

Identification and QTL mapping of physiological drivers of soybean yield under contrasting management systems using high-throughput phenotyping

by

Race Heith Higgins

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Plant Breeding

Program of Study Committee:
Asheesh K. Singh, Major Professor
Thomas Lubberstedt
Maria Salas Fernandez
Andrew VanLoocke
Michael Castellano

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

Copyright © Race Heith Higgins, 2018. All rights reserved.

DEDICATION

This dissertation is dedicated to Tylor Schlader, my parents, and siblings, all who encouraged me to pursue my passions and finish my dissertation.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	vi
LIST OF TABLES	viii
NOMENCLATURE	x
ABSTRACT	xi
CHAPTER 1. GENERAL INTRODUCTION	1
Dissertation Organization	1
Brief Literature Review for each chapter	1
Physiological drivers of soybean traits.....	1
Biomass partitioning and residue quality	4
QTL mapping of vegetative indices	6
CHAPTER 2. SEEDING RATE INFLUCENCE ON YIELD COMPONENTS OF DIVERSE SOYBEAN AND DETAILED PHENTOYPING TO ELUCIDATE YIELD RESPONSE.....	9
Abstract.....	9
1. Introduction	10
2. Materials and methods.....	13
2.1. Plant Materials.....	13
2.2. Plant Measurements	14
2.3. Data and Analysis.....	15
3. Results	17
3.1. Genotype, Seeding rate, and Genotype x SR effects.....	18
3.2. Physiological drivers of soybean yield under contrasting seeding rates	19
4. Discussion.....	21
4.1. Genotype, Seeding rate, and Genotype x SR effects.....	21
4.2. Physiological drivers of soybean yield under contrasting seeding rates	23
5. Conclusions	26
Acknowledgments	27

References	27
Tables and figures.....	34
CHAPTER 3. UNCOVERING BIOMASS PARTITIONING AND RESIDUE QUALITY TRAITS FOR SOYBEAN IMPROVEMENT.....	45
Abstract.....	45
1. Introduction	46
2. Materials and methods.....	49
2.1 Plant Materials.....	49
2.2 Experimental Design	49
2.3 Plant Measurements	50
2.4 Data Analysis	51
3. Results	52
3.1 Biomass partitioning.....	52
3.2 Residue quality.....	53
4. Discussion.....	54
4.1 Biomass partitioning.....	54
4.2 Residue quality.....	55
5. Conclusion	56
Acknowledgments	58
References	58
Tables and figures.....	61
CHAPTER 4. LINKAGE MAPPING OF QTL FOR YIELD-RELATED VEGETATION INDICES IN A SOYBEAN NESTED ASSOCIATION MAPPING POPULATION	71
Abstract.....	71
Abbreviations.....	72
1. Introduction	72
2. Materials and Methods	75
2.1 Plant Materials and Field Trials	75
2.2 Phenotyping.....	76
2.3 Phenotypic Data Analysis	77
2.4 Linkage Map Construction and QTL Analysis	78

3. Results	79
3.1 Phenotypic Evaluation.....	79
3.2 Correlations Between Traits.....	80
3.3 Linkage Map Construction.....	80
3.4 QTL Analysis of Grain Yield and Vegetative Indices	81
4. Discussion.....	82
4.1 Correlations Between Traits and Yield	82
4.2 Yield and Vegetative Index QTL.....	83
4.3 Potential Application of QTL in Soybean Breeding	84
5. Conclusion.....	85
Acknowledgments	85
References	86
Tables and figures.....	92
CHAPTER 5. GENERAL CONCLUSIONS.....	107
REFERENCES	110
ACKNOWLEDGMENTS	116

LIST OF FIGURES

		Page
Figure 2.1	Boxplot distributions by environment and bar chart by seeding rate	37
Figure 2.2	Heatmap of Pearson correlations at three reproductive growth stages between physiological traits, agronomic traits, seed components, and final yield.	38
Figure 2.3	Adaptive elastic net yield prediction for three seeding rates of low, optimal, and high across reproductive growth stages	40
Figure 2.4	Venn diagram of shared and unique physiological traits selected by adaptive elastic net models at three seeding rates	42
Figure 2.5	Frequency histogram of plant emergence for three target seeding rates of low, optimal, and high in five replicated environments	43
Figure 2.6	Relationship between seed yield, seed oil percentage, and seed protein percentage for five locations.	43
Figure 2.7	Mean lodging scores for three seeding rates of low, optimal, and high and three ancestry categories of elite, diverse, and plant introduction.	44
Figure 2.8	Feature selection stability for 10 repeated 5 k-fold cross-validations of adaptive elastic net models.	44
Figure 3.1	Flow diagram of biomass sample collection and partitioning.....	62
Figure 3.2	Boxplots of a) biomass weight (g) by experiment year-location and growth stage and b) average C:N ratio by experiment year-location.	63
Figure 3.3	Biomass component dry weights by reproductive growth stage and genotype.	65
Figure 3.4	Biomass component percentages by reproductive growth stage and genotype.	66
Figure 3.5	Plot of a) weighted C:N ratios by biomass composition for all soybean genotypes of the final residue biomass excluding seed and b) mean seed yield of the previously ordered genotypes by C:N ratio.....	69
Figure 3.6	Curvilinear relationship between nitrogen harvest index and harvest index.....	70
Figure 4.1	Boxplot distribution of vegetation indices and yield by environments (n=2-3) and soybean growth stages.....	93

Figure 4.2	Heatmap of Pearson’s correlations between vegetation index and yield BLUPs by growth stage.....	97
Figure 4.3	Bar chart of Pearson’s correlations between vegetative index BLUPs and grain yield BLUPs	98
Figure 4.4	Genetic linkage map of chromosomes 1, 3, 10, and 18 with QTL for vegetation indices and yield	100
Figure 4.5	Genetic linkage map of soybean chromosomes 1-20 with marker names.	101

LIST OF TABLES

	Page
Table 2.1	Origin, population number, and maturity grouping of the soybean NAM parental genotype subset assessed in this study. 34
Table 2.2	Summary of vegetation indices with abbreviation, general class type, and source..... 35
Table 2.3	Yield, seed components, and agronomic trait summaries with count, mean, standard deviation, range, and Broad-sense heritabilities. 35
Table 2.4	ANOVA of yield, seed components, and agronomic traits of five environments and three replications by genotype. 36
Table 2.5	ANOVA of yield, seed components, and agronomic traits of five environments and three replications by ancestry..... 36
Table 2.6	Summary data of physiological traits by growth stage 39
Table 2.7	List of physiological traits selected by adaptive elastic net regularized regression..... 41
Table 3.1	Origin, population number, and maturity grouping of the soybean NAM parental genotype subset assessed in this study. 61
Table 3.2	Mean, range, standard deviation, ANOVA genotype factor significance, and heritability for biomass weights and percentage for three growth stages.... 64
Table 3.3	Pearson’s correlations and significance for biomass percentages, total biomass weights, and agronomic traits at growth stages R1, R4, and R8. 67
Table 3.4	Pearson’s correlations and significance for seed weight and harvest index with biomass percentages at growth stages R1, R4, and R8 and agronomic traits. 68
Table 3.5	Mean, range, standard deviation, ANOVA genotype factor significance, and heritability for C:N ratios for three growth stages 68
Table 4.1	Summary of parental genotype information..... 92
Table 4.2	Summary of selected vegetation indices with abbreviation, general class type, equation, and source. 92
Table 4.3	Summary of vegetation indices and yield 94
Table 4.4	Analysis of Variance and broad-sense heritabilities for vegetation indices and yield by growth stage..... 95

Table 4.5	Pearson's correlations and p-values of vegetation index and yield BLUPs by growth stage,.....	96
Table 4.6	Summary of soybean chromosomes	99
Table 4.7	Summary of QTL identified for five vegetation indices and yield	105
Table 4.8	List of 41 QTL reported for traits related to vegetation indices on chromosomes 1, 3, 10, and 18	106

NOMENCLATURE

C:N ratio	Carbon:nitrogen ratio
GY	Grain yield
H^2	Broad-sense heritability
MAS	Marker-assisted selection
NAM	Nested association mapping
QTL	Quantitative trait locus/loci
RIL	Recombinant inbred line
SoyNAM	Soybean nested association mapping population
SR	Seeding rate
VI	Vegetation index

ABSTRACT

Rapid characterization of physiological traits driving yield are becoming desirable aides to breeding programs to increase the rate of genetic gain. Each chapter in this dissertation investigates areas related to high-throughput phenotyping and physiological traits driving soybean yield. Chapter 2 seeks to understand the response of diverse soybean germplasm to seeding rate. An evaluation of final plot seed yield, seed protein percentage, seed oil percentage, seed weight, height, maturity, and plant lodging revealed a significant genotype x seeding rate interaction only for lodging, suggesting current soybean germplasm and soybean of wide genetic ancestry respond similarly to seeding rate. Our second objective was to identify physiological traits at multiple growth stages predicting yield response under contrasting levels of seeding rate. Adaptive elastic net models characterized diverging traits between seeding rates and determined chlorophyll traits as the leading predictors across seeding rates. Chapter 3 quantifies biomass partitioning strategies and residue quality determined through carbon:nitrogen (C:N) ratios in the same diverse panel of SoyNAM genotypes in Chapter 2. Above-ground plant components were dissected at three reproductive stages and revealed significant differences in biomass partitioning by R4. Significant genetic variation in C:N residue quality was found with no apparent negative relationship to final grain yield. Optimal biomass partitioning strategies for yield and improved residue C:N ratios for whole-system nitrogen sustainability can be targeted for yield improvement. Lastly, Chapter 4 includes a QTL mapping study of vegetative indices used for yield prediction in Chapter 1 in four SoyNAM RIL populations derived from five of the 32 parent NAM genotypes evaluated in Chapters 1 and 2. Five QTL were detected for grain yield and vegetative indices NDVI, NMDI, NWIB, PSRI, and VREI2 measured at R5,

spanning chromosomes 1, 3, 10 and 18. These QTL can serve as aides to MAS in soybean breeding and inform future studies aimed at dissecting the physiology of soybean grain yield. The overall research provides insights on soybean biomass partitioning and evidence of the presence of genetic variation in residue traits; physiological traits to predict yield in diverse germplasm and row-density management systems; and genomic regions mapped to spectral wavelengths related to soybean seed yield.

CHAPTER 1. GENERAL INTRODUCTION

Dissertation Organization

This dissertation is organized into five chapters. Chapter one provides a brief literature review focused on the three individual experiments related to high-throughput phenotyping and physiological traits related to soybean yield. Chapters two, three and four describe original research and are written in manuscript format for submission to scientific journals. Chapter two explores agronomic trait responses to seeding rate in diverse soybean and identifies the underlying physiological traits driving yield under different seeding rates. Chapter three investigates above-ground biomass partitioning strategies for yield and residue quality determined through carbon:nitrogen (C:N) ratios for a whole-systems nitrogen sustainability approach. Chapter four includes a QTL linkage study of vegetative indices related to soybean yield calculated from remote sensing in a four RIL soybean nested association mapping population. Lastly, chapter five is a summary of general conclusions. The knowledge generated from characterization of diverse soybean germplasm in row spacing treatments enables soybean breeders to utilize the opportunities for more informed selection with the use of physiological drivers as additional predictor traits in early generation of population development and selection; residue traits as additional economic opportunity and selection target and spectral indices QTL for marker assisted selection.

Brief Literature Review for each chapter

Physiological drivers of soybean traits

Soybean [*Glycine max* L. (Merr.)] is globally one of the most widely grown crops, with the United States as the world's leading soybean producer and the second-leading exporter (USDA-NASS, 2016). Expanding soybean production will continue interest in

management decisions such as seeding rate for optimal economic return (Fawcett et al., 2017). In response, plant breeders are challenged to create soybean cultivars adapted to new environments (Chang et al., 2015; Zanon et al., 2016) and management practices (Wilson et al., 2014; Grassini et al., 2015). Genetic gain and adaptation is driven by genetic diversity, but incorporating diverse germplasm into elite breeding programs presents challenges (Duvick, 1984). How germplasm will respond to various modern management practices such as seeding rate will need to be addressed to facilitate future use in breeding programs.

One stress plants must overcome in any production environment is interplant competition for needed resources like water, light, nutrients, and use of environmental resources (Weiner and Thomas, 1986). Agronomic practices, such as seeding rate, determine plant spatial arrangement and inter-plant stress, and soybeans respond to these conditions with phenotypic plasticity. The observed phenomenon of phenotypic plasticity is the ability of one genotype to produce more than one phenotype when exposed to different environments. Several studies have confirmed phenotypic plasticity in soybean, both at vegetative and reproductive stages (Egli, 1993; Carpenter and Board, 1997; Green-Tracewicz et al., 2011), and specifically to seeding rate (Elmore, 1998; Board, 2000; Vega, 2000). Soybean's recorded phenotypic plasticity to altered environments has traditionally made it difficult to determine optimal seeding rates, as soybeans can produce relatively similar yields from large changes in seeding rate (Carpenter and Board, 1997; Board, 2000; Edwards et al., 2005). To achieve a clear understanding of soybean's yield response to seeding rate requires identifying the underlying causing physiological traits.

However, several physiological traits respond to various levels of seeding rate. Under increased seeding rates, De Bruin and Pedersen (2009) found soybean yield has not

responded to increased plant density over time, but attributed genetic gain in newer cultivars to greater light use efficiency. Lee et al. (2008) found achieving complete plant canopy cover, and therefore maximizing light interception, at R1 produced maximum yield in standard agronomic practices of optimal planting density in their experiment. At lower plant populations, plant architecture is an important yield component and genetically determined. Plant branching capacity under stress was an important yield determinant (Frederick et al., 2001; Ferreira et al., 2016), and branching under low seeding rates was related to increased yield (Suhre et al., 2014; Agudamu et al., 2016). Studies in general have addressed seeding rate response to elite lines and their recent modern-adapted predecessors, leaving a gap in knowledge for seeding rate response in diverse or un-adapted germplasm.

