DOI: 10.1002/tpg2.70018

ORIGINAL ARTICLE

Genetic analysis of predicted vegetative biomass and biomass-related traits from digital phenotyping of strawberry

Amr Abd-Elrahman² | Xu Wang⁴

Cheryl Dalid¹ Caiwang Zheng² Luis Osorio¹ Sujeet Verma³ Vance M. Whitaker¹ 💿

¹Horticultural Sciences Department, IFAS Gulf Coast Research and Education Center, University of Florida, Wimauma, Florida, USA

²School of Forest Resources and Conservation Geomatics, IFAS Gulf Coast Research and Education Center, University of Florida, Plant City, Florida, USA

³Fall Creek Farm and Nursery Inc., Lowell, Oregon, USA

⁴Agricultural and Biological Engineering Department, IFAS Gulf Coast Research and Education Center, University of Florida, Wimauma, Florida, USA

Correspondence

Vance M. Whitaker, Horticultural Sciences Department, IFAS Gulf Coast Research and Education Center, University of Florida, Wimauma, FL 33598, USA. Email: vwhitaker@ufl.edu

Assigned to Associate Editor Qingyi Yu.

Funding information

Florida Agricultural Experiment Station; Florida Strawberry Research and Education Foundation

Abstract

High-throughput digital phenotyping (DP) has been widely explored in plant breeding to assess large numbers of genotypes with minimal manual labor and reduced cost and time. DP platforms using high-resolution images captured by drones and tractorbased platforms have recently allowed the University of Florida strawberry (Fragaria \times ananassa) breeding program to assess vegetative biomass at scale. Biomass has not previously been explored in a strawberry breeding context due to the labor required and the need to destroy the plant. This study aims to understand the genetic basis of predicted vegetative biomass and biomass-related traits and to chart a path for the combined use of DP and genomics in strawberry breeding. Aboveground dry vegetative biomass was estimated by adapting a previously published model using ground-truth data on a subset of breeding germplasm. High-resolution images were collected on clonally replicated trials at different time points during the fruiting season. There was moderate to high heritability ($h^2 = 0.26-0.56$) for predicted vegetative biomass, and genetic correlations between vegetative biomass and marketable yield were mostly positive ($r_G = -0.13 - 0.47$). Fruit yield traits scaled on a vegetative biomass basis also had moderate to high heritability ($h^2 = 0.25-0.64$). This suggests that vegetative biomass can be decreased or increased through selection, and that marketable fruit yield can be improved without simultaneously increasing plant size. No consistent marker-trait associations were discovered via genome-wide association studies. On the other hand, predictive abilities from genomic selection ranged from 0.15 to 0.46 across traits and years, suggesting that genomic prediction will be an effective breeding tool for vegetative biomass in strawberry.

_____ This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). The Plant Genome published by Wiley Periodicals LLC on behalf of Crop Science Society of America.

Abbreviations: ACCB, accumulated predicted vegetative biomass; AWT, average fruit weight; AWTAC, average fruit weight on an accumulated predicted vegetative biomass basis; AWTB, average fruit weight on a predicted biomass basis; BIO, predicted vegetative biomass; DP, digital phenotyping; EMY, early marketable yield; FRM, firmness; GS, genomic selection; GWAS, genome-wide association studies; SNP, single nucleotide polymorphism; SSC, soluble solids content; TC, total culls; TMY, total marketable yield; UAV, unmanned aerial vehicle; UF, University of Florida.

Plain Language Summary

Vegetative biomass is the total mass of the leaves and stems of a plant that roughly reflects the overall size of the plant. Strawberries are hand harvested; therefore, it is important to select moderately sized plants for easier access to fruits. However, larger plants tend to have more fruit yield. Vegetative biomass is a trait that our breeding program has not measured before because it requires destruction of the plant. Now, we have estimated biomass throughout multiple seasons using cameras mounted on tractors and drones in a "digital phenotyping" approach. Combining digital, field, and DNA data, we show that estimated biomass traits are controlled by genetics and can be accurately predicted in future breeding populations with the aid of DNA markers.

1 | INTRODUCTION

Cultivated strawberry (*Fragaria* × *ananassa*) is the most widely consumed fruit crop in the world with an annual production of 9.5 million metric ton in 2022 (FAO-UN, 2022). Although the production of strawberries increased by 1168% from 1961 to 2022 due to an increase in acreage, only a 200% increase in yield was observed (FAO-UN, 2022). To meet the growing global demands for strawberries, it is vital for strawberry breeding programs to develop high-yielding cultivars while also improving ease of harvest, fruit quality, and disease resistance.

In the United States, most strawberries are produced in California and Florida. As of 2022, California produced 89.1% of strawberries in the United States, while Florida, known for its peak strawberry production during winter, produced 10.9% (USDA-NASS, 2023). The University of Florida (UF) breeding program develops cultivars that are adapted for fruit production during the winter season in west-central Florida. Currently, strawberry fruit production in west-central Florida begins in late November and peaks during February and March as the weather warms and the plants continue to accumulate crowns and leaves (Whitaker & Fan, 2020; Wu et al., 2015). Development of early yielding strawberries is an important breeding objective at UF to help Florida farmers take advantage of high prices from late November until the end of January.

Harvest efficiency is also an important breeding objective since harvesting is a labor-intensive and costly activity that represents about 57% of the total cost of strawberry production (Guan et al., 2018). Selecting for a moderate plant size helps increase efficiency of manual harvesting and allows strawberry plants to fit within desired in-row spacings. On the other hand, per plant yield tends to be positively correlated to leaf area and plant size (Agehara, 2021). Plant biomass is the total mass of a plant, including aboveground plant parts such as stems, leaves, flowers, and fruits, as well as the root system (Poorter & Nagel, 2000). Plant vegetative biomass, which is our focus here, refers to only leaves and stems. High vegetative biomass is associated with the potential for higher yields across plant species (Poorter et al., 2009). With an increase in leaf biomass, higher photosynthesis is expected, which in turn increases biomass accumulation that can lead to greater yield (Burgess et al., 2023; Poorter et al., 2009; Zhu et al., 2010). However, direct measurement of vegetative biomass is labor-intensive, time consuming, and requires the destruction of the plant. Implementation of high-throughput phenotyping approaches, including digital phenotyping (DP), can help address these limitations.

Recent advances in DP technologies have enabled rapid and detailed extraction of plant traits on a large scale. Imagebased DP using ground-based and aerial systems are the two most common approaches (Guan et al., 2020; Krause et al., 2020; Pérez-Ruiz et al., 2020; H. Zheng et al., 2019). In strawberry, Guan et al. (2020) used high-resolution imagery obtained from a ground-based imaging system mounted on a tractor to develop a model predicting vegetative biomass with an accuracy (R^2) of 0.79. This previous work has provided an opportunity to estimate vegetative biomass at scale in strawberry and thus to study the genetics of this trait.

