Crop Science

ORIGINAL RESEARCH ARTICLE

Crop Breeding & Genetics

Quantitative characterization of proximate sensing canopy traits in the SoyNAM population

Ben P. Hall[1] | **Alencar Xavier[1,2]** | **Fabiana Moreira[1]** | **Shaun Casteel[1]** | **Katy M. Rainey[1]**

[1] Dep. of Agronomy, Purdue Univ., West Lafayette, IN 47907, USA

[2] Corteva Agrisciences, 8305 NW 62nd Ave., Johnston, IA 50131, USA

Correspondence
Katy M. Rainey, Dep. of Agronomy, Purdue Univ., West Lafayette, IN 47907, USA.
Email: krainey@purdue.edu

Assigned to Associate Editor Zenglu Li.

Abstract

The average increase of soybean [*Glycine max* (L.) Merr.] grain yield per year is not enough to meet the estimated need by 2050. Contributing factors may include narrowing genetic diversity in cultivated soybean germplasm arising from direct selection for grain yield and insufficient knowledge of the genetic influences on physiological traits associated with grain yield. Canopy development plays a critical role in the light intercepted by soybean crops throughout the growing season. Rapid canopy closure facilitates complete light interception and optimizes growth dynamics, ultimately increasing biomass accumulation and grain yield. This study sought to use ground-based digital imagery as a phenotyping tool for quantitative characterization of seasonal canopy coverage and light interception in the genetically diverse Soybean Nested Association Mapping (SoyNAM) population. We took weekly measurements of canopy coverage from early vegetative to mid-reproductive growth for 5,600 recombinant inbred lines during the 2013 and 2014 seasons. An asymptotic logistic growth curve fit to the coverage data allowed us to estimate daily changes in canopy coverage in the intervals between the sampling dates, enabling calculation of canopy dynamic parameters. These parameters included average canopy coverage, cumulative intercepted photosynthetically active radiation for the total sampling period, vegetative growth period, reproductive growth period, and the number of days required to reach 30, 50, and 70% canopy coverage. Variance component estimation showed that genetic differences among families significantly influenced variation in phenotypic expression for all canopy parameters. The narrow-sense heritabilities for all canopy parameters were high (h^2 = .81–.90), suggesting opportunities for genetic gains in canopy development through selection of superior genotypes.

Abbreviations: ACC, average canopy coverage;; CC, canopy coverage; CIPAR, cumulative intercepted photosynthetically active radiation; CIPARr, cumulative intercepted photosynthetically active radiation for the reproductive phase; CIPARv, cumulative intercepted photosynthetically active radiation for the vegetative phase; DAP, days after planting; LI, light interception; PAR, photosynthetically active radiation; QTL, quantitative trait locus; RIL, recombinant inbred line; SNP, single nucleotide polymorphism; SoyNAM, Soybean Nested Association Mapping; Yp, yield potential.

1 | INTRODUCTION

In the past century, soybean [*Glycine max* (L.) Merr.] yields have improved at an average rate of ~31 kg ha^{-1} annually (Ainsworth, Yendrek, Skoneczka, & Long, 2012; Ray, Mueller, West, & Foley, 2013). However, an increase of approximately 25–70% above current production levels must occur to meet the estimated global needs in 2050 (Hunter, Smith, Schipanski, Atwood, & Mortensen, 2017). Ray et al. (2013) found that soybean yield would have to increase from a rate of 1.3 to 2.4% yr^{-1} to double global production by 2050. The low heritability of soybean grain yield (Anand & Torrie, 1963; Bartley & Weber, 1952; Hanson & Weber, 1962; Johnson, Robinson, & Comstock, 1955) and the narrowing genetic diversity of soybean germplasm, due to the general trend of direct selection for grain yield (Gizlice, Carter, & Burton, 1994; Rincker et al., 2014; Sneller, 1994), are challenges that must be overcome to improve soybean through breeding.

Incorporating phenotypic selection of secondary traits associated with grain yield into breeding programs may both improve the rate of genetic gains and increase the selection space to better exploit germplasm diversity (Kahlon, Board, & Kang, 2011; Richards, 2000). The effectiveness of this method of selection is dependent on the heritability (h^2) of both traits and the additive genetic correlation between both traits ($r_{A_{AB}}$). If the heritability of the secondary trait (*B*) is greater than that of the primary trait (*A*), and the two traits have a high additive genetic correlation to each other, then the response in *A* produced by selecting for *B* has the potential to be greater than directly selecting for *A* itself; this can be estimated by calculating the efficiency of indirect selection (R_e), per Falconer (1981). In addition, the use of indirect selection may be preferable when the secondary trait is easier or less expensive to measure than the primary trait, is less sensitive to genotype × environment interaction, and can be selected out-of-season or in earlier developmental stages or generations (Bernardo, 2010; Richards, 2000). Recent advances in field-based phenomics have improved the precision, efficiency, and cost of phenotyping and provide methods for indirectly selecting for grain yield by selecting for strategic physiological and developmental traits that may be responsible for variation in yield (White et al., 2012).

Soybean's capacity to intercept incident solar radiation is a determining factor for optimizing crop growth and maximizing yield potential (Yp) (Campillo, Prieto, Daza, Monino, & Garcia, 2008; Purcell, Ball, Reaper, & Vories, 2002). Light interception (LI) provides the energy demand required for soybean growth and thus influences many physiological processes such as photosynthesis, transpiration, crop growth rate, and total dry matter accumulation (Board, Kamal, & Harville,

1992; Board & Kahlon, 2011; Campillo et al., 2008). Thus, optimization of these physiological processes depends largely on the ability to maximize LI, which makes it an important target for selection of superior genotypes (Shibles & Weber, 1965; Wells, 1991). The effective capture of solar radiation depends on the proportion of light intercepted by developing crop canopies; thus, the percentage of canopy coverage (CC) and percentage of LI can be used interchangeably (Purcell, 2000). Canopy coverage can be interpreted as the physical expression of all the reactions and processes that occur during crop growth, making CC a facilitating parameter controlling the ability of soybean to achieve its inherent Yp (Bojacá, García, & Schrevens, 2011; Koester, Skoneczka, Cary, Diers, & Ainsworth, 2014). The capacity of canopies to intercept the seasonal cumulative photosynthetically active radiation (CIPAR) is determined by the speed and duration of canopy closure (Purcell et al., 2002). Richards (2000) suggested that longer crop growth and increased interception of solar radiation are the two components that have contributed most to increases in crop biomass, photosynthesis, and grain yield.

Selection for increased LI has been occurring since the beginning of crop domestication by selection of fast-growing plants with larger and more leaves (Richards, 2000). Evaluating cultivars released between 1923 and 2007, Koester et al. (2014) found that the greater Yp of newer cultivars was associated with increased LI, the efficiency of radiation use, and partitioning efficiency. They also suggested that CIPAR has increased in newer cultivars due to decreased lodging and later maturing cultivars. Despite the importance of canopy LI for optimizing soybean growth and grain yield, the majority of canopy LI studies have explored grain yield variation in soybean subjected to various planting densities and row spacing configurations, and have done so with great success (Board & Harville, 1992; Board et al., 1992; De Bruin & Pedersen, 2009; Edwards, Purcell, & Karcher, 2005; Purcell et al., 2002; Shibles & Weber, 1965; Wells, 1991). Studies have also shown the ability to develop accurate models using measurements of LI obtained during specific developmental periods to predict grain yield, stress-induced grain yield reduction, and total dry matter (Board & Kahlon, 2012; Board et al., 2011).