Numerous physiological traits in conjunction forecast soybean yield, and high-throughput phenotyping platforms have enabled data collection on a greater spatial and temporal scale (Araus and Cairns, 2014). Remote sensing is a promising tool that rapidly and non-destructively collects phenotypes correlated to yield, and vegetative indices related to chlorophyll content, carotenoids, vegetation, water content, and dry matter content are used in soybean to predict yield (Ma et al., 2001; Bolton and Friedl, 2013; Johnson, 2014) and measure plant response to stress (Carter, 1994; Nutter Jr et al., 2002; Huang et al., 2016). In addition to vegetative indices, physiological traits leaf area index (LAI), mean tilt angle (MTA), intercepted photosynthetically active radiation (iPAR), and estimated chlorophyll content with a SPAD meter have been shown to correlate with soybean yield (Ma et al., 1995; Board and Harville, 1996) and can be collected in a rapid manner that complements remote-sensing. Despite methods of more attainable multi-trait collection, it is unknown how

the contribution of these traits together, over experiments and plant development stages, would explain yield response under multiple seeding rates.

Biomass partitioning and residue quality

Decisions regarding crop choice revolve around profitability for the farmer, which is influenced not singularly by a single season's crop output, but by impact on the yield of the succeeding crop seasons as well. It is important to include whole systems approach in research studies, but as a first step, individual components including genotype responses and traits need to be measured particularly for seed yield and other economic parts. Soybean output has improved with greater soybean seed yields, and seed yield and its yield potential is described by its harvest index, the ratio of seed yield to total biomass. The maximum harvest index from small grains is estimated to be 0.6, and it was found that modern US Midwestern soybean cultivars partitioned 60% of its biomass into seed, indicating that breeding has already succeeded in maximizing harvest index in soybean (Zhu et al., 2010). Pedersen and Lauer, (2004) found the harvest index for soybean ranged at 56.2 – 58.0% for elite soybean cultivars in the Midwest, additionally implicating harvest index has been maximized. Increased harvest index in soybean has resulted from increased seed yield with little increase in total aboveground biomass (Morrison et al., 2000; Jin et al., 2010), but the relationship between the composition of the total aboveground biomass to grain yield is unknown. Srinivasan et al. (2016) showed evidence that modern crop genotypes produce more leaf than is optimal, and removing leaves resulted in an 8% increase in yield. Identifying genotypes with optimal biomass partitioning strategies could further advance soybean line development where harvest index has been maximized and finding the reproductive growth stage when differences in biomass partitioning first become evident can provide insight to when genotypes begin to physiologically diverge.

Complementing increasing soybean output for greater season profitability and decreasing input costs through reduced fertilizer requirements of the succeeding crop can be achieved by improving soybean residue quality. Nitrogen (N), is essential for plant growth and seed production (Lawlor, 2002) and is the main component of fertilizer. Ubiquitous maize-soy cropping system in the US Midwest may particularly benefit, as improving the amount and quality of the soybean residue can contribute to the yield increase of maize in the succeeding season (Green and Blackmer, 1995) through increasing plant available nitrogen. Gentry et al. (2013) found net soil nitrogen mineralization was the strongest predictor of yield difference in continuous corn systems, where net mineralization of soil nitrogen is influenced by both quality (C:N ratio) and quantity of residue from the previous crop (Gentry et al., 2001). We hypothesize that soybeans may be further adapted to the maize-soy rotation by improving the carbon-nitrogen (C:N) ratio in soybean residue, without penalizing grain yield. Genotypic differences in C:N were found within multiple crop species between wild and domesticated crops (García-Palacios et al., 2013), and significant genotypic variations in stem nitrogen traits at maturity were found in modern soybean cultivars (Fritschi et al., 2013). Dhanapal et al. (2015) further supported finding C:N ratio variation in a collection of 373 soybean genotypes at flowering (R2). A caveat to lower C:N ratios is the concern that higher amounts of nitrogen in the vegetative plant organs results in less N remobilization to the seed during grain fill, suggesting a yield penalty, but multiple studies have shown that direct nitrogen uptake and accumulation during seed fill could be a more important factor for high seed yield instead of N remobilization (Kumudini et al., 2001; Zhao et al., 2014). Unknown are genotype-specific C:N ratios of the whole composite residue and its

relationship with soybean yield, along with two important seed quality factors, seed protein and seed oil content.

QTL mapping of vegetative indices

Soybean grain yield is a complex quantitative trait, and maximum yield potential is constrained by physiological and agronomic traits such as light interception, photosynthetic capacity, and biomass partitioning (Monteith and Moss, 1977). Historically, the steady increase in soybean grain yield has been attained through empirical selection for grain yield over the past century. However, there is evidence that phenotyping for physiological traits, as a complement to agronomic traits, may help in identifying selectable features that accelerate breeding for yield potential (Araus and Cairns, 2014; Keep et al., 2016). Currently, the soybean genetic base is narrow with low diversity, due to a genetic bottleneck after introduction to the US (Rincker et al., 2014). Introgressing exotic germplasm into cultivars to increase genetic diversity within domesticated crops has been used to enhance complex traits such as yield (Tanksley and McCouch, 1997) and may have unknowingly introduced novel genetic variation for yield-related physiological traits. In soybean, Thompson and Nelson (1998) tested experimental lines derived from crossing North American cultivars with several plant introductions, and several of these lines were incorporated into the soybean nested association mapping (SoyNAM) parent panel, including LG90-2550 and LG94-1128 of this study. An experimental population of high-yielding elite lines is enriched by including lines of diverse ancestry because it increases morphological and genetic diversity.

Many changes in morphological and physiological traits in soybean have accompanied changes in grain yield. Potential soybean yield is closely associated with plant photosynthesis (Slattery et al., 2017) and chlorophyll concentration affects photosynthetic capacity and primary production, the rate at which a crop can capture and store chemical

energy (Gitelson et al., 2003; Koester et al., 2016). Changes in leaf relative water content affect total water potential, osmotic potential, and turgor pressure, and therefore influence whole-plant physiology. Only when there is sufficient turgor pressure can cells expand for vegetative growth and stomata to open to incorporate carbon dioxide to be used in the Calvin cycle. (Zygielbaum et al., 2012; Gray et al., 2016). Canopy water content is indicative of canopy transpiration and determines radiation use efficiency and biomass accumulation in soybean (Saryoko et al., 2018). Biomass accumulation has long been established as important driver of potential yield, and is often estimated at leaf area index, or LAI (Ma et al., 1995; Board and Harville, 1996). Although genetic improvement of physiological traits can certainly lead to increased grain yield (GY), high-throughput, nondestructive measurements are necessary to rapidly collect many phenotypes for large mapping populations.

Remote sensing is a promising tool that rapidly and non-destructively collects vegetative indices related to chlorophyll content, carotenoids, vegetation, water content, and dry matter content that are used in soybean to predict yield (Ma et al., 2001; Bolton and Friedl, 2013; Johnson, 2014) and measure plant response to stress (Carter, 1994; Nutter Jr et al., 2002; Huang et al., 2016). An enormous number of spectral reflectance indices have been created to monitor vegetation health and productivity (Heinrich et al., 2011). Some indices have served as the industry standard for analyzing canopy “greenness” and detection of vegetation, such as the normalized difference vegetation index (Rouse Jr et al., 1974). However, many different indices have been developed depending on the specific trait to be monitored, and great advances in remote and proximal sensing technologies are currently underway. One advance has been the development of hyperspectral reflectance instruments (Haboudane et al., 2004). The major advantage of hyperspectral reflectance is that it allows

users to calculate any number of desired spectral reflectance indices pertinent to a trait of interest (Heinrich et al., 2011).

Quantitative trait loci (QTL) mapping is a key approach for understanding the genetic architecture of yield components and physiological traits in crops. However, pinpointing QTL can be hampered by relatively large QTL intervals due to the limited number of markers. Nested association mapping is an alternative population design that was proposed to increase the resolution of QTL mapping (Yu and Buckler, 2006). Nested association mapping populations are developed by crossing multiple diverse founders to a common parent followed by the development of recombinant inbred lines (RILs) or progenies in each family. In comparison to traditional QTL mapping, which only uses limited genetic information from two parents, NAM can increase genetic variation across contributing parental lines, increase genetic resolution, reduce linkage disequilibrium, and control population structure through design (Rafalski, 2010). The NAM design has been used successfully in soybean to map QTL controlling a number of traits such as grain yield stability (Xavier et al., 2018) and canopy coverage (Xavier et al., 2017). Rapid collection and identification of physiological predictors driving yield is desirable to breeding programs because they are used as breeding aides.

CHAPTER 2. SEEDING RATE INFLUENCE ON YIELD COMPONENTS OF DIVERSE SOYBEAN AND DETAILED PHENOTYPING TO ELUCIDATE YIELD RESPONSE

A paper in preparation for submission to *Field Crops Research*

R.H. Higgins, A.K. Singh

Highlights

- Soybean genotypes of diverse ancestry respond similarly to seeding rate in terms of yield, seed components, and agronomic traits.
- Yield prediction was highest when physiological trait predictors were measured over multiple reproductive growth stages, enabled through high-throughput phenotyping.
- Selected physiological traits for future genetic improvement among and within seeding rate treatments were identified through adaptive elastic net feature selection and ranking.

Keywords

Soybean; Seeding rate; Seed yield; High-throughput phenotyping; Remote sensing; Adaptive elastic net

Abstract

Seeding rate impact on soybean [*Glycine max* L. (Merr.)] yield has steadfastly remained a subject of agronomic research for decades. Expanding soybean production sustains interest in management decisions like seeding rate, prompting incorporation of soybean diversity for future adaptation and genetic gain. However, there is limited information on the response of diverse soybean germplasm to seeding rate. The first

objective of this study was to investigate how final plot seed yield, seed protein percentage, seed oil percentage, seed weight, height, maturity, and plant lodging responds to seeding rate in germplasm of diverse ancestries. Our second objective was to identify physiological traits and the corresponding growth stage window predicting yield response under multiple levels of seeding rate. A subset of the diverse SoyNAM parent panel consisting of 32 genotypes was evaluated under three levels of seeding rate in standard yield plots. Replicated RCBD experiments were grown in five environments in Central Iowa in 2014-2015. Physiological traits of remote sensing vegetative indices, leaf area index (LAI), mean tilt angle (MTA), intercepted photosynthetically active radiation (iPAR), and SPAD were measured in three environments in 2015 over three reproductive growth stages of flowering (R1-2), pod development (R3-4), and seed development (R5-6). Here we report a significant genotype x seeding rate interaction was only detected for lodging, and not for yield, seed weight, seed oil percentage, seed protein percentage, height, or maturity. These results suggest that current soybean germplasm and soybean of wide genetic ancestry respond similarly to seeding rate and implies introgression of diverse material may not detrimentally affect yield in response to seeding rate variations. In addition, physiological traits predicting the yield response among and within seeding rates were selected and ranked by adaptive elastic net, with chlorophyll traits determined as the leading predictors across seeding rates. The characterization of diverging traits between the seeding rate yield models provides targets for soybean improvement for varied seeding rate practices.

1. Introduction

Soybean [*Glycine max* L. (Merr.)] is globally one of the most widely grown crops, with the United States as the world's leading soybean producer and the second-leading exporter (USDA-NASS, 2016). Expanding soybean production will continue interest in

management decisions such as seeding rate for optimal economic return (Fawcett et al., 2017). In response, plant breeders are challenged to create soybean cultivars adapted to new environments (Chang et al., 2015; Zanon et al., 2016) and management practices (Grassini et al., 2015; Wilson et al., 2014). Genetic gain and adaptation is driven by genetic diversity, but incorporating diverse germplasm into elite breeding programs presents challenges (Duvick, 1984). How germplasm will respond to various modern management practices such as seeding rate will need to be addressed to facilitate future use in breeding programs.

Interplant competition introduces stress environments for plants due to competition for resources: water, light, nutrients, and environmental resources (Weiner and Thomas, 1986). Agronomic practices, such as seeding rate, determine plant spatial arrangement and inter-plant stress, and soybeans respond to these conditions with phenotypic plasticity, the ability of one genotype to produce more than one phenotype when exposed to different environments. Several studies have confirmed phenotypic plasticity in soybean, both at vegetative and reproductive stages (Carpenter and Board, 1997; Egli, 1993; Green-Tracewicz et al., 2011), and specifically to seeding rate (Board, 2000; Elmore, 1998; Vega, 2000). Soybean's recorded phenotypic plasticity to altered environments has traditionally made it difficult to determine optimal seeding rates, as soybeans can produce relatively similar yields from large changes in seeding rate (Board, 2000; Carpenter and Board, 1997; Edwards et al., 2005). To achieve a clear understanding of soybean's yield response to seeding rate requires identifying the underlying causing physiological traits.

However, several physiological traits respond to various levels of seeding rate. Under increased seeding rates, De Bruin and Pedersen (2009) found soybean yield has not responded to increased plant density over time, but attributed genetic gain in newer cultivars

to greater light use efficiency. Lee et al. (2008) found achieving complete plant canopy cover, and therefore maximizing light interception, at R1 produced maximum yield. At lower plant populations, plant architecture is an important yield component and genetically determined. Plant branching capacity under stress was an important yield determinant (Ferreira et al., 2016; Frederick et al., 2001), and branching under low seeding rates was related to increased yield (Agudamu et al., 2016; Suhre et al., 2014). Studies in general have addressed seeding rate response to elite lines and their recent modern-adapted predecessors, leaving a gap in knowledge for seeding rate response in diverse or un-adapted germplasm.