Because the ability to measure vegetative biomass at scale is a new development in strawberry, the genetics of this trait have not yet been explored. Genome-wide association studies (GWAS) represent an effective method to discover genetic variants associated with traits, potentially leading to the development of tools for marker-assisted selection (Ibrahim et al., 2020). However, this approach is most beneficial when the trait is controlled by one or few large-effect loci. For complex traits, a genomic selection (GS) strategy has been effective for increasing genetic gains in plants (Crossa et al., 2010; de Bem Oliveira et al., 2019; Gezan et al., 2017; Merrick & Carter, 2021; Pincot et al., 2020; Tessema et al., 2020). GS uses genome-wide marker data to predict the performance of untested individuals (Meuwissen et al., 2001), a successful approach in various breeding programs, including in strawberry (Crain et al., 2018; O'Connor et al., 2021; Osorio et al., 2021; Peixoto et al., 2024).

To date, vegetative biomass has not been measured or studied at scale in any strawberry breeding program due to the lack of a high-throughput and nondestructive method. Now that we have a DP method to estimate vegetative biomass, we are able to explore the genetic basis of this trait in relation to important fruit yield and quality traits. Here, we have combined predicted vegetative biomass data with genomic information from the UF strawberry breeding program to estimate heritabilities and genetic correlations for vegetative biomass and yield-related traits. We also performed GWAS to examine the genetic architecture of vegetative biomass. Lastly, we evaluated the potential of GS, demonstrating the combined use of DP and genomics in strawberry breeding.

2 | MATERIALS AND METHODS

2.1 | Plant material and field trials

The genotypes used in this study represent the elite strawberry germplasm of the Strawberry Breeding and Genetics Program of the UF including released cultivars and advanced selections. These were planted in replicated yield trials during the 2017–2018, 2018–2019, 2020–2021, and 2021–2022 seasons on October 9, October 10, October 7, and October 6, respectively, at the UF/IFAS Gulf Coast Research and Education Center, in Balm, FL (27° 45′ 37.98″ N, 82° 13′ 32.49″ W) (Figure 1). Other data from these field trials were also utilized in previous studies (Fan & Whitaker, 2024; Fan et al., 2024; Osorio et al., 2021).

Bare-root clonal replicates of 1748 different entries across the years were planted in a randomized complete block design with five replications that consists of two beds per replication. The 1748 phenotypes represented 1460 unique individuals including advanced selections and varieties that were repeated across trials (Table 1). Each bed was divided into five to eight plots with one cultivar check per plot to account for variation within the bed (columns). Each replication consisted of a single plant (clonal replicate) per genotype. Overlapping genotypes were included across years. The trial site was prepared and maintained following the standard strawberry growing practices in west-central Florida. Fertilization, weed management, and pest and disease control varied among different seasons according to environmental conditions.

2.2 | Phenotypic data collection

2.2.1 | Yield and fruit quality

Weekly yield and fruit quality data were collected from mid-November until mid-March for all trials. During each harvest

Core Ideas

- University of Florida strawberry breeding program nondestructively measured biomass through digital phenotyping (DP) methods.
- These measurements allowed us to examine the inheritance of biomass and balance selection for yield and plant size.
- Predictive abilities from genomic selection were moderate and will allow effective predictive breeding for biomass.
- Improvement of DP methods for biomass estimation could also improve selection accuracy for biomass in the future.

week, all ripe fruits per plant were removed and evaluated for six yield and fruit quality traits that were divided into two timepoints, early and total season. Early season is the harvest period between late November and January, while the total season is the harvest period between late November and March. The traits evaluated were early marketable yield (EMY; g/plant), total marketable yield (TMY; g/plant), average fruit weight (AWT; g/fruit), total culls (TC; number/plant), soluble solids content (SSC; %), and firmness (FRM; 1-5 score). All marketable fruit were considered as EMY if harvested before the first day of February. TMY was calculated as the marketable fruit collected until the first week of March. AWT in grams was estimated as the TMY divided by the number of marketable fruits. TC, or unmarketable fruit, were counted and expressed as a proportion of the total number of fruits per plant. SSC was measured five times during the season within each trial and was calculated as the mean of all measurements. At each measurement, one ripe fruit from each plant was squeezed by hand onto a handheld digital refractometer. FRM was measured as a score from a scale of 1-5 (Osorio et al., 2021; Whitaker et al., 2012).

2.2.2 | Digital phenotyping

A ground-based imaging platform described by Guan et al. (2020) and Abd-Elrahman et al. (2020) was utilized to collect high-resolution RGB and NIR images from the field trials during the 2017–2018 and 2018–2019. Image data were collected at two timepoints that represent the end points of the early season and total season. Canopy size metrics such as plant canopy area, volume, average height, and standard deviation of height were computed using a model used for all datasets acquired throughout the season as described in Abd-Elrahman et al. (2020).



FIGURE 1 Location of the phenomics trial used to develop the model to predict vegetative biomass in relation to the advanced selection trial where vegetative biomass was predicted and used in this study.

TABLE 1 Number of tested genotypes across years.

Year	2017	2018	2020	2021	N^{a}
2017	390 ^b	64 ^c	19	9	394
2018	64	413	29	17	445
2020	19	29	413	102	443
2021	9	17	102	444	466

Note: The numbers in bold shows the number of unique genotypes that were included in the analysis.

^aTotal number of tested genotypes across years (total N = 1748).

^bNumber of analyzed genotypes (diagonal, 1460 unique genotypes).

^cNumber of common genotypes within years (off-diagonal).

In the 2020–2021 and 2021–2022 seasons, an unmanned aerial vehicle (UAV) platform was developed to replace the ground-based imaging platform for image acquisition. A MicaSense RedEdge-M multispectral camera with a down-

welling light sensor (DLS) mounted on a Da-Jiang Innovations (DJI) Inspire 2 drone was used in this study. Since the UAV is easy to operate and control remotely, it was used to collect image data at more time points (13 and 26 for the 2020–2021 and 2021–2022 seasons, respectively) to track the growth of strawberry plants over time. Canopy size metrics and related vegetation indexes were extracted from the UAV imagery for vegetative biomass prediction. Image data processing and comparison between the UAV and groundbased platform in this study were discussed in C. Zheng et al. (2022a).

2.2.3 | Vegetative biomass prediction

Biomass data were estimated using a modified regression model previously published by Guan et al. (2020). Canopy size metrics utilized in the model were generated from high-resolution images acquired from genotypes where dry biomass using destructive methods was also measured to serve as the ground truth. This experiment was done at the UF/IFAS Gulf Coast Research and Education Center, in Balm, FL, on a subset of the UF breeding germplasm called the "phenomics trial" (Figure 1) following the formula:

Vegetative biomass (g) = $-22.95 + (156.72 \times \text{Canopy area})$

 $+(266.50 \times \text{Mean canopy height})$

+ $(0.8 \times \text{Weeks after planting})$.

Vegetative biomass that represents the early season was labeled as EBIO. Predicted vegetative biomass at the end of the season was labeled as BIO. Fruit yield data such as EMY, TMY, and AWT were then scaled on a biomass basis by dividing EMY, TMY, and AWT by the predicted vegetative biomass, which resulted in new traits of interest, namely, early AWT on a predicted biomass basis (EAWTB), AWT on a predicted biomass basis (AWTB), early marketable yield on a predicted biomass basis (EMYB), and total marketable yield on a predicted biomass basis (TMYB).