Few genetic studies have derived variance components from measurements of soybean LI to evaluate genetic variation among crosses in soybean families in an effort to estimate heritability and additive genetic correlations with other traits. Jannink, Orf, Jordan, and Shaw (2000) reported low estimates of heritability for canopy LI ($h^2 = .19–.27$); however, the study only took measurements of early and midseason LI and therefore lacked extensive characterization of seasonal CC dynamics. Hoyos-Villegas, Houx, Singh, and Fritschi (2014) used ground-based digital imagery to acquire weekly

measurements of CC during early vegetative growth through complete leaf senescence to evaluate the responses of two cultivars subject to various rooting depth restrictions. Canopy coverage measurements at various crop stages showed phenotypic correlations with leaf biomass, aboveground biomass, plant height, net photosynthetic rate, and grain yield. Furthermore, multiple linear regression analysis incorporating only canopy data was able to accurately predict all of the abovementioned physiological variables, in addition to crop growth rate, across root restriction treatments. A genome-wide association analysis of CC across several sampling dates throughout the growing season detected, among others, a large quantitative trait locus (QTL) on soybean chromosome 19 (Xavier, Hall, Hearst, Cherkauer, & Rainey, 2017). This QTL resulted in an estimated increase in grain yield of 47.30 kg ha^{-1} with no increase in days to maturity (-0.24 d). In this study, the parameter average CC, consisting of the average of multiple observed values of CC during growing season, has a high heritability ($h^2 = .77$) and strong additive genetic correlation with grain yield (.87), making it a potential target for indirect selection of grain yield. Kaler et al. (2018) also found significant single nucleotide polymorphisms (SNPs) associated with CC obtained from two early-season samplings and suggested that those SNPs may be useful to obtain faster CC or improving radiation interception in breeding programs. Jarquin, Howard, Xavier, and Das Choudhury (2018) used imagery-based CC recorded between 14 and 71 d after planting (DAP) in genomic prediction models. They found that CC data improved the predictive ability of the genomic prediction models for yield, and that using only early-season canopy (14–33 DAP) had comparable improvements, suggesting that it would be a valuable trait to assist the selection process when yield data are missing.

an Nested Association Mapping (SoyNAM) population is a valuable germplasm resource for genetic analysis of the complex physiological traits that influence Yp because of its broad genetic diversity, consisting of high-yielding elite cultivars, experimental breeding lines, lines with exotic ancestry, and plant introductions spanning several geographic regions (Diers et al., 2018, Xavier et al., 2018). This study's aims were (a) to assess the use of digital imagery for periodic estimation of fractional CC and derive canopy dynamic parameters to evaluate seasonal variation and growth patterns of canopy development, (b) to evaluate the degree of genetic variation for canopy development parameters across genetically distinct SoyNAM families, and (c) to evaluate canopy parameters for yield improvement. The current study builds upon our previous association mapping study (Xavier, Hall, Hearst, et al., 2017) by assessing a set of canopy-related traits from both a breeding and physiology perspective, providing quantitative descriptors at the family and population level.

2 | MATERIALS AND METHODS

We conducted field experiments in 2013 and 2014 at the Purdue University Agronomy Center for Research and Education (ACRE), West Lafayette, IN (40°28′20.5″ N, 86°59′32.3″ W; 216 m asl) on a Chalmers silty clay loam (fine-silty, mixed, superactive, mesic Typic Endoaquolls) Raub–Brenton complex (fine-silty, mixed, superactive, mesic Aquic Argiudolls).

2.1 | Population

This study was collateral to the SoyNAM population main study for mapping yield and consisted of 40 families of F$_5$-derived lines developed through single-seed descent by crossing 40 founder lines (Maturity Groups I–V) to a common hub parent IA3023 (Maturity Group III), constituting an association panel suitable to the study of complex traits of agronomic interest (Diers et al., 2018). This generated 140 recombinant inbred lines (RILs) per family, for a total of 5,600 RILs. Individuals were genotyped with the Illumina SoyNAM BeadChip SNP array compressing 5,305 SNPs specifically designed for the SoyNAM with markers called from the parental sequence (Song et al., 2017).

2.2 | Experimental settings

The experimental design was a modified augmented design with control lines replicated in sets, and no replication of RILs. Each family of 140 RILs was divided equally into four sets composed of 35 RILs and five checks randomized within the sets. A total of 160 sets were randomized throughout the field.

Experiments were planted on 20 May 2013 and 23 May 2014 in two-row plots (273 × 152 cm) spaced at 76 cm, at a rate of 34 untreated seeds m^{-2}, seeded using an Almaco Seed pro 360 with Skytrip. Experiments were maintained free of weeds to ensure that digital images captured only soybean vegetation.

2.3 | Traits under evaluation

We recorded the following agronomic data in both years of the study: beginning bloom (R1), recorded as the DAP when 50% of the plants in a plot displayed an open flower at any main stem node; full maturity (R8), expressed as the days after 31 August when 95% of pods had reached mature color; flowering and maturity notes collected every 3 d, with back-and-forward scoring when the plots reached R1 and R8 between scoring dates; mean plant height, sampled at R8 by measuring

the distance from the soil surface to the highest point on the main stem for three randomly selected plants per plot; seed moisture and grain yield (g plot^{-1}) collected from the two-row plots at harvest; and grain yield converted to kilograms per hectare adjusted to 13% moisture.

We measured CC weekly using ground-based digital imagery (Purcell, 2000). Images were taken at the center of the plot using a Canon Powershot A4000 IS inclined 30° from the horizon and mounted to a monopod 1.5 m above the soil surface. The imagery was collected in the mornings, between 9:00 and 11:00 a.m., with one plot per image. To obtain a general assessment of the growth period, we performed phenological staging (Fehr, Caviness, Burmood, & Pennington, 1971) of the parent checks on the same day as CC measurements. In 2013, we collected images at 28, 35, 42, 50, 56, and 65 DAP, encompassing the developmental growth period V3–R4. A total of 38,400 images were recorded during the 2013 season. In 2014, we collected images at 20, 27, 34, 40, 48, and 56 DAP for a total of 38,400 images. These sampling dates spanned the developmental growth period V1–R3.

Automated batch analysis of digital images used the "Turf Analysis" macro for SigmaScan Pro 5.0 (SPSS) as described in Karcher and Richardson (2005) to distinguish soybean canopy from the image background when the appropriate threshold ranges for hue (0–255) and saturation (0–50) are selected. We performed the algorithm calibration, using random image samples, separately for each sampling date to account for variations in incident solar radiation (Karcher & Richardson, 2005).