Numerous physiological traits in conjunction forecast soybean yield, and high-throughput phenotyping platforms have enabled data collection on a greater spatial and temporal scale (Araus and Cairns, 2014). Remote sensing is a promising tool that rapidly and non-destructively collects phenotypes correlated to yield, and vegetative indices related to chlorophyll content, carotenoids, vegetation, water content, and dry matter content are used in soybean to predict yield (Bolton and Friedl, 2013; Johnson, 2014; Ma et al., 2001) and measure plant response to stress (Carter, 1994; Huang et al., 2016; Nutter Jr et al., 2002). In addition to vegetative indices, physiological traits leaf area index (LAI), mean tilt angle (MTA), intercepted photosynthetically active radiation (iPAR), and estimated chlorophyll content with a SPAD meter have been shown to correlate with soybean yield (Board and Harville, 1996; Ma et al., 1995) and can be collected in a rapid manner that complements remote-sensing. Despite methods of more attainable multi-trait collection, it is unknown how the contribution of these traits together, over experiments and plant development stages, would explain yield response under multiple seeding rates.

The objective of this investigation was to determine how germplasm of diverse ancestries responds to seeding rate, by collecting final plot seed yield, seed protein percentage, seed oil percentage, seed weight, height, maturity, and plant lodging. Our second objective was to identify the physiological traits and corresponding growth stage window predicting yield response under multiple levels of seeding rate. Here we report the results of an ANOVA analysis on a subset of diverse genotypes from the SoyNAM parent population under three seeding rates ranging from low to high. We further report which physiological traits of remote sensing vegetative indices, LAI, MTA, iPAR, and SPAD were identified as predictors of yield under three levels of seeding rate and three reproductive growth stages. We aim to further understanding of soybean response to management practices to provide the breeding and research community tools for yield improvement.

2. Materials and methods

2.1. Plant Materials

For this study, 32 of 41 parent genotypes of the soyNAM panel were selected to be maturity groups II-III adapted to Central Iowa (Table 2.1). This panel was originally designed to include a diverse range of germplasm (Stupar and Specht, 2013). Specifically, genetic ancestry was elite public lines, high-yielding lines of diverse ancestry, or direct plant introductions. As RIL populations are already developed, further characterization of the parent panel provides a valuable resource, as the genomic and multi-environmental data for the soybean nested association mapping dataset is public (Xavier et al., 2015).

Experiments were grown in Central Iowa at a total of five environments at Iowa State University's experiment stations farm network sites. Planting dates in 2014 were 29 May and 13 May at Worle (41.99, -93.69) and Milo (41.35, -93.40), respectively; while in 2015, experiments were established on 13 May at Milo (41.34, -93.40), 21 May at Lippert (42.04, -

93.73), and 25 May at Bruner (42.01, -93.73). Plots were four rows with 0.76 m row width and 4.6 m long rows. Three fixed treatments of seeding rate, low= 20k plants ha⁻¹, medium=57k plants ha⁻¹ near commercial seeding rate, and high=93k plants ha⁻¹ were planted in a randomized complete block design with three replications at each location. Seeds were treated with ApronMaxx® RTA® fungicide treatment to protect the seed during germination, and plant emergence was measured for the two middle 0.91 m sections of each yield plot.

2.2. Plant Measurements

Soybean seed yield and seed components protein percentage, oil percentage, and seed weight were recorded for each location in 2014 and 2015. Approximately three hundred grams of whole soybean seed was used to quantify seed protein and oil contents using near-infrared reflectance (NIR) spectroscopy (Infratec™ 1241 Grain Analyzer, FOSS) and seed weights were averaged over a hundred-seed count weight. Before harvest at R8, agronomic traits of plant height, lodging (score 1-5) and final plant maturity (days after planting, DAP) were collected in all environments, except for two environments missing maturity and one environment missing seed weight. Lodging was recorded on a scale from 1-5, with 1 as upright and 5 as prostrate. Height was the average of two plants representative of the entire plot from the middle two rows. Grain yield (GY) was harvested at each location for the middle two rows of each plot with a two-row ALMACO plot combine. GY was determined as weight of grain harvested per unit area (kg/ha).

High-throughput phenotyping trait physiological data were collected as non-destructive repeated measures for three environments in 2015. The middle two rows of each plot were phenotyped at three soybean reproductive growth stages: flowering (R1-2), pod set (R3-4), and seed fill (R5-6) per Fehr et al., (1971). Remote-sensing data at R1-2 was not

recorded for one environment (Milo 2015), therefore this timepoint was excluded from further analysis. All light measurements were performed at 1000 h to 1400 on clear days. Leaf chlorophyll content was estimated using a Minolta SPAD-502 Plus chlorophyll meter with data logger and sampled non-destructively on ten fully-expanded trifoliates in the upper canopy. Intercepted photosynthetically active radiation (iPAR) was measured with a LI-191R line quantum sensor (Li-Cor, Inc., Lincoln, NE) with below-canopy measurements transecting the alley between the middle rows. iPAR was calculated as:

$$F = \left(1 - \frac{I_0}{I_t}\right) \times 100\%$$

where F is the fractional amount of radiation interception, I_0 is the measured incident PAR on the surface of the ground, and I_t is the radiant flux density on top of the canopy. LAI and MTA were collected with a LAI-2200C plant canopy analyzer (Li-Cor, Inc., Lincoln, NE), that simultaneously logged iPAR. A single above-canopy measurement with four to six below-canopy measurements were made along spatially partitioned diagonal transects between the middle two rows. A canopy gap test was performed as directed in the LAI-2200C Instruction Manual and a minimum apparent clumping factor of 0.95 determined the view cap size and number of below-canopy readings. Absolute reflectance remote sensing data were measured using with a FieldSpec® 4 Hi-Res (ASD Inc., Boulder, CO), which ranges from 350 to 2500 nm with a single nanometer resolution. A white reference panel (Specralon® Labsphere Inc., North Dutton, NH) reading for remote-sensing and k-records for LAI were collected at the beginning of each replication within all sites served as controls.

2.3. Data and Analysis

Plant emergence counts were standardized across seeding rates before use as a covariate. Outlier analysis was determined on yield, seed components, and agronomic traits

with studentized residuals and Cook's distance performed in JMP v. 13 (SAS Institute, Cary, NC). Broad sense heritability for yield, seed components, agronomic traits, and physiological traits were calculated as:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GSr}^2}{j} + \frac{\sigma_{GE}^2}{k} + \frac{\sigma_{GSrE}^2}{jk} + \frac{\sigma_e^2}{jkr}}$$

where σ_G^2 is the genotypic variance, σ_{GSr}^2 is the genotype x seeding variance, σ_{GE}^2 is the genotype x environment variance, σ_{GSrE}^2 is the genotype x seeding rate x environment variance, σ_e^2 is the error variance, j is the number of seeding rates, k is the number of environments, and r is the number of replications. The estimation of REML variance components was performed in JMP with all effects as random. A mixed ANOVA with a covariate term was conducted to assess the impact of the fixed effects genotype, seeding rate, and the genotype x seeding rate interaction on yield, seed components, and agronomic traits. Random effect terms included environment, genotype x environment, seeding rate x environment, genotype x seeding rate x environment, and spatial nested terms range (environment) and pass (environment).

SPAD values were filtered using R-package library (MIPHENO) (Bell et al., 2012) and averaged for single value per plot, which is a median-based normalization method for use in datasets where there are no explicit controls. LAI and MTA were estimated with FV2200 software. ASD Spectra were processed using ViewSpec pro software, and single and multiple wavebands of absolute reflectance were extracted with R software to calculate the following vegetative indices (Table 2.2). The multivariate normal imputation utility in JMP imputed missing physiological trait values. This algorithm uses least squares imputation. Pairwise correlations of physiological traits by growth stage and yield, seed components, and

agronomic traits were determined using the multivariate procedure of JMP. Yield prediction and feature extraction was performed with the regularized regression technique adaptive Elastic Net in JMP v. 13. Elastic Net alpha was set to 0.09, number of grid points = 150, and minimum penalty fraction = 0. Yield prediction models were assembled with physiological traits from growth stages R1-2, R3-4, R5-6, and combinations of R3-6 and R1-6. Models including growth stage R1-2 consisted of two environments instead of three. Five-fold cross-validation was performed to avoid inflated estimates of predictive ability, and cross-validation was repeated 10 times to assess feature selection stability.

3. Results

Mean seed yield pooled across five locations was 3051.22 kg ha⁻¹, and the averages of seed protein and oil concentration were 35.3% and 18.7%, respectively (

Table 2.3). Broad sense heritabilities calculated for yield, seed components, and agronomic traits were high with a range from 0.89-0.99 (

Table 2.3). Environments ranked by seed yield from highest to lowest were 2015 Bruner, 2014 Worle, 2015 Milo, 2015 Lippert, and 2014 Milo (Figure 2.1A). There was no noticeably better yielding year, but there was a strong individual environment effect considering all locations were in Central Iowa. Although the range of values is large for all traits, as expected from the diverse panel selected, the values follow a normal distribution, excluding lodging score, which was skewed to the left (Figure 2.1A-C). The relationship between seed yield and seed oil was positive, and both were negative with seed protein (Figure 2.6), and these relationships are generally observed (Wilcox and Shibles, 2001). Plant emergence followed a normal distribution at each seeding rate across environments and

consistently varied more at higher seeding rates than lower. Almost entirely distinct clusters reaffirmed seeding rate effects (Figure 2.5).

3.1. Genotype, Seeding rate, and Genotype x SR effects

A mixed ANOVA with a covariate of plant emergence was used to determine the fixed treatment effects of yield, seed components, and agronomic traits of the pooled five location dataset. Genotype was significant for all traits, and seeding rate was significant at various levels of alpha for all traits excluding seed weight and maturity. The genotype x seeding rate interaction was not significant for any trait but lodging (Table 2.4).

. Plant emergence appeared to influence seed composition traits protein and oil, but no others (Table 2.4).

When determining which levels of seeding rate were significantly different from each other, yield, seed protein, and protein percentage, and plant height revealed a similar pattern. We found that the low seeding rate treatment was significantly different from optimal and high treatment, but the optimal and high seeding rate treatment were no different from each other. Yield, seed oil percentage, and height were greater in optimal and high seeding rates, and lower in the low seeding rate. The inverse observation was true for seed protein percentage, i.e., lower in optimal and high seeding rates, and higher in the low seeding rate. Lodging was significantly different for each level of seeding rate and increased from low to high seeding rate.

Lodging was the only trait to reveal a significant genotype x seeding rate interaction effect. Genotypes were clustered by ancestry to show lodging was lowest in elite lines, followed by diverse, and highest for plant introductions. An ANOVA substituting genotype with ancestry revealed the ancestry x seeding rate interaction was significant at alpha =

0.05 for lodging (Table 2.5). Yield was significantly different by ancestry, by seed traits oil percentage, protein percentage, and seed weight were not (Table 2.5). Our results suggest elite genotypes to be more resilient to lodging at a higher seeding rate (Figure 2.7).

3.2. Physiological drivers of soybean yield under contrasting seeding rates

Pearson correlations between traits confirmed high correlations of related vegetative indices (Figure 2.2). Vegetative indices at R1-2 and R3-4 were more strongly correlated than R3-4 and R5-6, and R1-2 and R5-6 did not display high correlation. Of the non-remote sensing traits, iPAR and LAI were more strongly correlated with the vegetative indices across all growth stages in contrast to SPAD or MTA. Yield appeared to be equally correlated with physiological traits across growth stages (Figure 2.2).

Broad sense heritability, or repeatability, was generally found to increase over the reproductive growth stages for most physiological traits, and the highest heritabilities were found in chlorophyll-related traits, followed by foliage and water content traits (Table 2.6). Regularized regression by the adaptive elastic net method was implemented for yield prediction models for each seeding rate and repeated over reproductive growth stages and combinations thereof. Overall, R² values ranged from 0.44-0.55 for R1-2, 0.55-0.59 for R3-4, and 0.52-0.64 for R5-6, increasing with later reproductive stages for all seeding rates. R² values were greatest in the combination of all reproductive growth stages together, R1-6, ranging from 0.77-0.82, followed by the combination of the latter two stages, R3-6, ranging from 0.67-0.74. Yield prediction was similar between seeding rates at each growth stage, but low seeding rate models were observed to be more predictive in most cases, apart from R3-4 (Figure 2.3). Physiological traits selected in the adaptive elastic net models were summarized by growth stage and seeding rates and ranked by standard least squares (

Table 2.7). Traits were reported if they were selected by a simple majority (greater than five) of the ten-repeated five-fold cross-validations. Solution stability, the reoccurrence of terms either selected or discarded by the model, was relatively high, with most traits entirely excluded or retained (Figure 2.8). Models by seeding rate and growth stage were further compared and summarized for feature overlap. Traits that were specific to one seeding rate were found at each growth stage, excluding optimal R3-4 and high R5-6 (Figure 2.4A). The likelihood that any given trait would be predictive across seeding rates increased at later growth stages, as more traits were shared in predictive models (Figure 2.4A). In contrast, when combining growth stages, the likelihood was equal for unique or shared for R3-6 and greater for unique than shared for R1-6, the most predictive model. If any trait was shared in the R1-6 model, it was most likely between all seeding rates, followed by high-optimal, optimal-low, and lastly high-low (Figure 2.4B). SPAD collected during seed development was the most predictive trait across seeding rates for the R5-6 and combined growth stages (Table 2.7) SPAD was selected in R1-2 only for high seeding rate and for all seeding rates in R3-4. In the R1-6 combined model, iPAR at flowering was the second most predictive trait for low and optimal seeding rates, but less predictive for high. Of the highest ranked traits selected for the three seeding rates in the R1-6 model, low seeding rate listed several chlorophyll related remote sensing indices. The optimal seeding rate model was distinguished with consecutive iPAR traits, and high included MTA at pod development (Table 2.7). Overall, the highest ranked traits were shared among seeding rate models, and traits descending in rank diverged between models.