2.2.4 | Accumulated predicted vegetative biomass

In the 2020–2022 seasons, vegetative biomass was estimated using machine learning models previously proposed by C. Zheng et al. (2022a, 2022b) using image-derived canopy size metrics and several vegetation indices extracted from a MicaSense RedEdge-M multispectral camera and a DLS mounted on a DJI Inspire 2 drone operated autonomously, with field flight parameters pre-set using the Litchi flight planning software to cover the whole strawberry experiment area, which were collected 13 times during 2020–2021 and 26 times during 2021–2022. Using the estimated vegetative biomass data at multiple time points, the logistic function was adopted to fit the growth curve of each strawberry plant, which contains three parameters, as shown below:

$$y = \frac{k}{1 + e^{\rho_0 + \gamma x}},$$

where *x* represents the days after planting, *y* represents the predicted dry biomass data, and there are three coefficients (k, ρ_0 , and γ). The logistic model can be dated back to the work of Verhulst (1838), who proposed a self-limiting model of population growth with a function or curve representing an S-shaped curve (Kawano et al., 2020).

As shown in Figure 2, the fitted logistic function can reflect the growth trend of strawberry plants. The accumulated



FIGURE 2 An example of strawberry plant growth curve where the solid red line is the fitted logistic curve, and the discrete blue points are the biomass estimated using machine learning methods. The green shaded area represents the accumulated biomass during the strawberry winter growing season, spanning from the initial to the final days of data collection.

biomass was defined as the sum of biomass between the first and last days of data collection during the strawberry winter growing season.

During the 2020–2021 and 2021–2022 seasons, the accumulated predicted vegetative biomass (ACCB) was recorded, and AWT as well as TMY were scaled on an ACCB basis (AWT/ACCB and TMY/ACCB, respectively) generating two new traits, AWT on an accumulated predicted vegetative biomass basis (AWTAC) and TMYAC.

2.3 | SNP genotyping

Two sets of single nucleotide polymorphism (SNP) data were utilized in this study. First, a reference population of 1228 individuals out of a total 1297 genotypes was genotyped using common markers across the Affymetrix Axiom IStraw90 (Bassil et al., 2015) and IStraw35 arrays (Verma et al., 2017). Quality control was implemented on a total of 14,332 segregating SNP markers in which SNPs with minor allele frequency (MAF) < 0.05, and missing marker data > 0.05 were removed, leaving a total of 5300 markers for the analyses. These markers were used for the 2017–2018 and 2018–2019 seasons. Missing values for each of the markers were imputed based on average allele frequency. Genotyping of the individuals from seasons 2020–2021 and 2021–2022 was performed utilizing 8749 molecular markers from the newer FanaSNP 50K array (Hardigan et al., 2020), and a similar quality control was implemented. Due to the difference in the marker platforms used, and since different DP methods were used in the first 2 years versus the last 2 years, we took the conservative approach of conducting two analyses, one for 2017–2018 and 2018–2019 seasons and another for 2020–2021 and 2021–2022 seasons, as opposed to combining data across 4 years.

2.4 | Genetic parameter estimation

The estimation of the genetic parameters from the advanced selection trials and multi-year trials followed the two-stage approach proposed by Smith, Cullis, and Thompson (2001) and implemented in the ASRtriala software (Gezan et al., 2022). These analyses were based on multiple linear mixed models, including the modeling and/or not of spatial model terms, row, column, autoregressive correlated errors, or independent residual units. All spatially adjusted phenotypic data used in this study are available in Table S1.

In the first-stage analysis, field observations of the single advanced trials were analyzed according to the following general linear mixed model:

$$y_{ijkl} = \mu + B_i + c(B)_{ij} + r(B)_{ik} + g_l + e_{ijkl},$$

where μ is the general mean, B_i is the fixed replication effect, $c(B)_{ij}$ is the column within replication random effect, $r(B)_{ik}$ is the random row within replication effect, g_l is the random clonal effect, and e_{ijkl} is the residual random effect of the l clone in the j column and k row of the i replication. The residual effect (R) for the spatial models of the t trial fitted the autoregressive variance-covariance matrix of order 1 of the form:

$$R_t = \sigma_t^2 \sum_{c_t} (\rho_{ct}) \otimes \sum_{r_t} (\rho_{rt})$$

where $\sum_{c_t} (\rho_{c_t})$ and $\sum_{r_t} (\rho_{r_t})$ are spatial correlation structures in the column and row directions, respectively.

There were 14 models tested for each trait within the early and total periods of each year, including the presence or absence of some model factors (Table S2). The best-fitted model for each trait, period, and year was selected based on the heritability of the prediction error variance (h_{PEV}^2) , a goodness of fit criterion that is commonly used to compare different models with and/or without spatial factors (Table S3). These single-stage analyses provide variance components for the random effects, genetic predictions for all genotypes and their corresponding weights (Gezan et al., 2022).

The genotype by year interaction $(G \times Y)$ of EBIO, BIO, and the scaled traits EAWTB, AWTB, EMYB, and TMYB

were estimated in the second-stage analyses. The genetic predictions from the first-stage model, with random genotype effects, were Best Linear Unbiased Predictions (BLUPs) and cannot be used in a multi-year trial because they were shrunken (Smith, Cullis, & Thompson, 2001). Therefore, the previous model was fitted with genotype as a random effect, and the variance components from the first stage were fixed, once it ran, the model is refitted again with genotype fixed to obtain unshrunken genotype predictions. These new predictions are Best Linear Unbiased Estimates (BLUEs), comparable across trials, and the estimated weights from the stage 1 analyses are included into a weighted linear mixed model in the across-year analyses (Gezan et al., 2022). Details of the statistical model can be found in Smith, Cullis, and Gilmour (2001) and Gogel et al. (2018). The same procedure was followed to estimate the $G \times Y$ interaction of the same traits from seasons 2020 to 2021 and 2021 to 2022, plus the additional ACCB trait and the scaled traits AWTAC and TMYAC.

The genetic correlations among multiple traits within a season were estimated using the Bayesian multi-environment model (BME) function of the Bayesian multi-trait and multi-environment (BMTME) R software developed by Montesinos-López et al. (2019). Details on the linear mixed model and the script used to estimate the genetic correlations among traits are presented in the cited article.

2.5 | Genome-wide association study for predicted vegetative biomass and accumulated biomass

Spatially adjusted data generated from the two-stage analysis of predicted vegetative biomass from 2017 to 2018, 2018 to 2019, and 2020 to 2021 and accumulated biomass from 2020 to 2021 and 2021 to 2022 through ASRtriala were used as the phenotype input for GWAS. Analyses were conducted separately by season, using SNP genotypes from a set of common markers across the Affymetrix Axiom IStraw90 (Bassil et al., 2015), IStraw35 (Verma et al., 2017), and FanaSNP 50k (Hardigan et al., 2020) SNP arrays. GWAS analysis was conducted using a Bayesian information and linkage disequilibrium iteratively nested keyway method implemented in GAPIT3 (Genome Association and Prediction Integrated Tool 3) R package (Wang & Zhang, 2021) for both predicted vegetative biomass and accumulated biomass with the inclusion of the top five principal components and markers with MAF > 0.05. Relatedness was accounted for using a kinship matrix estimated with the "VanRaden" algorithm implemented in GAPIT. Significant marker-trait associations were determined using both 1% and 5% Bonferroni correction.