2.4 | Spatial adjustment

To best represent the phenotypic values in unreplicated experimental designs, it was necessary to adjust the treatment observations for field variation by removing the spatial autocorrelation among plots through kriging (Banerjee, Finley, Waldmann, & Ericsson, 2010; Zas, 2006). Reproducing kernel Hilbert spaces fit a linear model based on two perceptron kernels, a linear kernel $\mathbf{G} = \mathbf{MM}'p^{-1}$, where $\mathbf{M}$ is the genotypic matrix coded as {−1, 0, 0} and p is the number of markers, that represented the genomic relationships among individuals (González-Camacho et al., 2012), and a Gaussian kernel $K = \exp(-\theta D^2)$ based on the exponential Euclidean distance (D) among treatments in the field (Xavier, Muir, Craig, & Rainey, 2016), normalized by the mean quadratic distance (θ). Thus, the spatial adjustment can be framed by the following linear mixed model:

$$y = \mu + u + s + e$$

where the response variable (y) is being modeled as a function of intercept (μ); genetics (u) assumed to be normally dis-

tributed, N(0, $G\sigma_a^2$), field plot variation (s) assumed to be normally distributed, N(0, $K\sigma_s^2$); and residuals (e) assumed to be normally distributed, N(0, $I\sigma_e^2$), where σ_a^2, σ_s^2, and σ_e^2 represent the genetic, spatial, and residual variances. In probabilistic terms, the phenotypes are assumed to follow a multivariate normal distribution as $y \sim N(1\mu, K\sigma_s^2 + G\sigma_a^2 + I\sigma_e^2)$. Subtracting the spatial coefficients from the primary observations provided adjusted values ($y^* = y - s$) used for all downstream analysis.

We fit the model in R using the package BGLR (Pérez & de los Campos, 2014). Our analysis evaluated bandwidth parameters (θ) of the Gaussian kernel, adjusting the degree to which neighboring plots were associated with the plot of interest at several levels to identify the most representative spatial correction. We then assessed the phenotypic repeatability of average canopy coverage (ACC) and grain yield over 2 yr of data to cross-validate our selection of bandwidth parameters (Supplemental Table S1).

2.5 | Longitudinal representation of canopy coverage

We fit an asymptotic logistic growth curve to the repeated measurements of CC for each plot using the generalized least squares (gnls) function of the nlme package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2015) in the R statistical language and environment (R Development Core Team, 2015). The form of the fitted model is presented by

$$Y = \frac{y_{max}}{1 + e^{b(DAP - c)}}$$

where Y is CC, y_{max} is maximum CC achieved, c corresponds to the time in DAP when maximum CC is achieved, and b is the slope or changing factor in the curve (Supplemental Figure S8). Fitting a logistic curve allows for derivation and calculation of several canopy growth parameters. The asymptote is the maximum coverage value estimated by the logistic model.

2.6 | Derived traits

Since CC has a linear relationship with daily LI, it is possible to estimate daily intercepted photosynthetically active radiation (PAR, MJ m^{-2} d^{-1}) for a given plot by taking the product of daily CC and total daily PAR (Purcell, 2000). We calculated daily PAR as one-half of the total solar radiation for a given day (Monteith, 1972). Total CIPAR (MJ m^{-2}) represented the sum of daily PAR measurements for the duration of the sampling period. We determined CIPAR for the vegetative phase (CIPARv) and the reproductive phase (CIPARr)

by adding the daily PAR from the initial sampling date to R1, and the daily PAR from R1 to the final sampling date, respectively. We used the logistic model for estimating t_i values to determine the number of days required for CC to reach 30% (t_{30}), 50% (t_{50}), and 70% (t_{70}) closure. Average canopy coverage (ACC) was computed by taking the mean value of daily CC from the initial to final sampling dates.

2.7 | Between- and within-family variance decomposition

We used an ANOVA in PROC GLM of SAS for the statistical analysis. A Fisher's protected LSD test at $\alpha = .05$ (FLSD$_{0.05}$) compared statistical differences among SoyNAM families. Due to differences in sampling dates, heteroskedasticity, and the stages sampled for canopy data, as well as the heterogeneity of variance between years, we performed ANOVA by year. Phenotypic correlations were calculated as pairwise Pearson correlation coefficients (r_{AB}) between any pair of traits A and B by year, implemented by the *cor* function in R (R Development Core Team, 2015).

Unbiased estimation of between- and within-family variance components fit linear mixed-effect models using restricted maximum likelihood (REML) through the Proc Mixed procedure in SAS version 9.4 (SAS Institute). Linear mixed-effect model estimates of parameter phenotypic repeatability at the family level (R_M) represented the total variance accounted for by differences among groups (Sokal & Rohlf, 1995), calculated as $R_M = \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\varepsilon^2)$, where σ_α^2 is the among-group variance and σ_ε^2 is the within-group variance as extracted from the linear mixed-effect model output that fit an intercept with individual identities, which were set as random (Nakagawa & Schielzeth, 2010).

2.8 | Estimation of heritability and genetic correlations on individual RILs

The following analyses were performed combining both 2013 and 2014 information. A Bayesian Gibbs sampler calculated the variance components through a multivariate mixed linear model (Sorensen & Gianola, 2002) using the software GIBBS3F90 (Misztal et al., 2002), which sampled regression coefficients from a multivariate normal distribution and (co)variance components from a scaled inverse-Wishart distribution, using the genomic relationship matrix (**G**) to define the additive relationship among all SoyNAM lines (Xavier, Hall, Casteel, Muir, & Rainey, 2017). The following mixed linear model was used to fit k traits simultaneously:

$$\mathbf{y}_k = \mathbf{X}_k\mathbf{b}_k + \mathbf{Z}_k\mathbf{u}_k + \mathbf{e}_k$$

where **y** is the vector of observations of the kth trait, $\mathbf{X}_k$ is the incidence matrix of year, $\mathbf{b}_k$ is the vector of regression coefficient of year effects, $\mathbf{Z}_k$ is the incidence matrix of genotypes, $\mathbf{u}_k$ is the polygenic effect associated with each line, and $\mathbf{e}_k$ is the residual term. Year was treated as fixed and genotypes were treated as random.

Using the derived genetic

$$\left(\mathbf{G}_0 = \mathbf{G} \otimes \begin{bmatrix} \sigma_{A_A}^2 & \cdots & \sigma_{A_{AB}} \\ \vdots & \ddots & \vdots \\ \sigma_{A_{AB}} & \cdots & \sigma_{A_B}^2 \end{bmatrix} \right)$$

and residual

$$\left(\mathbf{R}_0 = \mathbf{I} \otimes \begin{bmatrix} \sigma_{E_A}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{E_B}^2 \end{bmatrix} \right)$$

covariance matrices, we calculated narrow-sense heritability estimates (h^2) as

$$h^2 = \frac{\sigma_A^2}{(\sigma_E^2 + \sigma_A^2)}$$

for each trait and the additive genetic correlation ($r_{A_{AB}}$) as

$$r_{A_{AB}} = \frac{\sigma_{A_{AB}}}{\sqrt{\sigma_{A_A}^2 \sigma_{A_B}^2}}$$

between any pair of traits A and B, where σ_A^2 is the additive genetic variance, σ_E^2 is the environmental variance, and $\sigma_{A_{AB}}$ is the additive genetic covariance between the traits.