4. Discussion

4.1. Genotype, Seeding rate, and Genotype x SR effects

In the following, we discuss the results from our mixed ANOVA analysis for the first objective of this study. The large genotype effect and high trait heritabilities supported our observation that the population subset from the SoyNAM parent panel was effectively diverse. Efforts to include diverse genotypes outside the narrow genetic pool of elite cultivars allows researchers to expand our understanding to a broader soybean base. The second treatment of seeding rate followed genotype in magnitude of influence. It was purposefully selected to represent the extremes of low and high to overcome soybean's known variable response to seeding rate. Our overall observations of similar seeding rate response and no significant genotype x seeding rate interaction for nearly all traits in this study supports the claim that soybean, as of present, responds equally to seeding rate, regardless of ancestry. Implications of seeding rate on yield and yield related traits is discussed below.

The response variable of highest importance in this study was seed yield. A yield plateau was expected and observed at optimal seeding rate in this study because it is a general observation in elite germplasm. Our results suggest an anticipated yield plateau at optimal seeding rate can be further extended to germplasm of diverse and plant introduction ancestry. Importantly, inclusion of diverse germplasm in breeding programs may not negatively impact yield potential at higher seeding rates. A lack of genotype x seeding rate interaction in our study conflicts with significant interactions previously reported (Gan et al., 2002; Suhre et al., 2014). Ablett et al. (1991) and Beuerlein (1988) found that determinant types had a greater yield response to increased seeding rates than indeterminate or semi-determinate types, but no determinant types were included in our panel. Therefore, this may account for our lack of observed genotype x seeding rate interaction. However, Suhre et al.

(2014) found a genotype x seeding rate interaction at lower seeding rates between soybean lines of indeterminate growth and contributed this to compensatory yield on plant branches. Either branching ability and compensatory pod set were equal among our genotypes, or more likely, other mechanisms of yield potential left uncharacterized compensated for yield under low seeding rates. Moreover, a lack of significant genotype x seeding rate interaction for yield suggests when germplasm enhancement programs use diverse collections to obtain parental materials, they will witness a yield response to seeding rate no different from elite materials.

Our data revealed seeding rate does not affect seed traits equally, in view that seed weight was not influenced by seeding rate while seed oil and protein were impacted. Environment has some bearing on seed size (Borrás et al., 2004), but it can also be relatively stable, considering stable seed size QTL were identified across environments (Kato et al., 2014) and nutrient treatments (Hacisalihoglu et al., 2017). In contrast to seed weight, seed traits protein and oil concentration were significantly influenced by seeding rate and the only traits affected by plant emergence. Bellaloui et al. (2014) and Cober et al. (2005) found that protein concentration increased and oil concentration decreased with increasing seeding rate, supporting our results that seed composition traits are highly responsive to plant spacing. Because oil and protein response to seeding rate was similar for all genotypes in this study, farmers and producers may better estimate protein potential in consideration of seeding rate, since some regions have been shown to produce near insufficient protein percentage (Rotundo et al., 2016). We suggest that future cultivars may need to be evaluated at lower seeding rates to ensure the minimum protein percentage is attained.

Lodging potential determines yield potential by affecting photosynthetic ability, mechanical harvest losses, and disease pressure (Ustun et al., 2001) and positively correlates with plant height. In the current study, plant height increased with seeding rate and plateaued at optimal, but interestingly, lodging continued to increase past optimal seeding rate with no increase in height. Traits, e.g. stem thickness, not measured in this study, may have contributed to lodging, as thinner stems have been observed in high seeding rates (Lueschen and Hicks, 1977). Our observation that elite lines maintain low lodging compared to soybean lines of diverse or plant introduction ancestry supports previous observations that recent cultivars withstand lodging at higher plant densities (Rincker et al., 2014). A caveat for incorporation of diverse material into elite breeding programs includes increased lodging, most noticeable at higher seeding rates.

Noting environmental or genetic factors influencing maturity is essential, because extended maturity increases soybean yield potential. Soybean maturity is primarily determined by photoperiod, and neither shade stress (Egli, 1997) nor seeding rate in three genotypes (Gan et al., 2002) were found to have an effect on physiological maturity. These former studies suggest a maturity response to seeding rate is unlikely, and that is what we observed. However, Cober et al. (2005) found higher plant populations resulted in earlier plant maturity, implying while not detected in this study, maturity differences due to seeding rate may exist in other environments. Measuring maturity response to increased seeding rate will need to be further evaluated in more environments, as a shortened growing season will have implications for yield potential.

4.2. Physiological drivers of soybean yield under contrasting seeding rates

The second objective of this study was to identify physiological trait predictors of yield under treatments of seeding rate and determine when a given trait is most predictive of

yield. Infinite combinations of vegetative indices and their variations are used to forecast and predict crop traits (Bendig et al., 2015; Candiago et al., 2015; Kross et al., 2015), but vegetation indices in this study were selected to minimize the amount of pairwise correlation over growth stages. This minimization was important, because in the presence of strong correlation, traditional methods of feature selection, such as stepwise regression and generalized regression technique lasso, are not ideal. Feature selection is aided by independence between measured variables, because when a group of variables among which pairwise correlations are high, often one variable from the group is selected at random (Zou and Hastie, 2005). As our results indicated many traits were correlated, particularly among the remoting-sensing indices, we opted to implement adaptive elastic net, an extension of lasso, for our feature selection and yield prediction. This method encourages a grouping effect, where strongly correlated predictors tend to be in or out of the model together (Zou and Hastie, 2005). Therefore, our choice of method for feature selection facilitated identification of important physiological traits predicting yield under any given seeding rate.

Soybean yield was predicted at individual growth stages and reached significant levels, noticeably at later reproductive growth stages. However, assuming yield formation cannot be determined from any single growth stage, combinations of growth stages would provide greater yield prediction accuracy, as evidenced in our study. Wang et al. (2014) similarly found improved yield prediction in wheat using multi-temporal remote sensing data. We show yield prediction from a single growth stage is possible and suggest measuring traits at later reproductive periods for increased prediction. At single growth stage collections, traits were more likely to be equally predictive for yield across seeding rates, implying a single model would be sufficient to predict yield in varying levels of seeding rate.

Only combining growth stages and increasing the coefficient of determination made obvious traits that predicted yield under one seeding rate and not another. A comprehensive utilization of multiple growth stages is therefore emphasized for future studies.

The leading physiological traits predicting yield for the most predictive models were summarized in this study. SPAD is an indirect measurement of leaf chlorophyll and we found it to be the most predictive trait for all seeding rates at seed fill, R5-6. It is not surprising that chlorophyll related traits were ranked the leading yield predictors, for the reason that chlorophyll content has been demonstrated to linearly increase with cultivar year of release (Koester et al., 2016). Several of the remote sensing indices in this study are also used for prediction of chlorophyll content. However, these indices may still be sensitive to the combined response of several vegetation and environmental properties, such as canopy shadows and background soil reflectance (Haboudane et al., 2002). Furthermore, broad-sense heritabilities of the chlorophyll indices were lower than SPAD. Crain et al. (2017) observed an increase of trait heritability on a given day was a good indication of how well that dataset correlated to yield, suggesting increasing trait heritability would increase prediction. Progressing from SPAD to chlorophyll vegetation indices would be ideal for increased throughput of data collection. As our SPAD values were averaged over ten individual measurements and remote sensing from a single measurement, we theorize increasing our replications of spectral measurements would increase heritabilities. We hypothesize heritabilities increased at later reproductive growth stages for remote sensing traits because of a greater canopy to background soil ratio.

After chlorophyll content, our selected feature light interception (iPAR) at flowering, R1-2, agrees with a historical observation that full canopy coverage by flowering determines

yield (Board and Harville, 1993). We suggest future studies could gravitate towards remote-sensing for measurement collection, in that our leading traits can be substituted with remote sensing indices. However, we caution using remote sensing alone on the grounds that Glenn et al. (2008) argued vegetative indices should be used simply as a measurement of canopy light absorption rather than as a surrogate for detailed features of canopy architecture.

Remote-sensing may be limited in predicting in high seeding rates as mean tilt angle was the second ranked predictor, unless future improvements in estimating MTA from remote sensing are achieved. Overall, our most predictive seeding rate models shared common leading traits, indicating improvement of these physiological traits will lead to increased performance across seeding rates.

5. Conclusions

In this study, genotype and seeding rate interactions for yield, seed components, and agronomic traits were evaluated, and adaptive elastic net models identified the underlying physiological traits predicting yield response to three levels of seeding rate. A significant genotype x seeding rate interaction was only detected for lodging, and not for yield, seed weight, seed oil percentage, seed protein percentage, height, or maturity. These results suggest that current soybean germplasm and soybean of wide genetic ancestry respond similarly to seeding rate and implies introgression of diverse material may not detrimentally affect yield response to seeding rate. In addition, physiological traits predicting the yield response within and across seeding rates were summarized, with chlorophyll traits determined as the leading predictors across seeding rates in this study. Our further characterization of diverging traits between the seeding rate yield models will provide the research community targets for soybean improvement for current and future seeding rate practices. Moreover, further characterizing genotype x seeding rate across diverse

germplasm to understand the mechanisms underlying yield response to seeding rate is an important direction for future research and soybean improvement.

Acknowledgments

We sincerely appreciate the members of Singh Soybean Group at Iowa State University (ISU), particularly Kyle Parmley, for assistance with phenotyping; and Dr. Jim Rouse and staff of the Iowa Crop Improvement Association for initially assisting with project implementation with field planting. We thank Dr. Somak Dutta for his feedback on statistical analyses. We are thankful to ISU, Iowa Soybean Association, Monsanto Chair in Soybean Breeding at ISU, and the R. F. Baker Center for Plant Breeding at ISU for financial support. Race Higgins was supported by the Monsanto Chair in Soybean Breeding, Iowa Soybean association and R. F. Baker Center for Plant Breeding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Ablett, G.R., Beversdorf, W.D., Dirks, V.A., 1991. Row width and seeding rate performance of indeterminate, semideterminate, and determinate soybean. *J. Prod. Agric.* 4, 391–NP.
- Agudamu, Yoshihira, T., Shiraiwa, T., 2016. Branch development responses to planting density and yield stability in soybean cultivars. *Plant Prod. Sci.* 19, 331–339. <https://doi.org/10.1080/1343943X.2016.1157443>
- Araus, J.L., Cairns, J.E., 2014. Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. <https://doi.org/10.1016/j.tplants.2013.09.008>
- Bell, S.M., Burgoon, L.D., Last, R.L., 2012. MIPHENO: data normalization for high throughput metabolite analysis. *BMC Bioinformatics* 13, 10.
- Bellaloui, N., Mengistu, A., Walker, E.R., Young, L.D., 2014. Soybean seed composition as affected by seeding rates and row spacing. *Crop Sci.* 54, 1782–1795.
- Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., Bareth, G., 2015. Combining UAV-based plant height from crop surface models, visible, and

- near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs. Geoinformation* 39, 79–87. <https://doi.org/10.1016/j.jag.2015.02.012>
- Beuerlein, J.E., 1988. Yield of indeterminate and determinate semidwarf soybean for several planting dates, row spacings, and seeding rates. *J. Prod. Agric.* 1, 300–303.
- Blackburn, G.A., 1998. Spectral indices for estimating photosynthetic pigment concentrations: a test using senescent tree leaves. *Int. J. Remote Sens.* 19, 657–675.
- Board, J., 2000. Light interception efficiency and light quality affect yield compensation of soybean at low plant populations. *Crop Sci.* 40, 1285–1294.
- Board, J.E., Harville, B.G., 1996. Growth dynamics during the vegetative period affects yield of narrow-row, late-planted soybean. *Agron. J.* 88, 567–572.
- Board, J.E., Harville, B.G., 1993. Soybean yield component responses to a light interception gradient during the reproductive period. *Crop Sci.* 33, 772–777.
- Boegh, E., Soegaard, H., Broge, N., Hasager, C.B., Jensen, N.O., Schelde, K., Thomsen, A., 2002. Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. *Remote Sens. Environ.* 81, 179–193.
- Bolton, D.K., Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* 173, 74–84.
- Borrás, L., Slafer, G.A., Otegui, M.E., 2004. Seed dry weight response to source–sink manipulations in wheat, maize and soybean: a quantitative reappraisal. *Field Crops Res.* 86, 131–146. <https://doi.org/10.1016/j.fcr.2003.08.002>
- Candiago, S., Remondino, F., De Giglio, M., Dubbini, M., Gattelli, M., 2015. Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote Sens.* 7, 4026–4047.
- Carpenter, A.C., Board, J.E., 1997. Branch yield components controlling soybean yield stability across plant populations. *Crop Sci.* 37, 885–891.
- Carter, G.A., 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Remote Sens.* 15, 697–703.
- Chang, W.-S., Lee, H.-I., Hungria, M., 2015. Soybean Production in the Americas, in: *Principles of Plant-Microbe Interactions*. Springer, Cham, pp. 393–400. https://doi.org/10.1007/978-3-319-08575-3_41
- Chappelle, E.W., Kim, M.S., McMurtrey, J.E., 1992. Ratio analysis of reflectance spectra (RARS): an algorithm for the remote estimation of the concentrations of chlorophyll