2.6 | Genomic selection for predicted vegetative biomass and accumulated biomass

GS was implemented using 2017-2018 as a training population to predict 2018-2019 and using 2020-2021 to predict 2021-2022, in an independent validation approach (separate years as opposed to cross-validation). Common individuals between seasons were masked in the testing set, and genotypes not previously tested were predicted. The Bayesian ridge regression (BRR) model was fitted in the R programming language (R Core Team, 2021), and run in Rstudio (Rstudio Team, 2019), using the R package BGLR (Bayesian Generalized Linear Regression) (Pérez & De Los Campos, 2014). The package implemented a Markov Chain Monte Carlo method with 15,000 iterations, where the first 5000 were used as a burn-in. Predictive ability (PA), reported as the Pearson correlation coefficient r between the observed spatially adjusted phenotypes and the genomic-estimated breeding values (GEBVs) of the testing sets (Osorio et al., 2021), was estimated for both seasons 2018-2019 and 2021-2022. Genomic narrow-sense heritability was estimated using BRR, as follows:

$$h^2 = \frac{\sigma_{p-}^2 \sigma_e^2}{\sigma_p^2}$$

where σ_p^2 is the phenotypic variance and σ_e^2 is the estimated residual variance.

3 | RESULTS

3.1 | Phenotypic variability and trait correlations

The phenotypic variation of the predicted vegetative dry biomass during the early and the total season is represented in Table S4. Predicted EBIO ranged from 8.5 to 47 g/plant across years. Predicted BIO ranged from 12.2 to 75.7 g/plant across years. Scaled traits such as EAWTB, EMYB, AWTB, and TMYB ranged from 0.3 to 1.8 g fruit weight/g biomass, 0.0 to 18.7 g marketable weight/g biomass, 0.3 to 1.6 g fruit weight/g biomass, and 0.5 to 31.8 g marketable weight/g biomass, respectively.

The genetic correlations between predicted biomass and marketable yield were positive and low to moderate $(r_G = 0.16)$ for the total season in 2017–2018 (Figure 3). A similar trend was also seen in the total season in 2018–2019 ($r_G = 0.40$) but was slightly higher than the first season. The genetic correlations between the two seasons for predicted biomass ($r_{G\times Y} = 0.57$, 0.80), AWT ($r_{G\times Y} = 0.69$, 0.87), and marketable yield ($r_{G\times Y} = 0.60$, 0.75) indicated moderate to low genotype by year interac-

tions (Figure S1). Scatter plots, histograms, and phenotypic correlations for the predicted biomass and biomass-related traits with the fruit yield and quality traits are shown in Figure S2.

During 2020-2021 and 2021-2022, aside from a different SNP marker set, we were able to access weekly drone image data, which generated another trait, the accumulated predicted vegetative biomass (ACCB), a measure of biomass over several timepoints in a growing season. This is presented as a growth-related curve in Figure S3. The genetic correlations between biomass and marketable yield were negative and low $(r_G = -0.13)$ in 2020–2021 but positive and low in 2021–2022 ($r_G = 0.14$) season (Figure 3). The phenotypic correlation for 2020–2021 (Figure S2) between biomass and yield was positive and low, in contrast to the genetic correlation for that season. For ACCB, the genetic correlation with marketable yield was low in 2020–2021 ($r_G = 0.17$) and moderate in 2021–2022 ($r_G = 0.43$). The genetic correlations between the two seasons for ACCB ($r_{G \times Y} = 0.88$), predicted biomass ($r_{G \times Y} = 0.91$), AWT ($r_{G \times Y} = 0.9$), and marketable yield ($r_{G \times Y} = 0.61$) also indicated moderate to low genotype by year interactions (Figure S4).

The heritability of the predicted biomass and biomassrelated traits were moderate to high and comparable with yield traits, with $h^2 = 0.48$ and $h^2 = 0.49$ for the EBIO 2017– 2018 and 2018–2019 seasons, respectively, and with $h^2 = 0.28$ and $h^2 = 0.26$ for EBIO 2020–2021 and 2021–2022, respectively (Table 2). For BIO, heritabilities were similar, with $h^2 = 0.50$ and $h^2 = 0.56$ for the 2017–2018 and 2018–2019 seasons, respectively, and $h^2 = 0.34$ and $h^2 = 0.30$ for 2020– 2021 and 2021-2022, respectively. For ACCB, heritabilities were $h^2 = 0.27$ and $h^2 = 0.27$ for 2020–2021 and 2021–2022, respectively. For the predicted biomass scaled traits, the heritability estimates were also moderate to high, with heritability ranging from $h^2 = 0.25 - 0.58$ across seasons for EAWTB, $h^2 = 0.39-0.60$ across seasons for EMYB, $h^2 = 0.44-0.64$ across seasons for AWTB, and $h^2 = 0.49-0.64$ across seasons for TMYB.

3.2 | Genome-wide association studies

To assess the genetic architecture of predicted biomass, we conducted a genome-wide association analysis of BIO for the 2017–2018 and 2018–2019 seasons (Figure 4) and ACCB for the 2020–2021 and 2021–2022 seasons (Figure 5). Results showed low-level peaks with no consistent SNPs associated with BIO and ACCB across years. This is evident from the lack of consistent peaks across environments in the Manhattan plots shown in Figures 4 and 5 as compared to the high-level peaks observed from strawberry disease resistance studies like in *FaRCa1* (Salinas et al., 2019), a major locus that confers resistance to anthracnose fruit rot.

a)



FIGURE 3 Genetic correlations of predicted biomass and biomass-related traits (BIO, predicted vegetative biomass; ACCB, accumulated predicted vegetative biomass; AWTB, average fruit weight on a predicted biomass basis; MYB, marketable fruit yield on a predicted biomass basis; AWTAC, average fruit weight on an accumulated predicted vegetative biomass basis; TMYAC, total marketable fruit yield on an accumulated predicted vegetative biomass basis) with yield and fruit traits (AWT, average weight; TMY, total marketable yield) during (a) the total season 2017-2018, (b) total season 2018-2019, (c) total season 2020-2021, and (e) total season 2021-2022. Early season showed similar trend as the total season. TMYB, total marketable yield on a predicted biomass basis.



FIGURE 4 Genome-wide association studies (GWAS) analysis of predicted biomass during (a) the early season 2017–2018, (b) total season 2017–2018, (c) early season 2018–2019, and (d) total season 2018–2019 showing no consistent peaks across years.