The efficiency of indirect selection (R_e) was calculated according to Falconer (1981) as

$$R_e = \frac{CR_A}{R_A} = \frac{r_{A_{AB}} h_A}{h_B}$$

where CR_A is the amount of improvement in the primary trait (A) obtained by indirect selection for a secondary trait (B), R_A is the amount of improvement obtained by direct selection for a primary trait, $r_{A_{AB}}$ is the additive genetic correlation between the primary trait and the secondary character, and h_A and h_B are the square roots of the narrow-sense heritability for the secondary and primary trait, respectively.

3 | RESULTS

3.1 | Initial sampling dates

Variation within individual sampling dates tended to be greater in 2013 than in 2014 (Supplemental Table S14). In both years, the greatest amount of variation occurred at the sampling date where the majority of plots were at or near the R1 growth stage (42 DAP in 2013 and 40 DAP in 2014). The
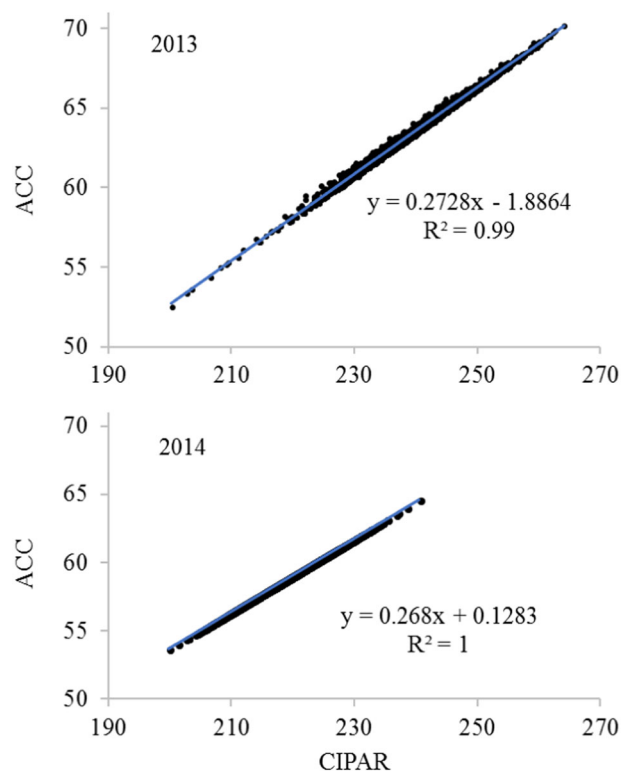
FIGURE 1 Relationship between average canopy coverage (ACC) and cumulative intercepted photosynthetically active radiation (CIPAR) for 2013 and 2014

ANOVA results from both years showed that CC values for each sampling date were significantly ($P < .0001$) controlled by genetic differences among families. However, as suggested by the larger R^2, the proportion of the phenotypic variance explained by the effect of family was greater in 2014 than in 2013. This is likely a result of the larger RMSE reported within families in 2013. While still low, the phenotypic correlation of grain yield to CC was consistently greater for the earlier season sampling dates in both years (Supplemental Tables S2 and S3). The large R^2 and low RMSE values suggest that the observed values of CC fit extremely well to the predicted values from the logistic fit model (Supplemental Table S4).

3.2 | ACC and CIPAR

Due to the nearly one-to-one relationship between ACC and CIPAR across all plots in both 2013 ($R^2 = .99$) and 2014 ($R^2 = 1$), we will discuss these parameters simultaneously (Figure 1). As a result of a later initial sampling date (and thus greater CC), mean ACC and CIPAR across the entire SoyNAM population were greater in 2013 than in the 2014 growing season (Supplemental Figures S2 and S3). Despite the inconsistency in values across the years, the relative pattern and ranking of ACC and CIPAR for the SoyNAM families

were very similar in 2013 and 2014 (Supplemental Tables S5 and S6, Supplemental Figures S2 and S3). For example, mean separation by $FLSD_{0.05}$ revealed that the ACC and CIPAR of 7 of the top 10 and 8 of the bottom 10 families were consistent across years. The high phenotypic family-level repeatability values for ACC ($R_M = 0.92$) and CIPAR ($R_M = 0.94$) (Table 1) further support the notion of consistency across years. As previously suggested by their one-to-one relationship, comparison of family rankings across ACC and CIPAR showed nearly the same order within each year, with slight differences occurring in 2013. These differences in 2013 are likely a result of rounding error when calculating CIPAR, as inconsistent rankings were typically a result of different mean family values at the tenth and hundredth decimal places. The ANOVA results revealed that ACC and CIPAR are highly influenced by genetic differences among families ($P < .0001$, Supplemental Table S15). As expected, due to the large among-family variance, both ACC ($h^2 = .85$) and CIPAR ($h^2 = .89$) proved to be highly heritable traits (Table 2), and thus selecting for these traits is likely to be effective.

The desirability of selecting for these canopy traits is highly dependent on associations with grain yield. The ACC and CIPAR phenotypically correlated with grain yield ($r = .27$) in both 2013 and 2014 (Supplemental Tables S7 and S8), whereas additive genetic correlations to grain yield across years for ACC ($r_A = .63$) and CIPAR ($r_A = .64$) were substantially greater (Table 2). On a within-family basis, the average phenotypic correlation between ACC and yield was $r = .33$ ($\pm .18$) in 2013 and $r = .31$ ($\pm .10$) in 2014. These results suggest that a large proportion of the variance shared among these traits and grain yield is genetically controlled.

Due to the moderate heritability of grain yield ($h^2 = .55$) (Table 2), the higher heritability of ACC and CIPAR, the strong additive genetic correlations between these parameters and grain yield, and the relative ease with which these canopy parameters can be measured, it is plausible that these traits can improve the efficiency of direct selection for grain yield when used as criteria for indirect selection (Fehr, 1987). Compared with direct selection for grain yield, indirect selection based on ACC and CIPAR had relative efficiencies of 0.78 and 0.81, respectively (Table 1). Because the additive genetic correlation between these parameters is 1, that difference is not meaningful and can be attributed to rounding error when calculating CIPAR in 2013.

3.3 | CIPARv and CIPARr

Previous studies have suggested that soybean grain yield is controlled by the amount of LI during specific developmental growth periods (Board & Kahlon, 2012; Board et al., 2011), and that dissection of highly quantitative traits into more heritable components helps reveal genetic and

TABLE 1 Phenotypic repeatability (R_M) and narrow-sense heritability (h^2), and efficiency of indirect selection (R_e) for grain yield using secondary traits, defined as the ratio of the amount of improvement in yield obtained by indirect selection for a secondary trait (CR_A) over the amount of improvement obtained by direct selection for grain yield (R_A)

Trait[a]	R_M	h^2	CR_A	R_A	R_e
CIPAR	0.92	.85	0.60	0.74	0.81
CIPARv	0.94	.88	0.63	0.74	0.84
CIPARr	0.99	.87	0.57	0.74	0.76
ACC	0.99	.86	0.58	0.74	0.78
t_{30}	0.97	.81	0.61	0.74	0.83
t_{50}	0.98	.89	0.62	0.74	0.84
t_{70}	0.98	.90	0.59	0.74	0.80

[a]CIPAR. cumulative intercepted photosynthetically active radiation; CIPARv, CIPAR for the vegetative period; CIPARr, CIPAR for the reproductive period; ACC, average canopy coverage; t_{30}, number of days to reach 30% closure; 5_0, number of days to reach 50% closure; t_{70}, number of days to reach 70% closure.