- a, chlorophyll b, and carotenoids in soybean leaves. *Remote Sens. Environ.* 39, 239–247.
- Cober, E.R., Morrison, M.J., Ma, B., Butler, G., 2005. Genetic Improvement Rates of Short-Season Soybean Increase with Plant Population. *Crop Sci.* 45, 1029. <https://doi.org/10.2135/cropsci2004.0232>
- Crain, J., Reynolds, M., Poland, J., 2017. Utilizing High-Throughput Phenotypic Data for Improved Phenotypic Selection of Stress-Adaptive Traits in Wheat. *Crop Sci.* 57, 648–659. <https://doi.org/10.2135/cropsci2016.02.0135>
- De Bruin, J.L., Pedersen, P., 2009. New and Old Soybean Cultivar Responses to Plant Density and Intercepted Light. *Crop Sci.* 49, 2225. <https://doi.org/10.2135/cropsci2009.02.0063>
- Duvick, D.N., 1984. Genetic diversity in major farm crops on the farm and in reserve. *Econ. Bot.* 38, 161–178.
- Edwards, J.T., Purcell, L.C., Karcher, D.E., 2005. Soybean Yield and Biomass Responses to Increasing Plant Population among Diverse Maturity Groups. *Crop Sci.* 45, 1778. <https://doi.org/10.2135/cropsci2004.0570>
- Egli, D.B., 1997. Cultivar maturity and response of soybean to shade stress during seed filling. *Field Crops Res.* 52, 1–8.
- Egli, D.B., 1993. RELATIONSHIP OF UNIFORMITY OF SOYBEAN SEEDLING EMERGENCE TO YIELD. *J. Seed Technol.* 17, 22–28.
- Elmore, R.W., 1998. Soybean cultivar responses to row spacing and seeding rates in rainfed and irrigated environments. *J. Prod. Agric.* 11, 326–331.
- Fawcett, J., Sievers, J., DeJong, J., Rossiter, L., 2017. On-Farm Corn and Soybean Planter Trials. *Farm Prog. Rep.* 2016, 112.
- Fehr, W.R., Caviness, C.E., Burmood, D.T., Pennington, J.S., 1971. Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. *Crop Sci.* 11, 929–931.
- Ferreira, A.S., Junior, B., Antonio, A., Werner, F., Zucareli, C., Franchini, J.C., Debiassi, H., Ferreira, A.S., Junior, B., Antonio, A., Werner, F., Zucareli, C., Franchini, J.C., Debiassi, H., 2016. Plant density and mineral nitrogen fertilization influencing yield, yield components and concentration of oil and protein in soybean grains. *Bragantia* 75, 362–370. <https://doi.org/10.1590/1678-4499.479>
- Frederick, J.R., Camp, C.R., Bauer, P.J., 2001. Drought-stress effects on branch and mainstem seed yield and yield components of determinate soybean. *Crop Sci.* 41, 759–763.

- Gan, Y., Stulen, I., van Keulen, H., Kuiper, P.J.C., 2002. Physiological response of soybean genotypes to plant density. *Field Crops Res.* 74, 231–241.
[https://doi.org/10.1016/S0378-4290\(01\)00212-X](https://doi.org/10.1016/S0378-4290(01)00212-X)
- Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58, 289–298.
- Glenn, E.P., Huete, A.R., Nagler, P.L., Nelson, S.G., 2008. Relationship Between Remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape. *Sensors* 8, 2136–2160.
- Grassini, P., Torrión, J.A., Yang, H.S., Rees, J., Andersen, D., Cassman, K.G., Specht, J.E., 2015. Soybean yield gaps and water productivity in the western US Corn Belt. *Field Crops Res.* 179, 150–163.
- Green-Tracewicz, E., Page, E.R., Swanton, C.J., 2011. Shade Avoidance in Soybean Reduces Branching and Increases Plant-to-Plant Variability in Biomass and Yield Per Plant. *Weed Sci.* 59, 43–49. <https://doi.org/10.1614/WS-D-10-00081.1>
- Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81, 416–426.
[https://doi.org/10.1016/S0034-4257\(02\)00018-4](https://doi.org/10.1016/S0034-4257(02)00018-4)
- Hacisalihoglu, G., Burton, A.L., Gustin, J.L., Eker, S., Asikli, S., Heybet, E.H., Ozturk, L., Cakmak, I., Yazici, A., Burkey, K.O., 2017. Quantitative trait loci associated with soybean seed weight and composition under different phosphorus levels. *J. Integr. Plant Biol.*
- Hardisky, M.A., Klemas, V., Smart, M., 1983. The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of. *Spartina Alterniflora* 77–83.
- Huang, Y., Yuan, L., Reddy, K.N., Zhang, J., 2016. In-situ plant hyperspectral sensing for early detection of soybean injury from dicamba. *Biosyst. Eng.* 149, 51–59.
<https://doi.org/10.1016/j.biosystemseng.2016.06.013>
- Johnson, D.M., 2014. An assessment of pre-and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sens. Environ.* 141, 116–128.
- Kato, S., Sayama, T., Fujii, K., Yumoto, S., Kono, Y., Hwang, T.-Y., Kikuchi, A., Takada, Y., Tanaka, Y., Shiraiwa, T., 2014. A major and stable QTL associated with seed weight in soybean across multiple environments and genetic backgrounds. *Theor. Appl. Genet.* 127, 1365–1374.

- Koester, R.P., Nohl, B.M., Diers, B.W., Ainsworth, E.A., 2016. Has photosynthetic capacity increased with 80 years of soybean breeding? An examination of historical soybean cultivars. *Plant Cell Environ.* n/a-n/a. <https://doi.org/10.1111/pce.12675>
- Kross, A., McNairn, H., Lapen, D., Sunohara, M., Champagne, C., 2015. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *Int. J. Appl. Earth Obs. Geoinformation* 34, 235–248. <https://doi.org/10.1016/j.jag.2014.08.002>
- Lee, C.D., Egli, D.B., TeKrony, D.M., 2008. Soybean Response to Plant Population at Early and Late Planting Dates in the Mid-South. *Agron. J.* 100, 971. <https://doi.org/10.2134/agronj2007.0210>
- Lueschen, W.E., Hicks, D.R., 1977. Influence of plant population on field performance of three soybean cultivars. *Agron. J.* 69, 390–393.
- Ma, B.L., Dwyer, L.M., Costa, C., Cober, E.R., Morrison, M.J., 2001. Early prediction of soybean yield from canopy reflectance measurements. *Agron. J.* 93, 1227–1234.
- Ma, B.L., Morrison, M.J., Voldeng, H.D., 1995. Leaf greenness and photosynthetic rates in soybean. *Crop Sci.* 35, 1411–1414.
- Merzlyak, M.N., Gitelson, A.A., Chivkunova, O.B., Rakitin, V.Y., 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.* 106, 135–141.
- Nutter Jr, F.W., Tylka, G.L., Guan, J., Moreira, A.J.D., Marett, C.C., Rosburg, T.R., Basart, J.P., Chong, C.S., 2002. Use of remote sensing to detect soybean cyst nematode-induced plant stress. *J. Nematol.* 34, 222.
- Peñuelas, J., Filella, I., Biel, C., Serrano, L., Save, R., 1993. The reflectance at the 950–970 nm region as an indicator of plant water status. *Int. J. Remote Sens.* 14, 1887–1905.
- Peñuelas, J., Filella, I., Gamon, J.A., 1995. Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytol.* 131, 291–296. <https://doi.org/10.1111/j.1469-8137.1995.tb03064.x>
- Rincker, K., Nelson, R., Specht, J., Sleper, D., Cary, T., Cianzio, S.R., Casteel, S., Conley, S., Chen, P., Davis, V., Fox, C., Graef, G., Godsey, C., Holshouser, D., Jiang, G.-L., Kantartzi, S.K., Kenworthy, W., Lee, C., Mian, R., McHale, L., Naeve, S., Orf, J., Poysa, V., Schapaugh, W., Shannon, G., Uniatowski, R., Wang, D., Diers, B., 2014. Genetic Improvement of U.S. Soybean in Maturity Groups II, III, and IV. *Crop Sci.* 0. <https://doi.org/10.2135/cropsci2013.10.0665>

- Romero, A., Aguado, I., Yebra, M., 2012. Estimation of dry matter content in leaves using normalized indexes and PROSPECT model inversion. *Int. J. Remote Sens.* 33, 396–414.
- Rotundo, J.L., Miller-Garvin, J.E., Naeve, S.L., 2016. Regional and Temporal Variation in Soybean Seed Protein and Oil across the United States. *Crop Sci.* 56, 797. <https://doi.org/10.2135/cropsci2015.06.0394>
- Rouse Jr, J., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS.
- Stupar, R.M., Specht, J.E., 2013. Insights from the soybean (*Glycine max* and *Glycine soja*) genome: past, present, and future. *Adv Agron* 118, 177–204.
- Suhre, J.J., Weidenbenner, N.H., Rowntree, S.C., Wilson, E.W., Naeve, S.L., Conley, S.P., Casteel, S.N., Diers, B.W., Esker, P.D., Specht, J.E., Davis, V.M., 2014. Soybean Yield Partitioning Changes Revealed by Genetic Gain and Seeding Rate Interactions. *Agron. J.* 106, 1631. <https://doi.org/10.2134/agronj14.0003>
- USDA-NASS, 2016. Crop production 2016 Summary. USDA-NASS, Washington, DC.
- Ustun, A., Allen, F.L., English, B.C., 2001. Genetic progress in soybean of the US Midsouth. *Crop Sci.* 41, 993–998.
- Vega, C., 2000. Reproductive Allometry in Soybean, Maize and Sunflower. *Ann. Bot.* 85, 461–468. <https://doi.org/10.1006/anbo.1999.1084>
- Vogelmann, J.E., Rock, B.N., Moss, D.M., 1993. Red edge spectral measurements from sugar maple leaves. *Int J Remote Sens* 14, 1563–1575.
- Wang, L., Qu, J.J., 2007. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophys. Res. Lett.* 34.
- Wang, L., Tian, Y., Yao, X., Zhu, Y., Cao, W., 2014. Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. *Field Crops Res.* 164, 178–188. <https://doi.org/10.1016/j.fcr.2014.05.001>
- Weiner, J., Thomas, S.C., 1986. Size Variability and Competition in Plant Monocultures. *Oikos* 47, 211. <https://doi.org/10.2307/3566048>
- Wilcox, J.R., Shibles, R.M., 2001. Interrelationships among seed quality attributes in soybean. *Crop Sci.* 41, 11–14.
- Wilson, E.W., Rowntree, S.C., Suhre, J.J., Weidenbenner, N.H., Conley, S.P., Davis, V.M., Diers, B.W., Esker, P.D., Naeve, S.L., Specht, J.E., 2014. Genetic Gain × Management Interactions in Soybean: II. Nitrogen Utilization. *Crop Sci.* 54, 340–348.

- Xavier, A., Beavis, W.D., Specht, J.E., Diers, B., Muir, W.M., Rainey, K.M., 2015. SoyNAM: Soybean Nested Association Mapping Dataset. R Package Version 1.
- Zanon, A.J., Streck, N.A., Grassini, P., 2016. Climate and management factors influence soybean yield potential in a subtropical environment. *Agron. J.* 108, 1447–1454.
- Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67, 301–320.

Tables and figures

Table 2.1 Origin, population number, and maturity grouping of the soybean NAM parental genotype subset assessed in this study.

NAM parent	Origin	NAM population	Ancestry	Growth habit	Maturity (DAP)
4J105-3-4	Purdue Univ.	NAM 03	Elite	Indeterminate	137
5M20-2-5-2	Purdue Univ.	NAM 04	Elite	Indeterminate	142
CLOJ095-4-6	Purdue Univ.	NAM 05	Elite	Indeterminate	133
CLOJ173-6-8	Purdue Univ.	NAM 06	Elite	Indeterminate	136
HS6-3976	Ohio State	NAM 08	Elite	Indeterminate	137
IA3023	Iowa State Univ.	Universal Parent	Elite	Indeterminate	139
LD01-5907	Univ. of Illinois	NAM 11	Elite	Indeterminate	141
LD02-4485	Univ. of Illinois	NAM 12	Elite	Indeterminate	130
LG00-3372	USDA-ARS	NAM 38	Diverse	Indeterminate	135
LG03-2979	USDA-ARS	NAM 24	Diverse	Indeterminate	135
LG04-4717	USDA-ARS	NAM 26	Diverse	Indeterminate	136
LG05-4464	USDA-ARS	NAM 29	Diverse	Indeterminate	141
LG05-4832	USDA-ARS	NAM 30	Diverse	Indeterminate	138
LG90-2550	USDA-ARS	NAM 31	Diverse	Semi-determinate	133
LG92-1255	USDA-ARS	NAM 32	Diverse	Indeterminate	127
LG94-1128	USDA-ARS	NAM 33	Diverse	Indeterminate	128
LG94-1906	USDA-ARS	NAM 34	Diverse	Indeterminate	132
LG97-7012	USDA-ARS	NAM 36	Diverse	Indeterminate	134
LG98-1605	USDA-ARS	NAM 37	Diverse	Indeterminate	123
Maverick	Univ. of Missouri	NAM 15	Elite	Indeterminate	138
NE3001	Univ. of Nebraska	NAM 18	Elite	Semi-determinate	133
PI 398.881	South Korea	NAM 40	Plant introduction	Indeterminate	131
PI 404.188A	China	NAM 54	Plant introduction	Indeterminate	131
PI 427.136	South Korea	NAM 41	Plant introduction	Indeterminate	134
PI 437.169B	Russia	NAM 42	Plant introduction	Indeterminate	129
PI 507.681B	-	NAM 46	Plant introduction	Indeterminate	131
PI 518.751	Serbia	NAM 48	Plant introduction	Indeterminate	131
PI 561.370	China	NAM 50	Plant introduction	Indeterminate	136
PI 574.486	China	NAM 64	Plant introduction	Indeterminate	141
Prohio	Ohio State Univ.	NAM 09	Elite	Indeterminate	141
Skylla	Mich. State Univ.	NAM 22	Elite	Indeterminate	124
U03-100612	Univ. of Nebraska	NAM 23	Elite	Indeterminate	120

Table 2.2 Summary of vegetation indices with abbreviation, general class type, and source.