	Heritability <i>h</i> ² (SE)			
Traits	2017-2018	2018-2019	2020-2021	2021-2022
	Early			
Predicted biomass traits				
EBIO	0.48 (0.40)	0.49 (0.44)	0.28 (0.41)	0.26 (0.41)
EAWTB	0.45 (0.09)	0.58 (0.08)	0.41 (0.10)	0.25 (0.09)
EMYB	0.44 (0.17)	0.39 (0.12)	0.47 (0.14)	0.60 (0.12)
ACCB	-	-	-	-
Yield and fruit traits				
EAWT	0.45 (0.16)	0.36 (0.21)	0.44 (0.35)	0.65 (0.30)
EMKN	0.44 (0.18)	0.36 (0.12)	0.41 (0.13)	0.54 (0.15)
EMY	0.39 (0.09)	0.33 (0.08)	0.41 (0.47)	0.49 (0.55)
ESSC	0.32 (0.09)	0.28 (0.08)	0.23 (0.09)	0.27 (0.09)
ETC	0.20 (0.39)	0.14 (0.23)	0.20 (0.08)	0.08 (0.06)
FRM	-	-	-	-
	Total			
Predicted biomass traits				
BIO	0.50 (0.90)	0.56 (0.55)	0.34 (0.73)	0.30 (0.50)
AWTB	0.57 (0.09)	0.64 (0.07)	0.50 (0.10)	0.44 (0.10)
ТМҮВ	0.54 (0.17)	0.49 (0.18)	0.63 (0.28)	0.64 (0.22)
ACCB	-	-	0.27 (0.38)	0.27 (0.61)
Yield and fruit traits				
AWT	0.56 (0.18)	0.70 (0.13)	0.46 (0.37)	0.83 (0.21)
MKN	0.53 (0.36)	0.44 (0.48)	0.55 (0.32)	0.50 (0.43)
TMY	0.48 (0.10)	0.43 (0.10)	0.50 (0.11)	0.46 (0.11)
SSC	0.37 (0.09)	0.37 (0.08)	0.44 (0.10)	0.11 (0.06)
TC	0.33 (0.41)	0.17 (0.18)	0.19 (0.08)	0.16 (0.07)
FRM	0.50 (0.08)	0.68 (0.07)	0.57 (0.09)	0.44 (0.09)

TABLE 2 Narrow-sense heritability (h^2) of predicted vegetative biomass, biomass-related traits, and yield traits in the University of Florida strawberry breeding population during the 2017–2018, 2018–2019, 2020–2021, and 2021–2022 seasons.

Note: Values given in parentheses are standard error.

Abbreviations: ACCB, accumulated predicted vegetative biomass; AWT, average fruit weight; AWTB, average fruit weight on a predicted vegetative biomass basis; BIO, predicted vegetative biomass; EAWTB, early average fruit weight on a predicted vegetative biomass basis; EBIO, vegetative biomass that represents the early season; EMKN, early marketable number; EMY, early marketable yield; EMYB, early marketable yield on a predicted vegetative biomass basis; ESSC, early solubale solidf content; ETC; early total culls; FRM, firmness; MKN, marketable number per plant; SSC, soluble solids content; TC, total culls; TMY, total marketable yield; TMYB, total marketable yield on a predicted vegetative biomass basis.

3.3 | Genomic selection

GS was tested as a breeding strategy for predicted vegetative biomass and biomass-related traits in our breeding program (Table 3). Since two different sets of SNP markers were used across years, we generated separate PAs with BRR as the model. The PA for BIO and BIO-related traits in 2017–2018 (2017–2018 season was used as the training set and 2018– 2019 season was used as the testing set) ranged from 0.24 to 0.46. The PA for BIO and BIO related traits in 2020–2021 (2020–2021 season was used as the training set and 2021– 2022 season was used as the testing set) ranged from 0.25 to 0.36. The correlation of heritability in 2018 and PA in 2017– 2018 was 0.67, while the correlation of heritability in 2021 and PA in 2020–2021 was 0.92, with ACCB included. PA for EAWT, EMY, AWT, and TMY were also included for comparison since those are traits routinely used in GS in the breeding program (Table 3).

4 | DISCUSSION

Vegetative biomass is an important trait to help assess a plant's yield potential. However, traditional measurement of vegetative biomass is destructive and time consuming. In a breeding program, it is essential to collect yield and fruit quality data at every harvest. Thus, destroying field plots to measure biomass manually is not a viable option. With the



FIGURE 5 Genome-wide association studies (GWAS) analysis of accumulated biomass during (a) the total season 2020–2021 and (b) total season 2021–2022 showing no consistent peaks across years.

emergence of DP technologies, it has become possible to estimate vegetative biomass at scale.

4.1 | Trait correlations and heritability

The phenotypic and genetic correlations between predicted biomass and marketable yield in both the early season and total season were generally positive but low in magnitude (Figure 3). This suggests that fruit yield can be improved without a corresponding increase in plant biomass. The moderate heritability of biomass-scaled yield also suggests that improvements in yield on a plant biomass basis are quite feasible. Overall, heritability estimates for biomass traits were in similar ranges with yield traits, suggesting that directional selection for biomass will also result in genetic gains (Table 2). This is noteworthy since the biomass traits are not ground-truth measurements. EBIO and BIO are based on size metrics extracted from high-resolution images that were used to develop a regression model modified from Guan et al. (2020) with an adjusted R^2 of 0.77. If the accuracy of biomass estimation can be increased, we should expect increases in heritability above what are reported here.

It is important to reiterate that our goal in strawberry breeding is not to increase vegetative biomass. In the strawberry industry, cost of labor is a significant bottleneck with the majority being used on harvesting. Thus, the goal is to select for high-yielding varieties without simultaneously increasing plant biomass. More biomass could mean a larger canopy and/or increased canopy density. A moderate plant size allows strawberry plants to fit within the desired in-row spacings that could maximize yield. In addition, plant density must not be increased too much, as air circulation within the plant and the ability for harvesters to visualize fruit under the canopy is critical. Currently, the recommended plant populations in Florida are 16,000-20,000 plants per acre on a two-row bed with distance of 12-16 in. within row and 12-14 in. between row (Whitaker et al., 2023). A moderate plant size also allows fruits to be more visible and thus helps increase efficiency of manual harvesting in contrast to larger plant size, which tends to hide fruits under its canopy. Thus, it was important to examine scaled traits representing yield-related traits on plant biomass basis. These scaled traits are meant to capture a roughly similar concept to harvest index, which is the ratio of yield to total harvested aboveground biomass in grain crops (Brandán et al., 2009; Porker et al., 2020; Van Hooijdonk et al., 2018). However, we should also reiterate that the yield traits that we measure are only from marketable fruits, and that our ground truth biomass data do not include reproductive parts and runners but only leaves and stems. The heritability

TABLE 3Predictive ability for predicted vegetative biomass,
biomass-related traits, and yield and fruit traits in the University of
Florida strawberry breeding population during the 2017–2018,
2018–2019, 2020–2021, and 2021–2022 seasons.