TABLE 2 Additive genetic correlations (above diagonal) and narrow-sense heritability (on the diagonal) for grain yield (Yld), cumulative intercepted photosynthetically active radiation (CIPAR), CIPAR for the vegetative period (CIPARv), CIPAR for the reproductive period (CIPARr), average canopy coverage (ACC), number of days to reach 30% (t_{30}), 50% (t_{50}), and 70% (t_{70}) closure, and days to maturity (R8)

Trait	Yld	CIPAR	CIPARv	CIPARr	ACC	t_{30}	t_{50}	t_{70}	R8
Yld	.55	.64	.67	.61	.63	−.68	−.66	−.63	.52
CIPAR	–	.89	.98	1.00	1.00	−.98	−1.00	−1.00	.30
CIPARv	–	–	.87	.96	.98	−1.00	−.99	−.97	.34
CIPARr	–	–	–	.86	1.00	−.95	−.99	−1.00	.28
ACC	–	–	–	–	.85	−.97	−1.00	−1.00	.29
t_{30}	–	–	–	–	–	.81	.99	.97	−.36
t_{50}	–	–	–	–	–	–	.89	.99	−.33
t_{70}	–	–	–	–	–	–	–	.90	−.29
R8	–	–	–	–	–	–	–	–	.73

Note. All terms are considered significant at $P < .0001$.

physiological mechanisms that account for trait variability, as well as their influence on grain yield (Tardieu & Tuberosa, 2010). To evaluate these concepts as they pertain to LI, we divided total CIPAR into two periods encompassing the vegetative (CIPARv) and reproductive (CIPARr) periods of soybean growth. It is notable that our estimates of CIPARr are lower than what is typically reported because most studies typically measure CIPARr until R6. Our studies stop at R4 due to concerns of low accuracy above ~90% canopy cover. The smaller values of CIPARr that we obtained influence interpretation of the relationship between CIPARr and grain yield, but our results are useful for identifying parameters for selection in breeding pipelines using high-throughput phenotyping of CIPAR, CIPARr, and CIPARv from image analysis, and especially using unmanned aerial systems.

In 2014, CIPARv accounted for a larger proportion of total CIPAR due to an earlier initial sampling date. Relative to CIPAR, phenotypic repeatability was greater for both CIPARv ($R_M = 0.99$) and CIPARr ($R_M = 0.99$) (Table 1). However, contrary to our expectations, heritability decreased slightly

for CIPARv ($h^2 = .87$) and CIPARr ($h^2 = .86$) (Table 2). Family rankings followed similar patterns in both years, with 7 and 8 of the top 10 performing families repeated across years for CIPARv and CIPARr, respectively (Supplemental Tables S9 and S10, Supplemental Figures S4 and S5). In both years, phenotypic correlations with grain yield were greater for CIPARv ($r = .35$, 2013; $r = .30$, 2014) and less for CIPARr ($r = .19$, 2013; $r = .23$, 2014) (Supplemental Tables S7 and S8). Additive genetic correlations with grain yield followed the same tendencies as phenotypic correlations relative to CIPAR; there was a stronger additive genetic correlation with CIPARv ($r_A = .67$) and a weaker one with CIPARr ($r_A = .61$) (Table 2). Both phenotypic and additive genetic correlations are metrics that describe the linear relationship between traits, which is an important assumption for studying associations from the breeding standpoint. Linearity was also used to describe the relationship between CIPAR and yield by Gaspar and Conley (2015). Evaluating the relative efficiency of these parameters for indirect selection of grain yield suggests that selecting for CIPARv ($R_e = 0.84$) would improve the

efficiency of indirect selection compared with selecting for CIPARr ($R_e = 0.76$) (Table 1). In light of similar estimates of heritability for these two parameters, the increased efficiency of indirect selection for CIPARv is due to the stronger additive genetic correlation with grain yield than that of CIPARr.

3.4 | Rate of canopy coverage

To further evaluate the influence of CC rate on grain yield, we determined the number of days required for each genotype to reach 30, 50, and 70% coverage for both years of the study. Mean values of t_{30}, t_{50}, and t_{70} were all higher in 2013 than in 2014, suggesting that canopies closed at a more rapid rate in 2014 (Supplemental Table S15). Mean separation of families by $FLSD_{0.05}$ revealed that 5, 7, and 9 of the top 10 fastest closing families were consistent across years for t_{30}, t_{50}, and t_{70}, respectively (Supplemental Tables S11–S13, Supplemental Figures S6–S8). The phenotypic consistency across years was further supported by the high repeatability ($R_M \geq 0.97$) reported for all three of these parameters (Table 1). As suggested by the large amount of among-family variance compared with within-family variance, heritabilities were high for t_{30} ($h^2 = .81$), t_{50} ($h^2 = .89$), and t_{70} ($h^2 = .90$) (Table 2). Phenotypic correlations to grain yield were greatest for t_{30} ($r = -.36$, 2013; $r = -.32$, 2014) and decreased thereafter for t_{50} ($r = -.32$, 2013; $r = -.29$, 2014) and t_{70} ($r = -.25$, 2013; $r = -.25$, 2014) (Supplemental Tables S7 and S8). The observed negative relationships suggest that greater grain yield was associated with a faster rate of closure. Additive genetic correlations with grain yield followed the same general trend. Genetic correlations with t_{30} ($r_A = -.68$) showed the strongest genetic association, and t_{70} ($r_A = -.63$), with the weakest with t_{50} ($r_A = -.66$) falling in between (Table 2). The relative efficiencies of indirect selection for t_{30} ($R_e = 0.83$), t_{50} ($R_e = 0.84$), and t_{70} ($R_e = 0.80$) suggest higher efficiencies seen during periods of early canopy development (t_{30} and t_{50}), which could be a result of their increased additive genetic correlations to grain yield (Table 1).

4 | DISCUSSION

All the calculated canopy parameters in this study were highly heritable ($h^2 = .81–.90$) and associated with grain yield both phenotypically ($r = .19–.36$) and genetically ($r_A = .61–.68$), which would provide correlated responses to selection in a breeding pipeline. In the same population, Xavier, Hall, Casteel, et al. (2017) applied unsupervised learning methods to investigate interaction among various traits including ACC, a canopy trait also evaluated in the present study. Although Xavier, Hall, Casteel, et al. (2017) derived ACC from a single-parameter logistic curve, the estimated heritability and additive genetic correlation between yield and ACC were consistent with our results based on a more generalized estimator of ACC, which was computed from a multiparameter logistic curve. To build upon the possible avenues of exploiting CC in soybeans, the current study exploits multiple canopy-derived traits with meaningful physiological basis for yield improvement.