Vegetation index description	Abbreviation	Type	Source
Photochemical reflectance index	PRI	Carotenoids	Peñuelas et al., 1995
Plant senescence reflectance index	PSRI	Carotenoids	Merzlyak et al., 1999
Pigment specific simple ratio chlorophyll A	PSSRa	Chlorophyll	Blackburn, 1998
Ratio analysis of reflectance spectra chlorophyll A	RARSa	Chlorophyll	Chappelle et al., 1992
Ratio analysis of reflectance spectra chlorophyll B	RARSb	Chlorophyll	Chappelle et al., 1992
Vogelmann red edge index 2	VREI2	Chlorophyll	Vogelmann et al., 1993
Dry matter content index	DMCI	Dry matter	Romero et al., 2012
Green average (505 to 595)	Green	Vegetation	Gitelson et al., 1996
Leaf area index (vegetation index)	LAI (VI)	Vegetation	Boegh et al., 2002
Normalized difference vegetation index	NDVI	Vegetation	Rouse Jr et al., 1974
Normalized difference moisture index	NDMI	Water content	Hardisky et al., 1983
Normalized multi-band drought index	NMDI	Water content	Wang and Qu, 2007
Normalized water index 1	NWIA	Water content	Peñuelas et al., 1993

Table 2.3 Yield, seed components, and agronomic trait summaries with count, mean, standard deviation, range, and Broad-sense heritabilities.

Trait	n	Mean	Std Dev	Range	$H^{2†}$
Yield (kg/ha)	1393	3051.2	992.8	458-5461	0.90
Seed protein %	1363	35.3	1.6	30-40	0.92
Seed oil %	1363	18.7	0.9	16-22	0.89
100 seed weight (g) [‡]	1092	14.7	2.2	9-23	0.95
Maturity (DAP) [§]	844	133.7	5.6	118-144	0.99
Lodging (score 1-5)	1401	2.4	1.3	1-5	0.91
Height (cm)	1400	86.7	16.8	34-146	0.98

[†] H^2 , broad-sense heritability, calculated on an entry mean basis

[‡]4 environments

[§]3 environments

Table 2.4 ANOVA of yield, seed components, and agronomic traits of five environments and three replications by genotype.

Source of variation	Effect	df	F value and significance level of fixed effects						
			Height	Lodging	Maturity	Seed weight	Seed oil %	Seed protein %	Yield
Genotype	FE	31	45.23**	15.25**	1036.29**	23.11**	9.87**	12.48**	8.66**
Seeding rate	FE	2	44.06**	45.53**	<1	1.05	14.53**	8.55*	31.71**
G x SR	FE	62	1.03	1.81**	<1	1.26	1.19	1.2	1.09
Emergence	FE	1	2.12	<1	<1	<1	11.76**	5.41*	1.06

* Significant at the 0.05 level

** Significant at the 0.01 level

Table 2.5 ANOVA of yield, seed components, and agronomic traits of five environments and three replications by ancestry.

Source of variation	Effect	df	F value and significance level of fixed effects						
			Height	Lodging	Maturity	Seed weight	Seed oil %	Seed protein %	Yield
Ancestry	FE	2	<1	12.88***	<1	1.01	2.03	2.04	17.5***
Seeding rate	FE	2	41.28***	46.25***	<1	<1	14.49***	8**	38.44***
Anc x SR	FE	4	<1	2.51*	1.41	2.13	<1	1.3	1.02
Emergence	FE	1	<1	<1	<1	<1	16.55***	8.25***	<1

* Significant at the 0.05 level

** Significant at the 0.01 level

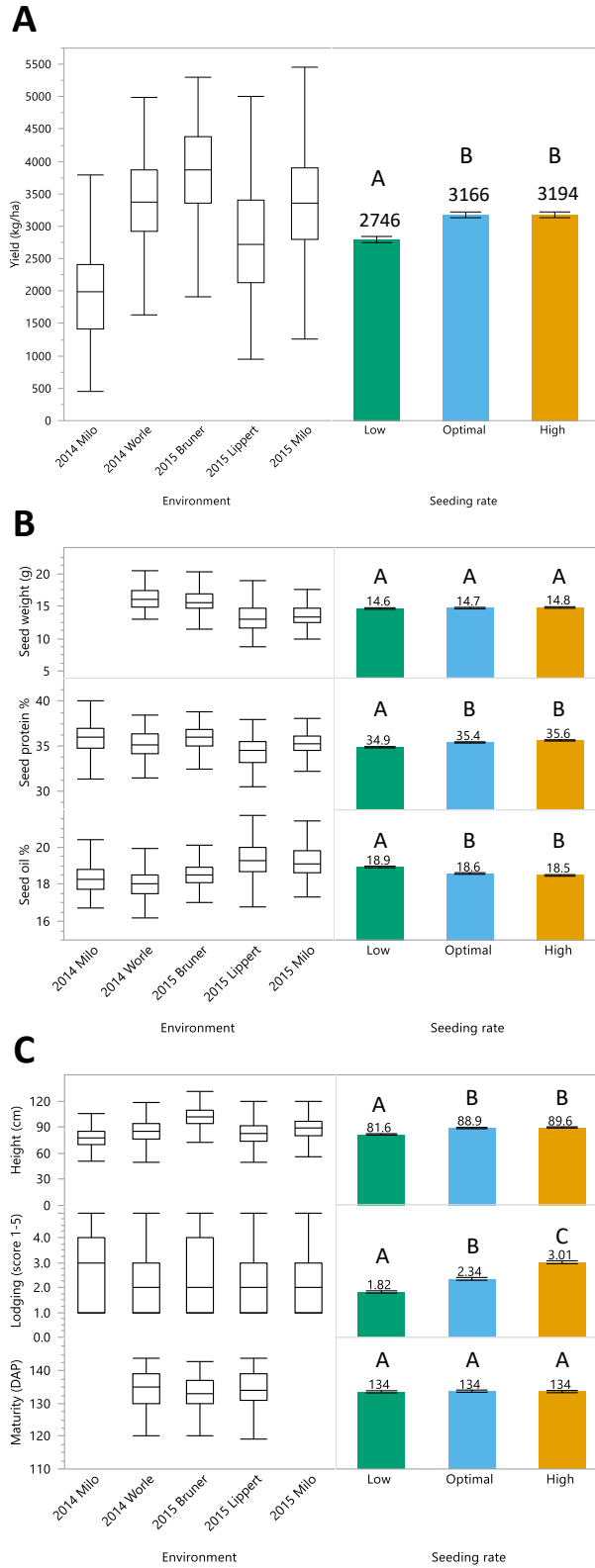


Figure 2.1 Boxplot distributions by environment and bar chart by seeding rate with standard error, mean, and letter grouping for A) yield B) seed components and C) agronomic

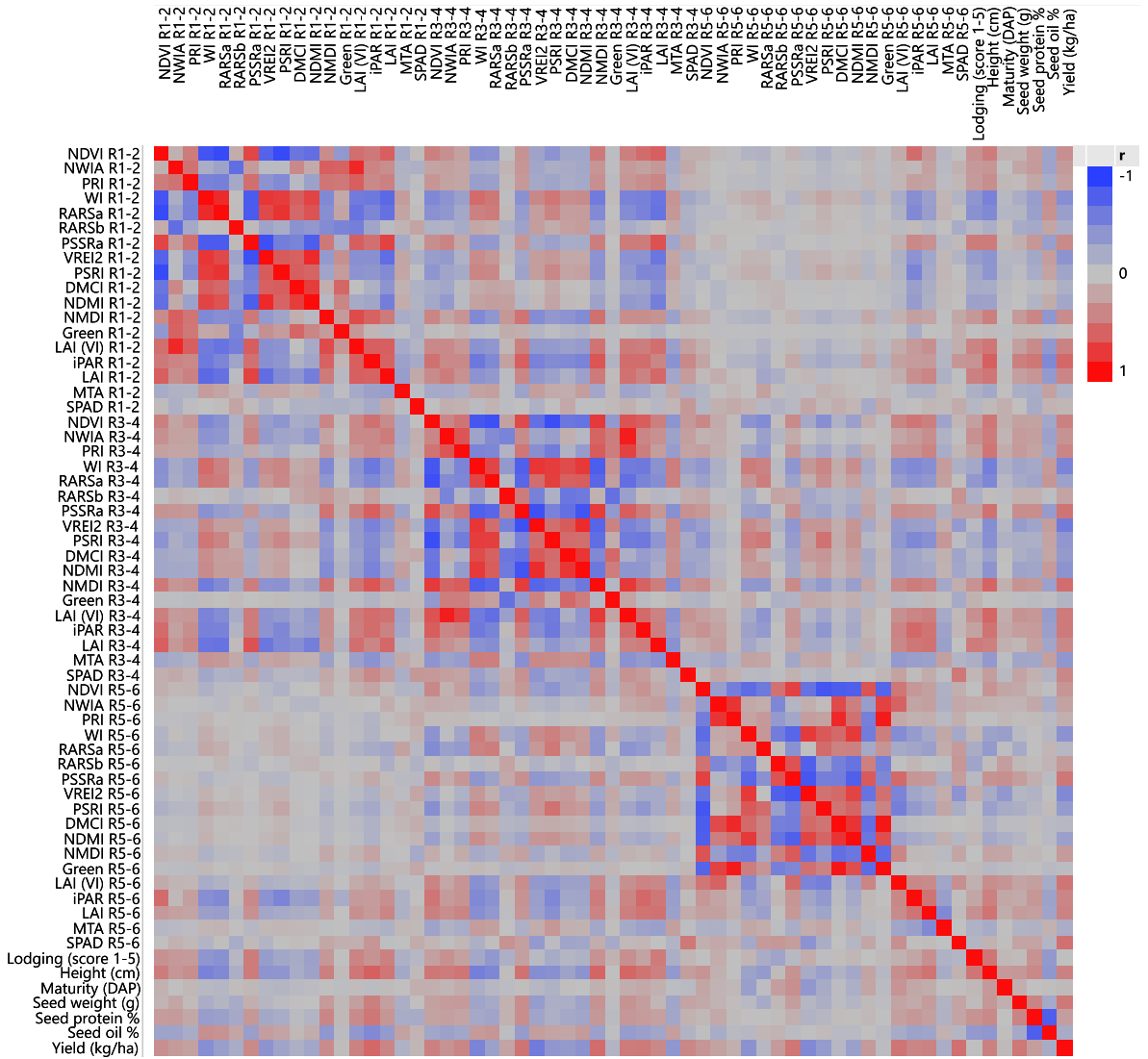


Figure 2.2 Heatmap of Pearson correlations at three reproductive growth stages between physiological traits, agronomic traits, seed components, and final yield.

Table 2.6 Summary data of physiological traits by growth stage with number of observations, percent of data imputed, mean, standard deviation, range, and heritability.