	Predictive ability (PA)		
Tue:44	2017-2018/	2020-2021/	
Traits	2018-2019	2021-2022	
	Early		
Predicted biomass traits			
EBIO	0.24	0.20	
EAWTB	0.31	0.15	
EMYB	0.24	0.31	
ACCB			
Yield and fruit traits			
EAWT	0.35	0.30	
EMY	0.22	0.28	
	Total		
Predicted biomass traits			
BIO	0.41	0.23	
AWTB	0.46	0.21	
TMYB	0.33	0.36	
ACCB	-	0.24	
Yield and fruit traits			
AWT	0.41	0.45	
TMY	0.25	0.23	

Note: The first year as the training set and the second year as the testing set for independent validation.

Abbreviations: ACCB, accumulated predicted vegetative biomass; AWT, average fruit weight; AWTB, average fruit weight on a predicted biomass basis; BIO, predicted vegetative biomass; EAWTB, early average fruit weight on a predicted biomass basis; EBIO, vegetative biomass that represents the early season; EMY, early marketable yield; EMYB, early marketable yield on a predicted biomass basis; TMY, total marketable yield; TMYB, total marketable yield on a predicted biomass basis.

estimates for the biomass-scaled traits are moderate and thus can be improved in our breeding program.

In 2020–2021 and 2021–2022 seasons, we were able to obtain UAV-derived biomass data (C. Zheng et al., 2022a). Unlike the ground-based method, the UAV could be flown weekly to collect image data. This allowed the calculation of ACCB. This trait better reflects the vegetative growth trajectory of a strawberry plant over the season, which may give more in-depth insights into how vegetative biomass relates to fruit production. Interestingly, in this study, the accumulated biomass traits did not have higher heritabilities or predictive abilities compared to the biomass estimates at single time points. Thus, it is not clear whether weekly UAV imaging can provide additional value over single flights at the end of the early season and total season when predicting vegetative biomass. If more time points do not provide

an advantage, they would represent a waste of personnel and computing resources. A study on the rate of growth of individual genotypes using weekly UAV imaging could be a potential trait to explore that could be used in phenomic selection.

4.2 | Genome-wide studies

GWAS analysis did not reveal consistent peaks across trials. This was not unexpected, as biomass should be controlled by a large number of physiological, biochemical, and genetic factors. This was also a strong indication that genomic prediction would be the best strategy for DNA-informed breeding for these traits (Clark et al., 2019; Tilhou & Casler, 2022; Toda et al., 2021). Predictive abilities for the predicted biomass and biomass-related traits in this study were similar to the yield traits in a scenario in which one field trial was used to predict the next year's trial. This suggests that genomic predicting untested genotypes in new environments. Therefore, inclusion of vegetative biomass in the multi-trait index or the use of biomass-scaled yield traits in GS should help balance genetic gains for fruit and plant traits.

5 | CONCLUSIONS

With the advent of DP technologies, we were able to measure aboveground vegetative biomass on a large scale, a trait that we have not measured before in the UF strawberry breeding program. This approach is allowing our program to efficiently balance selection for fruit yield and plant size within the context of GS for the first time. We expect that as additional data are accumulated over seasons, increasing the training population size, predictive abilities will continue to improve. Further improvements in the accuracy of DP methods for biomass estimation could also help improve the accuracy of selection for vegetative biomass traits in the future.

AUTHOR CONTRIBUTIONS

Cheryl Dalid: Conceptualization; data curation; formal analysis; project administration; visualization; writing—original draft. Caiwang Zheng: Data curation; formal analysis; writing—review and editing. Luis Osorio: Data curation; formal analysis; writing—review and editing. Sujeet Verma: Formal analysis; writing—review and editing. Amr Abd-Elrahman: Resources; writing—review and editing. Xu Wang: Resources; writing—review and editing. Vance M. Whitaker: Conceptualization; funding acquisition; project administration; supervision; writing—original draft.

ACKNOWLEDGMENTS

We thank the staff of the UF/IFAS Gulf Coast Research and Education Center Strawberry Breeding and Genetics laboratory and former interns, Juan David Jimenez Pardo and Mariel Andrea Gallardo Salazar for their work in data collection and trial maintenance. We also thank the UF/IFAS Geomatics and Plant Phenomics laboratories for their work in high-resolution image data collection and processing. This research was partially supported by grants awarded to Vance M. Whitaker from the Florida Strawberry Research and Education Foundation (FSREF) and support from the Florida Agricultural Experiment Station (FAES).

CONFLICT OF INTEREST STATEMENT The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All phenotypic data and spatial analysis models are included in the Supporting Information tables.

ORCID

Cheryl Dalid https://orcid.org/0000-0001-6327-404X *Vance M. Whitaker* https://orcid.org/0000-0002-2172-3019

REFERENCES

- Abd-Elrahman, A., Guan, Z., Dalid, C., Whitaker, V., Britt, K., Wilkinson, B., & Gonzalez, A. (2020). Automated canopy delineation and size metrics extraction for strawberry dry weight modeling using raster analysis of high-resolution imagery. *Remote Sensing 2020*, 12(21), 3632. https://doi.org/10.3390/RS12213632
- Agehara, S. (2021). Characterizing early-season nitrogen fertilization rate effects on growth, yield, and quality of strawberry. *Agronomy*, *11*(5), 905. https://doi.org/10.3390/AGRONOMY11050905
- Bassil, N. V., Davis, T. M., Zhang, H., Ficklin, S., Mittmann, M., Webster, T., Mahoney, L., Wood, D., Alperin, E. S., Rosyara, U. R., Koehorst-vanc Putten, H., Monfort, A., Sargent, D. J., Amaya, I., Denoyes, B., Bianco, L., van Dijk, T., Pirani, A., Iezzoni, A., ... van de Weg, E. (2015). Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry Fragaria × ananassa. *BMC Genomics*, *16*(1), Article 155. https://doi.org/10.1186/S12864-015-1310-1
- Brandán, E. Z., Fernández, R. R., Villagra, E. L., Salazar, S. M., Cruz, C., & Ponce De León, R. (2009). Effects of latitude in morphological and physiologic parameters and plant harvest index in fresh plants of strawberry cv. "Camarosa." *Acta Horticulturae*, 842, 711–714. https://doi.org/10.17660/ActaHortic.2009.842.154
- Burgess, A. J., Masclaux-Daubresse, C., Strittmatter, G., Weber, A. P. M., Taylor, S. H., Harbinson, J., Yin, X., Long, S., Paul, M. J., Westhoff, P., Loreto, F., Ceriotti, A., Saltenis, V. L. R., Pribil, M., Nacry, P., Scharff, L. B., Jensen, P. E., Muller, B., Cohan, J. P., ... Baekelandt, A. (2023). Improving crop yield potential: Underlying biological processes and future prospects. *Food and Energy Security*, *12*(1), e435. https://doi.org/10.1002/FES3.435
- Clark, L. V., Dwiyanti, M. S., Anzoua, K. G., Brummer, J. E., Ghimire, B. K., Głowacka, K., Hall, M., Heo, K., Jin, X., Lipka, A. E., Peng,