An important result from this study is that parameters encompassing the vegetative growth period (CIPARv, t_{30}, and t_{50}) showed the highest phenotypic ($r = .29–.36$) and genetic ($r_A = .66–.68$) correlations with grain yield while still maintaining high heritability ($h^2 \geq .81$). These results suggest that rapid CC during early- season growth is highly variable among SoyNAM genotypes and has a greater influence on grain yield than does CC during the early reproductive periods. Ball, Purcell, and Vories (2000) indicated that early canopy development can be particularly important for short-season production systems. Jarquin et al. (2018) results also support the importance of early-season canopy as a good parameter to improve the predictive ability for yield in genomic selection. Both Xavier, Hall, Hearst, et al. (2017) and Kaler et al. (2018) identified common significant SNPs associated with early-season CC in soybean, and the former also showed that these SNPs lead to an increase in grain yield. These results cannot be directly compared with Koester et al. (2014) due to a lack of CC measurements beyond the mid-reproductive period. Nevertheless, it is interesting that although they found no difference in the speed of closure among cultivars, our results show wide variation in early-season canopy development within an unselected population.

Assessing the phenotypic correlations of CC from individual sampling dates, Hoyos-Villegas et al. (2014) showed negative correlations with grain yield during vegetative growth (33, 37, and 47 DAP) that generally increased with later sampling dates. However, our results showed the opposite trend, with the highest correlations with grain yield occurring during early vegetative periods and decreasing thereafter. These contrasting findings may be due to the yield-limiting treatments imposed by the Hoyos-Villegas et al. (2014) study. Our results may be an indicator of faster emergence or improved seedling vigor during these early vegetative periods; however, we did not collect data to assess this supposition.

The heritabilities reported by Jannink et al. (2000) for early LI ($h^2 = .19$) and mid-LI ($h^2 = .27$) were generated under a different measurement technique and provided much lower than any of the heritabilities reported in our study. Their results were based on single measurements of LI and therefore do not fully characterize the genotype's performance throughout the growing season. Furthermore, their study evaluated the weed suppressive ability of soybean genotypes, whereas we removed weeds, so it is possible that the influence of weed pressure on canopy development and LI caused a large

variance within each cross (family), thereby reducing heritability. It is important to note that the larger additive genetic correlation to weed suppressive ability that they found for early LI ($r_A = -.55$) compared with mid-LI ($r_A = -.34$) further supports the importance of rapid CC during early vegetative growth. Although our study and Xavier, Hall, Hearst, et al. (2017) used the same SoyNAM population, they found a lower heritability for ACC ($h^2 = .77$), which may be the result of the use of an extra year of observation and differences in calculating the variance components.

Although CIPAR in the vegetative period showed a stronger additive genetic correlation with grain yield, the heritability of CIPARv was slightly less than that of CIPAR. The CIPARv showed both a smaller heritability and additive genetic correlation to grain yield than CIPAR. Although this difference in additive genetic correlations may be due to differences in the important periods of LI for grain yield determination, we expected that the heritabilities associated with these divided periods of CIPAR would increase compared with the whole. These results may arise from insufficient phenological staging when dividing the periods. Furthermore, the divided periods did not encompass the entire vegetative and reproductive periods, and thus it was not possible to assess the full influence of LI on grain yield during each period of growth. Therefore, in this study, CIPAR may be a better indicator of a genotype's performance than CIPARv or CIPARr.

Ground-based digital imagery offers an efficient and precise method for evaluating differences in CC dynamics in genetically diverse populations. Furthermore, the observed differences provide insight into the physiological parameters responsible for grain yield variation among genetically distinct lines. Any of the canopy parameters in this study could be used as an indirect selection criterion, yet none of the canopy traits under evaluation reached an efficiency of indirect selection (>1) that would suggest it was equivalent to or more efficient than selecting directly for grain yield. This is unsurprising, as there are very few instances in which indirect selection for one secondary trait is superior to direct selection for grain yield (Gallais, 1984; Bernardo, 2010). Xavier, Hall, Hearst, et al. (2017) found that indirect selection for yield using ACC was more efficient than the direct selection for yield by 14%, probably due to (a) the use of one extra year of data (2015, collected from drones) for the genetic parameters' estimation and (b) the additional use of drove phenotypes for 2014, whereas the current study is based on ground imagery only.

By incorporating several traits known to influence grain yield (canopy cover, maturity, seed size, and seed number), index selection offers the advantage of selecting for several traits simultaneously on the basis of a single index value (Bernardo, 2010). Our results indicate that canopy traits are highly heritable and have significant additive genetic correlations with grain yield (Table 2), and therefore incorporating these traits into a single index value may benefit the indirect selection efficiency beyond that of selecting for grain yield directly (Hazel & Lush, 1942).

The additive genetic correlations between R8 and canopy parameters (Table 2) were just around half of the magnitude of their correlation to yield, and it is important to consider those correlations in a selection scheme. However, Xavier, Hall, Hearst, et al. (2017) found that the ACC QTL on chromosome 19, which was significant throughout the entire assessment period, had a positive effect on yield and did not increase R8. In addition, a graphical model among soybean traits showed that although yield and ACC had direct connection, there was no direct genetic link between ACC and maturity (Xavier, Hall, Casteel, et al., 2017).

The value in creating selection indices is especially important during the early stages of a soybean breeding pipeline, such as progeny row testing, when limited seed results in poor-quality data that provide an unreliable assessment of genotype performance in terms of grain yield and other quantitative traits. Creating an index of highly heritable traits known to have strong phenotypic and additive genetic correlations to grain yield result in a more informed selection of superior progeny and could possibly accommodate screening of more material.

Future studies of CC should expand the sampling period to encompass the entire growing season from emergence to complete senescence and include more environments. Much of the previous work on soybean CC suggests that the improvements in LI resulting from longer leaf duration is a primary contributor to the greater grain yield of newer elite cultivars (Koester et al., 2014; Richards, 2000). However, our results reveal a significant amount of genetic variation for the early canopy development that associates with greater grain yield. Further research that evaluates canopy development from emergence to harvest maturity would enable researchers to parse out the influence on soybean grain yield of both the early and late developmental periods. In addition, evaluating CC in more environments would improve the estimation of genetic parameters and result in robust conclusions. Identifying genotypes with superior canopy closure and canopy duration may enable the genetic characterization necessary for optimizing canopy development, LI, and grain yield. It would also be interesting to further evaluate the influence of interactions between genotypes and environment.

Selection experiments using the parameters described in this study as criteria are necessary for validating their efficacy in improving genetic gains in soybean. The heritabilities reported in this study suggest that selection of canopy parameters would be highly effective, but further assessment over several years and environments is necessary to confirm the applicability of these results. Should these parameters remain effective across selection trials, the next step is to further evaluate morphological and physiological traits

responsible for rapid canopy development. Seedling and pre-emergence traits are likely responsible for rapid seedling establishment and canopy development, such as long coleoptiles, broad seedling leaves, fast leaf expansion rate, and large seeds (Richards, 2000). Evaluating these traits in genotypes with consistently superior canopy development would provide breeders with further selection criteria for improving genetic gains.