Physiological trait	R1-2 Flowering						R3-4 Pod development						R5-6 Seed development					
	n [‡]	Percent imputed	Mean	Std Dev	Range	H^2 [†]	n	Percent imputed	Mean	Std Dev	Range	H^2 [†]	n	Percent imputed	Mean	Std Dev	Range	H^2 [†]
iPAR	570	0	0.66	0.21	0.95	0.18	841	1.3	0.9	0.09	0.67	0.56	841	0	0.93	0.09	0.61	0.50
LAI	570	0	2.31	0.9	5.94	0.35	841	1.3	5.02	1.14	6.89	0.18	841	35.2	5.94	0.91	7.27	0.36
MTA	570	3.2	45.33	9.37	88.00	0.02	841	1.3	44.99	4.93	35.00	0.12	841	35.2	43.47	3.69	39.00	0.38
SPAD	570	0.2	35.49	4.08	27.40	0.80	841	14.3	42.36	3.14	24.20	0.76	841	15.3	43.81	3.52	21.40	0.93
DMCI	570	0	-1.12	0.21	1.66	0.00	841	0	-1.07	0.15	1.07	0.06	841	0	-1.06	0.34	2.81	0.16
Green	570	0	0.02	0.02	0.38	0.13	841	0	0.04	0.02	0.35	0.57	841	0	0.05	0.07	0.63	0.26
LAI (VI)	570	0	0.43	0.28	1.33	0.43	841	0	0.9	0.27	1.44	0.52	841	0	0.89	0.2	1.47	0.57
NDMI	570	0	-3.18	1.49	8.05	0.03	841	0	-3.55	1.28	9.15	0.38	841	0	-3.56	1.22	8.19	0.65
NDVI	570	0	0.83	0.15	0.89	0.10	841	0	0.9	0.09	0.71	0.20	841	0	0.89	0.09	0.90	0.44
NMDI	570	0	-0.33	0.21	1.70	0.00	841	0	-0.15	0.14	1.00	0.43	841	0	-0.18	0.16	1.29	0.67
NWIA	570	0	-0.65	0.21	1.25	0.42	841	0	-0.23	0.26	1.49	0.45	841	0	-0.19	0.28	2.09	0.26
PRI	570	0	-0.93	0.06	0.92	0.48	841	0	-0.89	0.05	0.78	0.63	841	0	-0.87	0.14	1.40	0.30
PSRI	570	0	0.02	0.05	0.31	0.02	841	0	0.01	0.02	0.35	0.33	841	0	0.01	0.02	0.32	0.20
PSSRa	570	0	17.45	10.24	67.48	0.08	841	0	24.79	10.95	74.89	0.27	841	0	22.56	8.48	51.84	0.80
RARSa	570	0	0.49	0.15	0.71	0.13	841	0	0.42	0.11	0.69	0.30	841	0	0.37	0.08	0.53	0.43
RARSb	570	0	46.46	32.4	197.45	0.33	841	0	22.67	9.21	97.22	0.76	841	0	17.81	7.24	53.84	0.87
VREI2	570	0	-0.2	0.08	0.43	0.53	841	0	-0.26	0.09	0.53	0.86	841	0	-0.21	0.09	0.50	0.90
WI	570	0	0.95	0.06	0.33	0.16	841	0	0.93	0.03	0.33	0.61	841	0	0.93	0.03	0.19	0.81

[†] H^2 , broad-sense heritability, calculated on an entry mean basis

[‡]2 environments

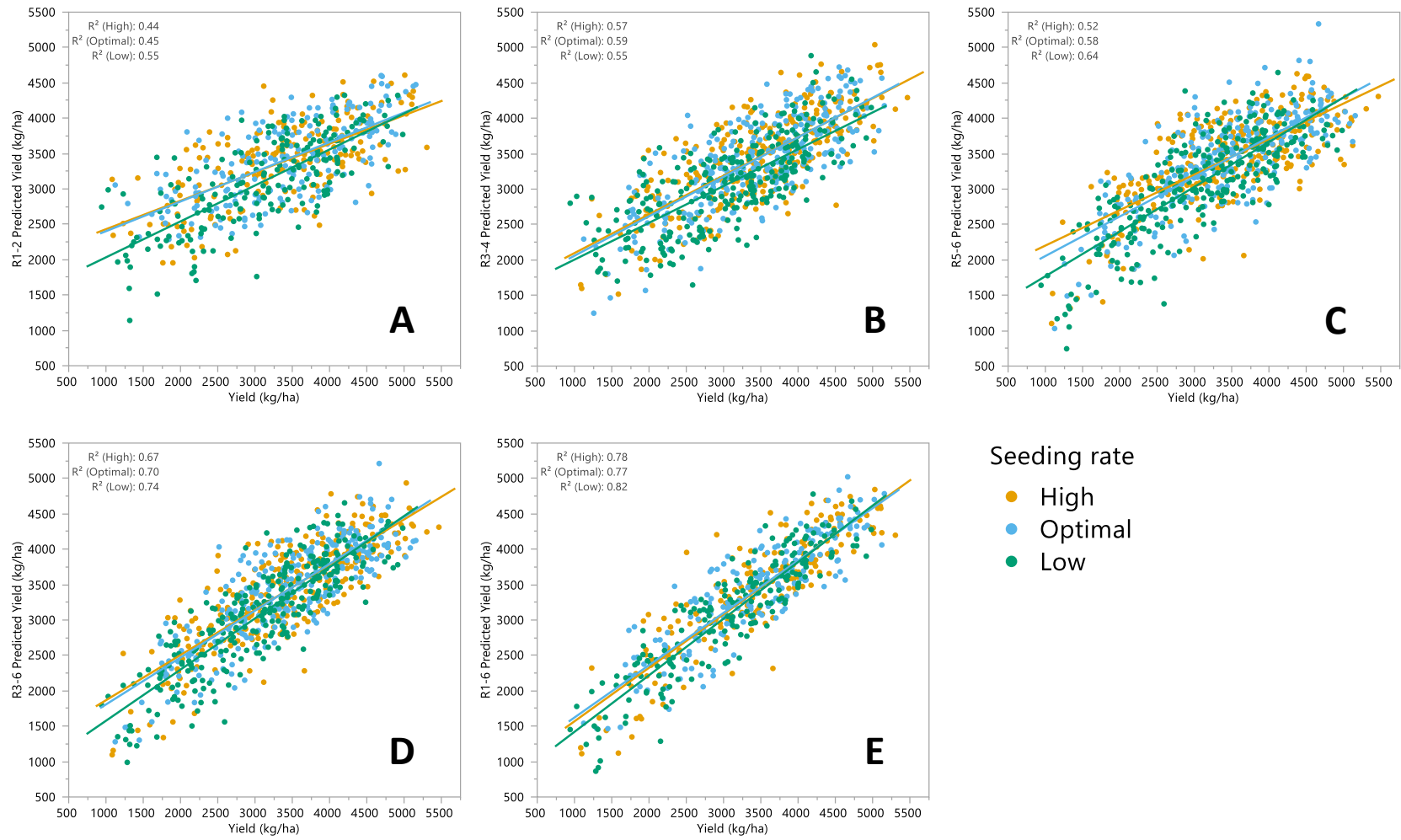


Figure 2.3 Adaptive elastic net yield prediction for three seeding rates of low, optimal, and high across reproductive growth stages A) R1-2, B) R3-4, and C) R5-6, and for two growth stage combinations of D) R3-6, and E) R1-6

Table 2.7 List of physiological traits selected by adaptive elastic net regularized regression for three soybean reproductive growth stages and two combinations of growth stages. Traits are ranked by standard least squares.

Rank	Single growth stages									Combination growth stages					
	R1-2 Flowering [†]			R3-4 Pod development			R5-6 Seed development			R3-6 Pod-Seed [†]			R1-6 Flowering-Seed		
	Low	Optimal	High	Low	Optimal	High	Low	Optimal	High	Low	Optimal	High	Low	Optimal	High
1	iPAR	iPAR	VREI2	NDMI	Green	LAI	SPAD	SPAD	SPAD	SPAD R5-6	SPAD R5-6	SPAD R5-6	SPAD R5-6	SPAD R5-6	SPAD R5-6
2	RARSb	WI	DMCI	DMCI	MTA	DMCI	iPAR	LAI (VI)	LAI	VREI2 R5-6	iPAR R3-4	LAI R3-4	iPAR R1-2	iPAR R1-2	MTA R3-4
3	NDVI	NWIA	iPAR	LAI	iPAR	Green	VREI2	NWIA	NWIA	PSSRa R3-4	MTA R3-4	NWIA R5-6	PSSRa R3-4	iPAR R3-4	LAI R1-2
4	VREI2	VREI2	PSSRa	VREI2	LAI	NDMI	RARSa	LAI	LAI (VI)	LAI R5-6	Green R3-4	iPAR R3-4	VREI2 R5-6	DMCI R3-4	Green R1-2
5	RARSa	PSRI	PSRI	SPAD	SPAD	VREI2	LAI	NMDI	Green	iPAR R5-6	NWIA R5-6	VREI2 R3-4	RARSa R3-4	WI R1-2	NDMI R1-2
6	LAI	RARSb	RARSb	RARSb	DMCI	PRI	NDVI	RARSb	iPAR	PRI R5-6	LAI (VI) R5-6	DMCI R3-4	PSRI R5-6	iPAR R5-6	PSSRa R1-2
7	DMCI	NMDI	NDMI	PRI	NWIA	MTA	Green	DMCI	NMDI	WI R3-4	RARSa R5-6	LAI (VI) R5-6	SPAD R1-2	PRI R5-6	LAI R3-4
8	Green	PSSRa	SPAD	NDVI	NDMI	PSSRa	NWIA	RARSa	RARSa	DMCI R3-4	RARSa R3-4	Green R5-6	NDVI R3-4	NDMI R5-6	PRI R1-2
9	LAI (VI)	MTA	LAI	WI	VREI2	iPAR	-	iPAR	VREI2	NMDI R5-6	DMCI R5-6	MTA R3-4	PRI R5-6	SPAD R1-2	iPAR R3-4
10	-	-	PRI	RARSa	PRI	SPAD	-	Green	-	SPAD R3-4	RARSb R5-6	LAI R5-6	NMDI R1-2	NMDI R1-2	NWIA R5-6
11	-	-	MTA	PSSRa	RARSa	RARSa	-	-	-	RARSa R3-4	iPAR R5-6	PSSRa R3-4	PRI R3-4	NWIA R1-2	SPAD R1-2
12	-	-	LAI (VI)	NWIA	PSRI	PSRI	-	-	-	NWIA R5-6	PRI R5-6	PSSRa R5-6	Green R5-6	NDMI R1-2	NMDI R1-2
13	-	-	-	-	WI	LAI (VI)	-	-	-	PSRI R5-6	LAI R3-4	NMDI R5-6	NDVI R5-6	Green R5-6	LAI R5-6
14	-	-	-	-	-	-	-	-	-	PRI R3-4	NWIA R3-4	RARSa R3-4	NDVI R1-2	Green R3-4	NWIA R3-4
15	-	-	-	-	-	-	-	-	-	NWIA R3-4	LAI (VI) R3-4	NDVI R3-4	Green R1-2	PSRI R1-2	DMCI R3-4
16	-	-	-	-	-	-	-	-	-	-	-	NWIA R3-4	NMDI R5-6	LAI (VI) R3-4	Green R5-6
17	-	-	-	-	-	-	-	-	-	-	-	-	NWIA R3-4	LAI (VI) R5-6	LAI (VI) R5-6
18	-	-	-	-	-	-	-	-	-	-	-	-	-	NDVI R3-4	Green R3-4
19	-	-	-	-	-	-	-	-	-	-	-	-	-	VREI2 R5-6	LAI (VI) R1-2
20	-	-	-	-	-	-	-	-	-	-	-	-	-	RARSa R3-4	RARSa R3-4
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	RARSa R1-2
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	VREI2 R5-6
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	PSSRa R5-6

[†]2 environments

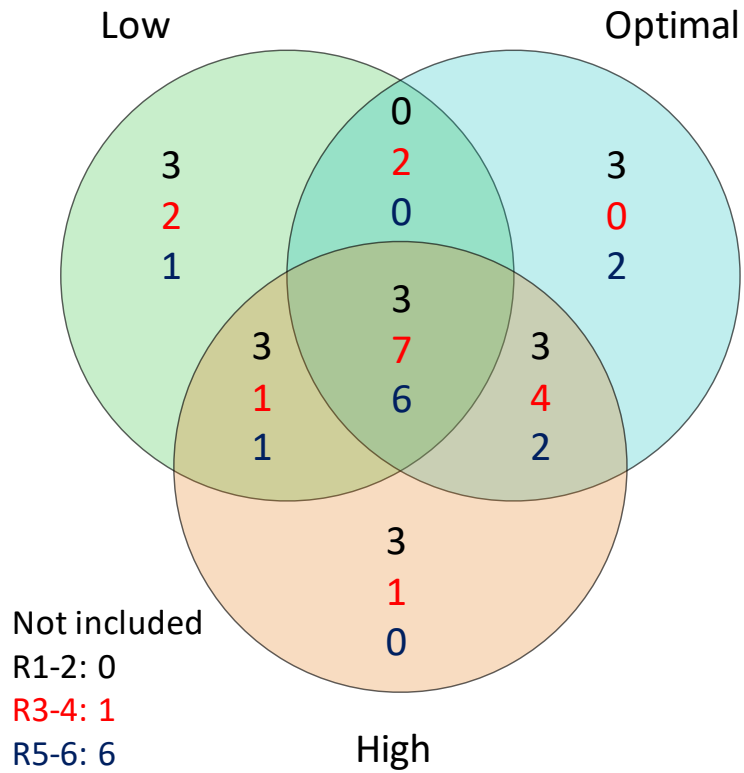
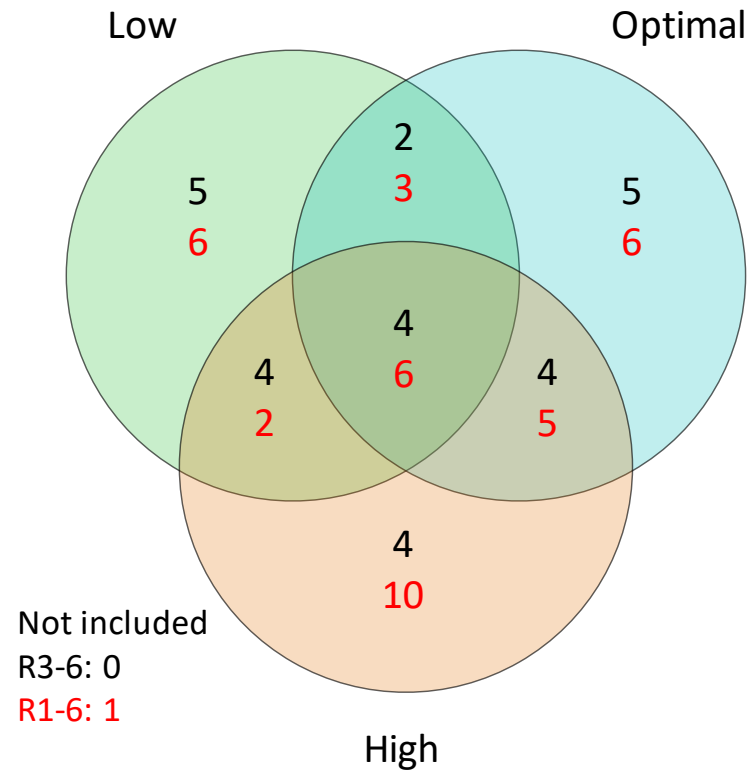
A**B**

Figure 2.4 Venn diagram of shared and unique physiological traits selected by adaptive elastic net models at three seeding rates and A) three reproductive growth stages R1-2, R3-4, and R5-6 and B) two combinations of growth stages, R3-6 and R1-6.