- J., Yamada, T., Yoo, J. H., Yu, C. Y., Zhao, H., Long, S. P., & Sacks, E. J. (2019). Genome-wide association and genomic prediction for biomass yield in a genetically diverse *Miscanthus sinensis* germplasm panel phenotyped at five locations in Asia and North America. *GCB Bioenergy*, *11*(8), 988–1007. https://doi.org/10.1111/GCBB.12620
- Crain, J., Mondal, S., Rutkoski, J., Singh, R. P., & Poland, J. (2018). Combining high-throughput phenotyping and genomic information to increase prediction and selection accuracy in wheat breeding. *The Plant Genome*, 11(1), 170043. https://doi.org/10.3835/ PLANTGENOME2017.05.0043
- Crossa, J., De Los Campos, G., Pérez, P., Gianola, D., Burgueño, J., Araus, J. L., Makumbi, D., Singh, R. P., Dreisigacker, S., Yan, J., Arief, V., Banziger, M., & Braun, H. J. (2010). Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics*, 186(2), 713–724. https://doi.org/10. 1534/genetics.110.118521
- de Bem Oliveira, I., Resende, M. F. R., Ferrão, L. F. V., Amadeu, R. R., Endelman, J. B., Kirst, M., Coelho, A. S. G., & Munoz, P. R. (2019). Genomic prediction of autotetraploids; influence of relationship matrices, allele dosage, and continuous genotyping calls in phenotype prediction. *G3: Genes, Genomes, Genetics*, 9(4), 1189–1198. https://doi.org/10.1534/g3.119.400059
- Fan, Z., Verma, S., Lee, H., Jang, Y. J., Wang, Y., Lee, S., & Whitaker, V. M. (2024). Strawberry soluble solids QTL with inverse effects on yield. *Horticulture Research*, 11(2), uhad271. https://doi.org/10. 1093/HR/UHAD271
- Fan, Z., & Whitaker, V. M. (2024). Genomic signatures of strawberry domestication and diversification. *The Plant Cell*, 36(5), 1622–1636. https://doi.org/10.1093/PLCELL/KOAD314
- FAO-UN. (2022). Data. FAO. https://www.fao.org/faostat/en/#data
- Gezan, S. A., Galli, G., & Murray, D. (2022). User's manual for ASRtriala v. 1.0.0: An R package with complementary functions for single and multi-environment trial analyses. VSN International Ltd. https:// vsni.co.uk/free-software/asrtriala
- Gezan, S. A., Osorio, L. F., Verma, S., & Whitaker, V. M. (2017). An experimental validation of genomic selection in octoploid strawberry. *Horticulture Research*, 4, 16070. https://doi.org/10.1038/hortres. 2016.70
- Gogel, B., Smith, A., & Cullis, B. (2018). Comparison of a one- and two-stage mixed model analysis of Australia's National Variety Trial Southern Region wheat data. *Euphytica*, 214(2), Article 44. https:// doi.org/10.1007/S10681-018-2116-4
- Guan, Z., Abd-Elrahman, A., Fan, Z., Whitaker, V. M., & Wilkinson, B. (2020). Modeling strawberry biomass and leaf area using objectbased analysis of high-resolution images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *163*, 171–186. https://doi.org/10. 1016/j.isprsjprs.2020.02.021
- Guan, Z., Wu, F., & Whidden, A. J. (2018). Florida strawberry production costs and trends. *EDIS*, 2018(1), 5. https://doi.org/10.32473/ EDIS-FE1013-2017
- Hardigan, M. A., Feldmann, M. J., Lorant, A., Bird, K. A., Famula, R., Acharya, C., Cole, G., Edger, P. P., & Knapp, S. J. (2020). Genome synteny has been conserved among the octoploid progenitors of cultivated strawberry over millions of years of evolution. *Frontiers in Plant Science*, 10, Article 1789. https://doi.org/10.3389/fpls.2019.01789
- Ibrahim, A. K., Zhang, L., Niyitanga, S., Afzal, M. Z., Xu, Y., Zhang, L., Zhang, L., & Qi, J. (2020). Principles and approaches of association mapping in plant breeding. *Tropical Plant Biology*, 13(3), 212–224. https://doi.org/10.1007/S12042-020-09261-4

- Kawano, T., Wallbridge, N., & Plummer, C. (2020). Logistic models for simulating the growth of plants by defining the maximum plant size as the limit of information flow. *Plant Signaling & Behavior*, 15(2), 1709718. https://doi.org/10.1080/15592324.2019.1709718
- Krause, M. R., Mondal, S., Crossa, J., Singh, R. P., Pinto, F., Haghighattalab, A., Shrestha, S., Rutkoski, J., Gore, M. A., Sorrells, M. E., & Poland, J. (2020). Aerial high-throughput phenotyping enables indirect selection for grain yield at the early generation, seedlimited stages in breeding programs. *Crop Science*, 60(6), 3096–3114. https://doi.org/10.1002/CSC2.20259
- Merrick, L. F., & Carter, A. H. (2021). Comparison of genomic selection models for exploring predictive ability of complex traits in breeding programs. *The Plant Genome*, 14(3), e20158. https://doi.org/10.1002/ tpg2.20158
- Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4), 1819–1829. https://doi.org/10.1093/GENETICS/157.4. 1819
- Montesinos-López, O. A., Montesinos-López, A., Luna-Vázquez, F. J., Toledo, F. H., Pérez-Rodríguez, P., Lillemo, M., & Crossa, J. (2019). An R package for Bayesian analysis of multi-environment and multi-trait multi-environment data for genome-based prediction. *G3: Genes, Genomes, Genetics*, 9(5), 1355–1369. https://doi.org/10.1534/ G3.119.400126
- O'Connor, K. M., Hayes, B. J., Hardner, C. M., Alam, M., Henry, R. J., & Topp, B. L. (2021). Genomic selection and genetic gain for nut yield in an Australian macadamia breeding population. *BMC Genomics*, 22(1), Article 370. https://doi.org/10.1186/S12864-021-07694-Z
- Osorio, L. F., Gezan, S. A., Verma, S., & Whitaker, V. M. (2021). Independent validation of genomic prediction in strawberry over multiple cycles. *Frontiers in Genetics*, 11, Article 596258. https://doi.org/10. 3389/FGENE.2020.596258
- Peixoto, M. A., Leach, K. A., Jarquin, D., Flannery, P., Zystro, J., Tracy, W. F., Bhering, L., & Resende, M. F. R. (2024). Utilizing genomic prediction to boost hybrid performance in a sweet corn breeding program. *Frontiers in Plant Science*, *15*, 1293307. https://doi.org/10. 3389/FPLS.2024.1293307
- Pérez, P., & De Los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, 198(2), 483– 495. https://doi.org/10.1534/GENETICS.114.164442
- Pérez-Ruiz, M., Prior, A., Martinez-Guanter, J., Apolo-Apolo, O. E., Andrade-Sanchez, P., & Egea, G. (2020). Development and evaluation of a self-propelled electric platform for high-throughput field phenotyping in wheat breeding trials. *Computers and Electronics in Agriculture*, 169, 105237. https://doi.org/10.1016/J.COMPAG.2020. 105237
- Pincot, D. D. A., Hardigan, M. A., Cole, G. S., Famula, R. A., Henry, P. M., Gordon, T. R., & Knapp, S. J. (2020). Accuracy of genomic selection and long-term genetic gain for resistance to Verticillium wilt in strawberry. *The Plant Genome*, *13*(3), e20054. https://doi.org/10. 1002/tpg2.20054
- Poorter, H., & Nagel, O. (2000). The role of biomass allocation in the growth response of plants to different levels of light, CO₂, nutrients and water: A quantitative review. *Functional Plant Biology*, 27(12), 1191. https://doi.org/10.1071/PP99173_CO
- Poorter, H., Niinemets, Ü., Poorter, L., Wright, I. J., & Villar, R. (2009). Causes and consequences of variation in leaf mass per area (LMA): A meta-analysis. *New Phytologist*, 182(3), 565–588. https://doi.org/10. 1111/J.1469-8137.2009.02830.X