AUTHOR CONTRIBUTIONS

B.P.H., F.M., and A.X. wrote the manuscript. B.P.H. and K.M.R. designed and implemented the experiment. B.P.H. and A.X. collected the imagery and conducted the statistical analysis. S.C. provided theoretical guidance of agronomic and physiological relevance throughout the study. K.M.R. provided guidance on breeding and genetics. F.M. contributed to the interpretation and breeding applications of imagery-based phenotypes.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Alencar Xavier https://orcid.org/0000-0001-5034-9954
Katy M. Rainey https://orcid.org/0000-0001-8541-5851

REFERENCES

Ainsworth, E. A., Yendrek, C. R., Skoneczka, J. A., & Long, S. P. (2012). Accelerating yield potential in soybean: Potential targets for biotechnological improvement. *Plant, Cell & Environment*, *35*, 38–52. https://doi.org/10.1111/j.1365-3040.2011.02378.x

Anand, S. C., & Torrie, J. H. (1963). Heritability of yield and other traits and interrelationships among traits in the F_3 and F_4 generations of soybean crosses. *Crop Science*, *3*, 508–511. https://doi.org/10.2135/cropsci1963.0011183X000300060015x

Ball, R. A., Purcell, L. C., & Vories, E. D. (2000). Optimizing soybean plant population for a short-season production system in the southern USA. *Crop Science*, *40*, 757–764. https://doi.org/10.2135/cropsci2000.403757x

Banerjee, S., Finley, A. O., Waldmann, P., & Ericsson, T. (2010). Hierarchical spatial process models for multiple traits in large genetic trials. *Journal of the American Statistical Association*, *105*, 506–521. https://doi.org/10.1198/jasa.2009.ap09068

Bartley, B. G., & Weber, C. R. (1952). Heritable and nonheritable relationships and variability of agronomic characters in successive generations of soybean crosses. *Agronomy Journal*, *44*, 487–493. https://doi.org/10.2134/agronj1952.00021962004400090008x

Bernardo, R. (2010). *Breeding for quantitative traits in plants* (2nd ed.). Woodbury, MN: Stemma Press.

Board, J., & Kahlon, C. (2012). A proposed method for stress analysis and yield prediction in soybean using light interception and developmental timing. *Crop Management*, *11*(1). https://doi.org/10.1094/CM-2012-0712-02-RS

Board, J. E., & Harville, B. G. (1992). Explanations for greater light interception in narrow- vs. wide-row soybean. *Crop Science*, *32*, 198–202. https://doi.org/10.2135/cropsci1992.0011183X003200010041x

Board, J. E., & Kahlon, C. S. (2011). Soybean yield formation: What controls it and how it can be improved. In H. A. El-Shemy (Ed.), *Soybean physiology and biochemistry* (pp. 1–36). London: InTech Open. https://doi.org/10.5772/17596

Board, J. E., Kamal, M., & Harville, B. G. (1992). Temporal importance of greater light interception to increased yield in narrow-row soybean. *Agronomy Journal*, *84*, 575–579. https://doi.org/10.2134/agronj1992.00021962008400040006x

Board, J., Kumudini, S., Omielan, J., Prior, E., Kahlon, C., & Weaver, D. (2011). Light interception as a yield-loss prediction tool for defoliation-induced yield loss in soybean. *Crop Management*, *10*(1). https://doi.org/10.1094/CM-2011-0722-01-RV

Bojacá, C. R., García, S. J., & Schrevens, E. (2011). Analysis of potato canopy coverage as assessed through digital imagery by nonlinear mixed effects models. *Potato Research*, *54*, 237–252. https://doi.org/10.1007/s11540-011-9189-y

Campillo, C., Prieto, M. H., Daza, C., Monino, M. J., & Garcia, M. I. (2008). Using digital images to characterize canopy coverage and light interception in a processing tomato crop. *HortScience*, *43*, 1780–1786. https://doi.org/10.21273/HORTSCI.43.6.1780

De Bruin, J. L., & Pedersen, P. (2009). New and old soybean cultivar responses to plant density and intercepted light. *Crop Science*, *49*, 2225–2232. https://doi.org/10.2135/cropsci2009.02.0063

Diers, B. W., Specht, J., Rainey, K. M., Cregan, P., Song, Q., Ramasubramanian, V., … Beavis, William D. (2018). Genetic architecture of soybean yield and agronomic traits. *G3: Genes, Genomes, Genetics*, *8*, 3367–3375. https://doi.org/10.1534/g3.118.200332

Edwards, J. T., Purcell, L. C., & Karcher, D. E. (2005). Soybean yield and biomass responses to increasing plant population among diverse maturity groups: II. light interception and utilization. *Crop Science*, *45*, 1778–1785. https://doi.org/10.2135/cropsci2004.0570

Falconer, D. S. (1981). *Introduction to quantitative genetics* (2nd ed.). London and New York: Longmans Green.

Fehr, W. R. (1987). *Principal of cultivar development: Theory and technique*. New York: Macmillan Publishing Company.

Fehr, W. R., Caviness, C. E., Burmood, D. T., & Pennington, J. S. (1971). Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. *Crop Science*, *11*, 929–931. https://doi.org/10.2135/cropsci1971.0011183X001100060051x

Gaspar, A. P., & Conley, S. P. (2015). Responses of canopy reflectance, light interception, and soybean seed yield to replanting suboptimal stands. *Crop Science*, *55*, 377–385. https://doi.org/10.2135/cropsci2014.03.0200

Gallais, A. (1984). Use of indirect selection in plant breeding. In W. Lange (Ed.), et al. *Efficiency in plant breeding. Proceedings of the 10th Congress of the European Association for Research on Plant Breeding* (pp. 45–60). Wageningen, the Netherlands: Pudoc.

Gizlice, Z., Carter, T. E., & Burton, J. W. (1994). Genetic base for North American public soybean cultivars released between 1947 and 1988. *Crop Science*, *34*, 1143–1151. https://doi.org/10.2135/cropsci1994.0011183X003400050001x

González-Camacho, J. M., De Los Campos, G., Pérez, ., Gianola, D., Cairns, J. E., Mahuku, G., … Crossa, J. (2012). Genome-enabled prediction of genetic values using radial basis function neural networks. *Theoretical and Applied Genetics*, *125*, 759–771. https://doi.org/10.1007/s00122-012-1868-9

Hanson, W. D., & Weber, C. R. (1962). Analysis of genetic variability from generations of plant-progeny lines in soybeans. *Crop Science*, *2*, 63–67. https://doi.org/10.2135/cropsci1962.0011183X000200010021x

Hazel, L. N., & Lush, J. L. (1942). The efficiency of three methods of selection. *Journal of Heredity*, *33*, 393–399. https://doi.org/10.1093/oxfordjournals.jhered.a105102

Hoyos-Villegas, V., Houx, J. H., Singh, S. K., & Fritschi, F. B. (2014). Ground-based digital imaging as a tool to assess soybean growth and yield. *Crop Science*, *54*, 1756–1768. https://doi.org/10.2135/cropsci2013.08.0540

