Gündüz K, Rosyara U, Hancock JF, et al. Genotype by environment interactions and combining ability for strawberry families grown in diverse environments. *Euphytica*. 2017;213:112.
77. Clark S, van de Werf J. Genomic best unbiased linear prediction (gBLUP) for the estimation of genomic breeding values. In: Gondro C, van de Werf J, Hayes B, editors. *Genome-wide Assoc stud genomic predict*. Springer; 2013. p. 321–30.
78. Heslot N, Jannink J-L, Sorrells ME. Perspectives for genomic selection applications and research in plants. *Crop Sci*. 2015;55(1):12.
79. Dirlewanger E, Quero-García J, Le Dantec L, Lambert P, Ruiz D, Dondini L, et al. Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three *Prunus* species: peach, apricot and sweet cherry. *Heredity*. 2012;109:280–92.
80. Kim Y, Kimball JS, Didan K, Henebry GM. Response of vegetation growth and productivity to spring climate indicators in the conterminous United States derived from satellite remote sensing data fusion. *Agric For Meteorol*. 2014;194:132–43.
81. Zhang G, Sebolt A, Sooriyapathirana S, Wang D, Bink M, Olmstead J, et al. Fruit size QTL analysis of an F1 population derived from a cross between a domesticated sweet cherry cultivar and a wild forest sweet cherry. *Tree Genet Genomes*. 2010;6:25–36.
82. De Franceschi P, Stegmeir T, Cabrera A, van der Knaap E, Rosyara UR, Sebolt AM, et al. Cell number regulator genes in *Prunus* provide candidate genes for the control of fruit size in sweet and sour cherry. *Mol Breed*. 2013;32:311–26.
83. Campoy JA, Le Dantec L, Barreneche T, Dirlewanger E, Quero-García J. New insights into fruit firmness and weight control in sweet cherry. *Plant Mol Biol Report*. 2015;33:783–96.
84. Bassi D, Bartolozzi F, Muzzi E. Patterns and heritability of carboxylic acids and soluble sugars in fruits of apricot (*Prunus armeniaca* L.). *Plant Breed*. 1996;115:67–70.
85. Brooks SJ, Moore JN, Quantitative MJB. Qualitative changes in sugar content of peach genotypes [*Prunus persica* (L.) Batsch]. *J Am Soc Hortic Sci*. 1993;118:97–100.
86. Genard M, Lescouret F, Gomez L, Habib R. Changes in fruit sugar concentrations in response to assimilate supply, metabolism and dilution: a modeling approach applied to peach fruit (*Prunus persica*). *Tree Physiol*. 2003;23:373–85.
87. Morandi B, Corelli Grappadelli L, Rieger M, Lo Bianco R. Carbohydrate availability affects growth and metabolism in peach fruit. *Physiol Plant*. 2008;133:229–41.
88. Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, et al. Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch). *Theor Appl Genet*. 1999;98:18–31.
89. Olmstead J, Lang GA. Leaf disk assay for screening sweet cherry genotypes for susceptibility to powdery mildew. *Hortscience*. 2000;35:274–7.
90. Olmstead J, Lang G, Grove G. Inheritance of powdery mildew resistance in sweet cherry. *Hortscience*. 2001;36:337–40.
91. Wilfert L, Schmid-Hempel P. The genetic architecture of susceptibility to parasites. *BMC Evol Biol* 2008;8:187–187.
92. Divya B, Biswas A, Robin S, Rabindran R, Joel AJ. Gene interactions and genetics of blast resistance and yield attributes in rice (*Oryza sativa* L.). *J Genet*. 2014;93:415–24.
93. Zhang J, Singh A, Mueller DS, Singh AK. Genome-wide association and epistasis studies unravel the genetic architecture of sudden death syndrome resistance in soybean. *Plant J Cell Mol Biol*. 2015;84:1124–36.
94. Cach NT, Perez JC, Lenis JJ, Calle F, Morante N, Ceballos H. Epistasis in the expression of relevant traits in cassava (*Manihot esculenta* Crantz) for subhumid conditions. *J Hered*. 2005;96:586–92.
95. Oakey H, Verbyla A, Pitchford W, Cullis B, Kuchel H. Joint modeling of additive and non-additive genetic line effects in single field trials. *Theor Appl Genet*. 2006;113:809–19.
96. Bai W, Zhang H, Zhang Z, Teng F, Wang L, Tao Y, et al. The evidence for non-additive effect as the main genetic component of plant height and ear height in maize using introgression line populations. *Plant Breed*. 2010;129:376–84.
97. Wang D, Salah El-Basyoni I, Stephen Baenziger P, Crossa J, Eskridge KM, Dweikat I. Prediction of genetic values of quantitative traits with epistatic effects in plant breeding populations. *Heredity*. 2012;109:313–9.
98. Dudley JW, Johnson GR. Epistatic models and pre-selection of markers improve prediction of performance in corn. *Mol Breed*. 2013;32:585–93.
99. Nazarian A, Gezan SA. Integrating nonadditive genomic relationship matrices into the study of genetic architecture of complex traits. *J Hered*. 2016;107:153–62.
100. Cockerham CC. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics*. 1954;39:859–82.
101. Holland J. Epistasis and plant breeding. In: Janick J, editor. *Plant Breed Rev*. Oxford, UK: John Wiley & Sons, Inc.; 2010. p. 27–92.
102. Tao R, lezzoni AF. The S-RNase-based gametophytic self-incompatibility system in *Prunus* exhibits distinct genetic and molecular features. *Sci Hortic*. 2010;124:423–33.
103. Heffner EL, Sorrells ME, Jannink J-L. Genomic selection for crop improvement. *Crop Sci*. 2009;49(1):12.
104. Ratcliffe B, El-Dien OG, Klapste J, Porth I, Chen C, Jaquish B, et al. A comparison of genomic selection models across time in interior spruce (*Picea engelmannii* x *glauca*) using unordered SNP imputation methods. *Heredity*. 2015;115:547–55.
105. Muranty H, Troglio M, Sadok IB, Rifai MA, Auwerkerken A, Banchi E, et al. Accuracy and responses of genomic selection on key traits in apple breeding. *Hortic Res*. 2015;2:15060.
106. Biscarini F, Nazzicari N, Bink M, Arús P, Aranzana MJ, Verde I, et al. Genome-enabled predictions for fruit weight and quality from repeated records in European peach progenies. *BMC Genomics*. 2017;18:432.
107. Garcia MR, Carbonell EA, Asins MJ. QTL analysis of yield and seed number in Citrus. *Theor Appl Genet*. 2000;101:487–93.
108. Resende MFR, Muñoz P, Acosta JJ, Peter GF, Davis JM, Grattapaglia D, et al. Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *New Phytol*. 2012;193:617–24.
109. Resende RT, Resende MDV, Silva FF, Azevedo CF, Takahashi EK, Silva-Junior OB, et al. Assessing the expected response to genomic selection of individuals and families in *Eucalyptus* breeding with an additive-dominant model. *Heredity*. 2017;119:245–55.
110. Marulanda JJ, Mi X, Melchinger AE, Xu J-L, Würschum T, Longin CFH. Optimum breeding strategies using genomic selection for hybrid breeding in wheat, maize, rye, barley, rice and triticale. *Theor Appl Genet*. 2016;129:1901–13.
111. Jung S, Ficklin SP, Lee T, Cheng C-H, Blenda A, Zheng P, et al. The genome database for Rosaceae (GDR): year 10 update. *Nucleic Acids Res*. 2014;42:D1237–44.