- Porker, K., Straight, M., & Hunt, J. R. (2020). Evaluation of G × E × M interactions to increase harvest index and yield of early sown wheat. *Frontiers in Plant Science*, *11*, Article 994. https://doi.org/10.3389/FPLS.2020.00994
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
- RStudio Team. (2019). *RStudio: Integrated development for R*. RStudio. http://www.rstudio.com/
- Salinas, N., Verma, S., Peres, N., & Whitaker, V. M. (2019). FaRCa1: A major subgenome-specific locus conferring resistance to *Colletotrichum acutatum* in strawberry. *Theoretical and Applied Genetics*, 132(4), 1109–1120. https://doi.org/10.1007/s00122-018-3263-7
- Smith, A., Cullis, B., & Gilmour, A. (2001). Applications: The analysis of crop variety evaluation data in Australia. Australian & New Zealand Journal of Statistics, 43(2), 129–145. https://doi.org/10. 1111/1467-842X.00163
- Smith, A., Cullis, B., & Thompson, R. (2001). Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics*, 57(4), 1138–1147. https://doi.org/10.1111/J.0006-341X.2001.01138.X
- Tessema, B. B., Liu, H., Sørensen, A. C., Andersen, J. R., & Jensen, J. (2020). Strategies using genomic selection to increase genetic gain in breeding programs for wheat. *Frontiers in Genetics*, 11, Article 578123. https://doi.org/10.3389/fgene.2020.578123
- Tilhou, N. W., & Casler, M. D. (2022). Biomass yield improvement in switchgrass through genomic prediction of flowering time. GCB Bioenergy, 14(9), 1023–1034. https://doi.org/10.1111/GCBB.12983
- Toda, Y., Kaga, A., Kajiya-Kanegae, H., Hattori, T., Yamaoka, S., Okamoto, M., Tsujimoto, H., & Iwata, H. (2021). Genomic prediction modeling of soybean biomass using UAV-based remote sensing and longitudinal model parameters. *The Plant Genome*, 14(3), e20157. https://doi.org/10.1002/tpg2.20157
- USDA-NASS. (2023). Noncitrus fruits and nuts 2022 summary. https:// usda.library.cornell.edu/concern/publications/zs25x846c
- Van Hooijdonk, B. M., Tustin, D. S., Dayatilake, G. A., Oliver, M. J., & Richardson, A. (2018). Harvest index of apple and kiwifruit—How high can we go by optimising plant architecture and canopy management? *Acta Horticulturae*, *1228*, 307–313. https://doi.org/10.17660/ ActaHortic.2018.1228.46
- Verhulst, P. F. (1838). Notice sur la loi que la population poursuit dans son accroissement. *Correspondance mathématique et physique*, 10, 113–121. https://www.sciepub.com/reference/6960
- Verma, S., Bassil, N. V., Van De Weg, E., Harrison, R. J., Monfort, A., Hidalgo, J. M., Amaya, I., Denoyes, B., Mahoney, L. L., Davis, T. M., Fan, Z., Knapp, S., & Whitaker, V. M. (2017). Development and evaluation of the Axiom® IStraw35 384HT array for the allo-octoploid cultivated strawberry Fragaria ×ananassa. *Acta Horticulturae*, 1156, 75–81. https://doi.org/10.17660/ACTAHORTIC.2017.1156.10
- Wang, J., & Zhang, Z. (2021). GAPIT version 3: Boosting power and accuracy for genomic association and prediction. *Genomics, Pro*teomics & Bioinformatics, 19(4), 629–640. https://doi.org/10.1016/ j.gpb.2021.08.005
- Whitaker, V. M., Boyd, N. S., Peres, N. A., Desaeger, J., Lahiri, S., & Agehara, S. (2023). *Strawberry production* (HS736). University of Florida Institute of Food and Agricultural Sciences. https://doi.org/ 10.32473/edis-cv134-2023

- Whitaker, V. M., & Fan, Z. (2020). What will Florida strawberry varieties look like in the future? A breeder's perspective. *International Journal of Fruit Science*, 20(Suppl 1), 992–996. https://doi.org/10.1080/ 15538362.2020.1755770
- Whitaker, V. M., Osorio, L. F., Hasing, T., & Gezan, S. (2012). Estimation of genetic parameters for 12 fruit and vegetative traits in the University of Florida strawberry breeding population. *Journal* of the American Society for Horticultural Science, 137(5), 316–324. https://doi.org/10.21273/JASHS.137.5.316
- Wu, F., Guan, Z., & Whitaker, V. (2015). Optimizing yield distribution under biological and economic constraints: Florida strawberries as a model for perishable commodities. *Agricultural Systems*, 141, 113– 120. https://doi.org/10.1016/J.AGSY.2015.10.002
- Zheng, C., Abd-Elrahman, A., Whitaker, V., & Dalid, C. (2022a). Prediction of strawberry dry biomass from UAV multispectral imagery using multiple machine learning methods. *Remote Sensing*, 14(18), 4511. https://doi.org/10.3390/rs14184511
- Zheng, C., Abd-Elrahman, A., Whitaker, V. M., & Dalid, C. (2022b). Deep learning for strawberry canopy delineation and biomass prediction from high-resolution images. *Plant Phenomics*, 2022, 9850486. https://doi.org/10.34133/2022/9850486
- Zheng, H., Cheng, T., Zhou, M., Li, D., Yao, X., Tian, Y., Cao, W., & Zhu, Y. (2019). Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. *Pre-Pre-*

cision Agriculture, 20(3), 611–629. https://doi.org/10.1007/s11119-018-9600-7

Zhu, X. G., Long, S. P., & Ort, D. R. (2010). Improving photosynthetic efficiency for greater yield. *Annual Review of Plant Biol*ogy, 61, 235–261. https://doi.org/10.1146/ANNUREV-ARPLANT-042809-112206

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Dalid, C., Zheng, C., Osorio, L., Verma, S., Abd-Elrahman, A., Wang, X., & Whitaker, V. M. (2025). Genetic analysis of predicted vegetative biomass and biomass-related traits from digital phenotyping of strawberry. *The Plant Genome*, *18*, e70018. https://doi.org/10.1002/tpg2.70018