Hunter, M. C., Smith, R. G., Schipanski, M. E., Atwood, L. W., & Mortensen, D. A. (2017). Agriculture in 2050: Recalibrating targets for sustainable intensification. *Bioscience*, *67*, 386–391 https://doi.org/10.1093/biosci/bix010

Jannink, J. L., Orf, J. H., Jordan, N. R., & Shaw, R. G. (2000). Index selection for weed suppressive ability in soybean. *Crop Science*, *40*, 1087–1094. https://doi.org/10.2135/cropsci2000.4041087x

Jarquin, D., Howard, R., Xavier, A., & Das Choudhury, S. (2018). Increasing predictive ability by modeling interactions between environments, genotype and canopy coverage image data for soybeans. *Agronomy*, *8*(4). https://doi.org/10.3390/agronomy8040051

Johnson, H. W., Robinson, H. F., & Comstock, R. E. (1955). Estimates of genetic and environmental variability in soybeans. *Agronomy Journal*, *47*, 314. https://doi.org/10.2134/agronj1955.00021962004700070009x

Kahlon, C. S., Board, J. E., & Kang, M. S. (2011). An analysis of yield component changes for new vs. old soybean cultivars. *Agronomy Journal*, *103*, 13–22. https://doi.org/10.2134/agronj2010.0300

Kaler, A. S., Ray, J. D., Schapaugh, W. T., Davies, M. K., King, C. A., & Purcell, L. C. (2018). Association mapping identifies loci for canopy coverage in diverse soybean genotypes. *Molecular Breeding*, *38*(5). https://doi.org/10.1007/s11032-018-0810-5

Karcher, D. E., & Richardson, M. D. (2005). Batch analysis of digital images to evaluate turfgrass characteristics. *Crop Science*, *45*, 1536–1539. https://doi.org/10.2135/cropsci2004.0562

Koester, R. P., Skoneczka, J. A., Cary, T. R., Diers, B. W., & Ainsworth, E. A. (2014). Historical gains in soybean (*Glycine max* Merr.) Seed yield are driven by linear increases in light interception, energy conversion, and partitioning efficiencies. *Journal of Experimental Botany*, *65*, 3311–3321. https://doi.org/10.1093/jxb/eru187

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., & Lee, D. H. (2002). BLUPF90 and related programs (BGF90). In *Proceedings of the 7th World Congress on genetics applied to livestock production* (pp. 1–2). Montpellier, France: Institut national de la recherche agronomique.

Monteith, J. L. (1972). Solar radiation and productivity in tropical ecosystems. *Journal of Applied Ecology*, *9*, 747–766. https://doi.org/10.2307/2401901

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews*, *85*, 935–956.

Pérez, P., & de los Campos, G. (2014). Genome-wide regression & prediction with the BGLR statistical package. *Genetics*, *198*, 483–495. https://doi.org/10.1534/genetics.114.164442

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Core Team, R. (2015). *nlme: Linear and nonlinear mixed effects models*. R package version 3.1-120. Vienna: Comprehensive R Archive Network.

Purcell, L. C. (2000). Soybean canopy coverage and light interception measurements using digital imagery. *Crop Science*, *40*, 834–837. https://doi.org/10.2135/cropsci2000.403834x

Purcell, L. C., Ball, R. A., Reaper, J. D., & Vories, E. D. (2002). Radiation use efficiency and biomass production in soybean at different plant population densities. *Crop Science*, *42*, 172–177. https://doi.org/10.2135/cropsci2002.1720

R Development Core Team. (2015). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLOS ONE*, *8*(6). https://doi.org/10.1371/journal.pone.0066428

Richards, R. A. (2000). Selectable traits to increase crop photosynthesis and yield of grain crops. *Journal of Experimental Botany*, *51*, 447–458. https://doi.org/10.1093/jexbot/51.suppl_1.447

Rincker, K., Nelson, R., Specht, J., Sleper, D., Cary, T., Cianzio, S. R., … Diers, B. (2014). Genetic improvement of US soybean in Maturity Groups II, III, and IV. *Crop Science*, *54*, 1419–1432. https://doi.org/10.2135/cropsci2013.10.0665

Shibles, R. M., & Weber, C. R. (1965). Leaf area, solar radiation interception and dry matter production by soybeans. *Crop Science*, *5*, 575–577. https://doi.org/10.2135/cropsci1965.0011183X000500060027x

Sneller, C. H. (1994). Pedigree analysis of elite soybean lines. *Crop Science*, *34*, 1515–1522. https://doi.org/10.2135/cropsci1994.0011183X003400060019x

Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: The principals and practice of statistics in biological research* (3rd ed.). New York: W.H. Freeman and company.

Song, Q., Yan, L., Quigley, C., Jordan, B. D., Fickus, E., Schroeder, S., … Cregan, P. (2017). Genetic characterization of the soybean nested association mapping population. *Plant Genome*, *10*(2). https://doi.org/10.3835/plantgenome2016.10.0109

Sorensen, D., & Gianola, D. (2002). *Likelihood, Bayesian, and MCMC methods in quantitative genetics*. New York: Springer.

Tardieu, F., & Tuberosa, R. (2010). Dissection and modelling of abiotic stress tolerance in plants. *Current Opinion in Plant Biology*, *13*, 206–212. https://doi.org/10.1016/j.pbi.2009.12.012

Wells, R. (1991). Soybean growth response to plant density: Relationships among canopy photosynthesis, leaf area, and light interception. *Crop Science*, *31*, 755–761. https://doi.org/10.2135/cropsci1991.0011183X003100030044x

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

Xavier, A., Hall, B., Casteel, S., Muir, W., & Rainey, K. M. (2017). Using unsupervised learning techniques to assess interactions among complex traits in soybeans. *Euphytica*, *213*, 200–2018. https://doi.org/10.1007/s10681-017-1975-4

Xavier, A., Hall, B., Hearst, A. A., Cherkauer, K. A., & Rainey, K. M. (2017). Genetic architecture of phenomic-enabled canopy coverage in *Glycine max*. *Genetics*, *206*, 1081–1089. https://doi.org/10.1534/genetics.116.198713

Xavier, A., Jarquin, D., Howard, R., Ramasubramanian, V., Specht, J. E., Graef, G. L., … Nelson, R. (2018). Genome-wide analysis of grain yield stability and environmental interactions in a multiparental soybean population. *G3: Genes, Genomes, Genetics*, *8*, 519–529. https://doi.org/10.1534/g3.117.300300

Xavier, A., Muir, W. M., Craig, B., & Rainey, K. M. (2016). Walking through the statistical black boxes of plant breeding. *Theoretical and Applied Genetics*, *129*, 1933–1949. https://doi.org/10.1007/s00122-016-2750-y

Zas, R. (2006). Iterative kriging for removing spatial autocorrelation in analysis of forest genetic trials. *Tree Genetics & Genomes*, *2*, 177–185. https://doi.org/10.1007/s11295-006-0042-4

SUPPORTING INFORMATION

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

How to cite this article: Hall BP, Xavier A, Moreira F, Casteel S, Rainey KM. Quantitative characterization of proximate sensing canopy traits in the SoyNAM population. *Crop Science*. 2020;1–12. https://doi.org/10.1002/csc2.20132