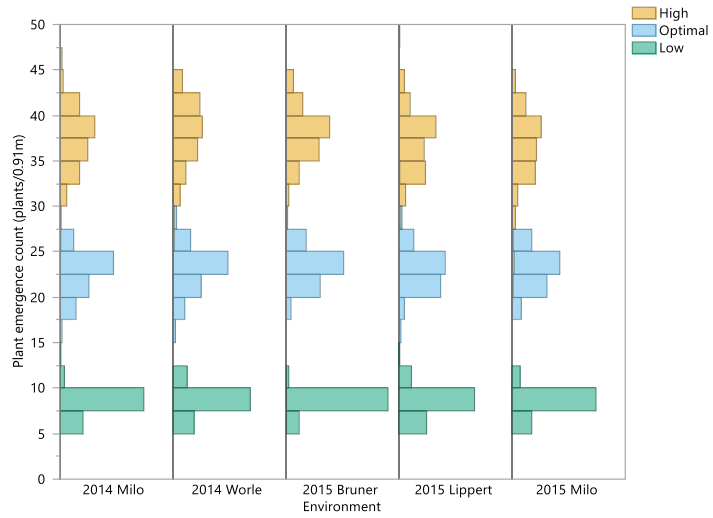


Figure 2.5 Frequency histogram of plant emergence for three target seeding rates of low, optimal, and high in five replicated environments

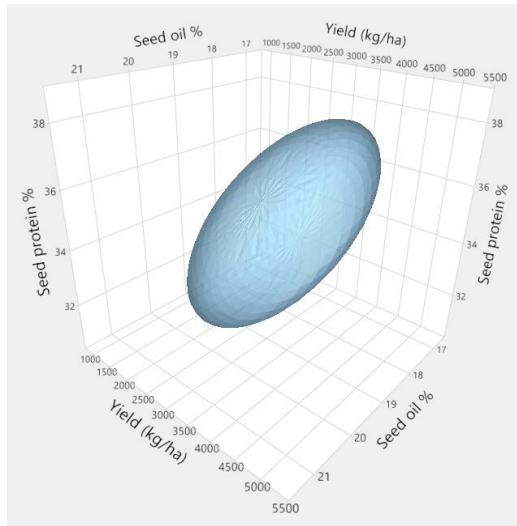


Figure 2.6 Relationship between seed yield, seed oil percentage, and seed protein percentage for five locations. A positive relationship was observed between seed yield and seed oil percentage, but a negative relationship to seed protein percentage.

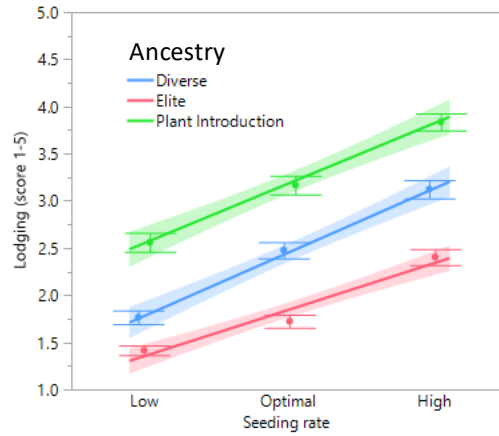


Figure 2.7 Mean lodging scores for three seeding rates of low, optimal, and high and three ancestry categories of elite, diverse, and plant introduction. Means are marked with standard error bars.

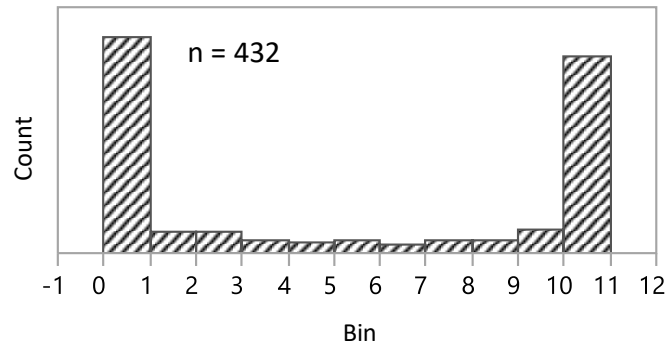


Figure 2.8 Feature selection stability for 10 repeated 5 k-fold cross-validations of adaptive elastic net models. The bin number represents the frequency of trait predictor inclusion in the models, and the high frequency of zero and ten bin groupings indicates the majority of traits were either completely retained or discarded in every model cross-validation instance, demonstrating solution stability of the adaptive elastic net.

CHAPTER 3. UNCOVERING BIOMASS PARTITIONING AND RESIDUE QUALITY TRAITS FOR SOYBEAN IMPROVEMENT

A paper in preparation for submission to *Field Crops Research*

R.H. Higgins, A.K. Singh

Highlights

- Soybean genotypes of diverse ancestry partition biomass similarly at flowering (R1) and diverge at pod development (R4).
- A higher percentage of petioles and lower percentage of stem at full pod (R4) correlated positively seed weight
- The reported range of carbon:nitrogen (C:N) ratios in soybean were extended, with lower C:N genotypes identified for residue quality improvement.

Keywords

harvest index, soybean biomass, soybean partitioning, carbon:nitrogen (C:N) ratio

Abstract

Improving economic return is one of the most important objectives in a soybean breeding program. Traits such as higher harvest index improved return by greater output of grain yield, while soybean residue quality as carbon:nitrogen content can unintentionally decreased inputs of nitrogen fertilizer for succeeding crops in rotation. Harvest index is a simple ratio of grain weight to total biomass, but the vegetative proportions and residue quality of the total biomass and ultimate relationship to grain yield is unknown. The objectives of this study were to 1) quantify the difference in biomass partitioning strategies using diverse soybean genotypes of the SoyNAM parent panel, and further understand the temporal physiological basis of biomass partitioning through sampling at three reproductive

growth stages: R1, R4, and R8, and 2) characterize the genetic variation in residue quality of final biomass by measuring carbon:nitrogen (C:N) ratio. Results showed significant differences in biomass partitioning and plant component percentages of stem, pod, seed petiole, and leaves between soybean genotypes at three reproductive stages. Significant genetic variation in C:N residue quality was found for each residue component, with no negative relationship to final grain yield. A genotype “Prohio” was identified as having exceptional residue quality. These findings indicate optimal biomass partitioning strategies for yield and improved residue C:N ratios for whole-system nitrogen sustainability can be targeted for yield improvement.

1. Introduction

Soybeans [*Glycine max* (L.) Merr.] are the second most planted crop in the United States (USDA-NASS, 2016) and fourth globally after wheat (*Triticum aestivum* L.), rice (*Oryza sativa* L.), and maize (*Zea mays* L.). As the world's largest source of animal protein feed and the second largest source of vegetable oil, various endeavors have been sought to enhance soybean production around the globe.

Crop profitability is influenced by crop inputs along with realized economic part yield (i.e., seed yield in soybean). Farming decisions are therefore ideally made to maximize yield with minimal input costs and optimally using all economic parts to drive up the profitability. Harvest index is one important factor as it is described as the ratio of seed yield to total biomass and impacts crop output. It has been suggested that harvest index of 0.6 is the theoretical maximum, and maximum harvest index in soybean has been achieved (Zhu et al., 2010). In a research study, researchers found the harvest index for soybean ranged from 56.2 – 58.0% for elite soybean cultivars in the US Midwest, which is near theoretical maximum (Pedersen and Lauer, 2004). Increased harvest index in soybean has resulted from increased

seed yield with little increase in total aboveground biomass (Jin et al., 2010; Morrison et al., 2000), but the relationship between the composition of the total aboveground biomass to grain yield is unknown. (Srinivasan et al., 2016) showed evidence that modern crop genotypes produce more leaf than is optimal, and removing leaves resulted in an 8% increase in yield. Identifying genotypes with optimal biomass partitioning strategies could further advance soybean line development where harvest index has been maximized and finding the reproductive growth stage when differences in biomass partitioning first become evident can provide insight to when genotypes begin to physiologically diverge.

Complementing increasing soybean output for greater season profitability and decreasing input costs through reduced fertilizer requirements of the succeeding crop can be achieved by improving soybean residue quality. Nitrogen (N), is essential for plant growth and seed production (Lawlor, 2002) and is the main component of fertilizer. Ubiquitous maize-soy cropping system in the US Midwest may particularly benefit, as improving the amount and quality of the soybean residue can contribute to the yield increase of maize in the succeeding season (Green and Blackmer, 1995) through increasing plant available nitrogen. Gentry et al. (2013) found net soil nitrogen mineralization was the strongest predictor of yield difference in continuous corn systems, where net mineralization of soil nitrogen is influenced by both quality (C:N ratio) and quantity of residue from the previous crop (Gentry et al., 2001). Cotrufo et al. (2013) developed a framework (MEMS) on the hypothesis that labile plant constituents, dependent on residue quality, are the dominant source of microbial products because they are utilized more efficiently by microbes, influencing soil mineralization. These beg the question if soybean value may be further enhanced in the maize-soy rotation by improving the carbon-nitrogen (C:N) ratio in soybean residue, without

penalizing grain yield. Genotypic differences in C:N were found within multiple crop species between wild and domesticated crops (García-Palacios et al., 2013), and significant genotypic variations in stem nitrogen traits at maturity were found in modern soybean cultivars (Fritschi et al., 2013). (Dhanapal et al., 2015) further supported finding C:N ratio variation in a collection of 373 soybean genotypes at flowering (R2). A caveat to lower C:N ratios is the concern that higher amounts of nitrogen in the vegetative plant organs results in less N remobilization to the seed during grain fill, suggesting a yield penalty, but multiple studies have shown that direct nitrogen uptake and accumulation during seed fill could be a more important factor for high seed yield instead of N remobilization (Kumudini et al., 2001; Zhao et al., 2014). Unknown are genotype-specific C:N ratios of the whole composite residue and its relationship with soybean yield, along with two important seed quality factors, seed protein and seed oil content.

A preliminary study suggested an inverse relationship in C:N ratio of whole composite residue samples for two elite soybean cultivars with different seed protein content. To further elucidate this relationship and capture the genetic diversity of biomass partitioning and residue quality through carbon/nitrogen content in soybeans, a 32 parent panel of genotypes in this study was selected from the soyNAM population, which represents high-yielding lines, lines with diverse ancestry, and plant introductions (Song et al., 2017) and was expected to cover the diversity of soybean. The parent panel was subsetted based on maturity adapted to Central Iowa.

Through these analyses, we aimed to identify the biomass partitioning strategies of diverse genotypes over multiple reproductive growth stages for potential future application to increase yield. To expand upon knowledge on estimated C:N ratio ranges and variation

within soybean, we further investigate the quality of the residual biomass through estimating C:N of each biomass component: stem, fallen residue (petioles and leaves), pod, and seed. Future studies may demonstrate that favorable C:N ratios may impact the succeeding crop, especially nitrogen responsive crops such as maize in the soy-maize rotation, ultimately reducing nitrogen inputs amended to the soil in the Midwest.

2. Materials and methods

2.1 Plant Materials

A 32-genotype subset of the soyNAM panel was selected based on maturity adapted to Central Iowa (maturity groups II-III) (Table 3.1). This diverse sample varied in growth habit and genetic ancestry. Specifically, growth habits were either indeterminate or semideterminate; and genetic ancestry were elite lines, high-yielding lines of diverse ancestry, or direct plant introductions. Experiments were grown in one central Iowa location in 2014 (Worle -environment 1) and two central Iowa locations in 2015 (Agronomy and Burkey, environments 2 and 3). The soil type found at environments 2 and 3 are Nicollet loam series of Aquic Hapludoll whereas environment 1 is Clarion loam series of Typic Hapludoll (Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture, 2016).

2.2 Experimental Design

Experiments at three locations were designed as a randomized complete block with three replications where each of the 32 genotypes was fixed effect treatment. Fields were planted in the month of May: 13th, 21st, and 29th for environments 2, 3, and 1 respectively. Planting density was targeted at 31 plants m⁻² in 0.76-m rows and plots were designed as four rows, 6 m long each. Destructive sampling measurements were taken within the two central

rows. Procedures were the same at each farm, except for environment 1, where samples were harvested from the border rows. Weeds were chemically controlled at planting and hand removed during crop growth. Pests were controlled as needed with standard agronomic practices. Disease prevalence was noted and recorded per plot basis.

2.3 Plant Measurements

Above ground biomass samples were collected in 0.91 m sections per row and randomized within six designated sections that were 0.3 m apart and at least 0.61 m from the beginning and end of the plot. Plants were harvested at three reproductive (R) soybean growth stages on a 1-8 scale: beginning bloom (R1), full pod (R4), and physiological (R8) per (Fehr et al., 1971). Because genotypes varied in maturity, samples were collected at multiple time-points during the season until the targeted growth stage was achieved. R1 and R4 plant samples were partitioned into stem, petiole, and leaves, and R4 with an additional section of pods. R8 samples were dissected into a modified category from the previous growth stages of stem, pods, seed, and fallen residue (leaves and petioles). Fallen residue was too dry and brittle to accurately partition into leaves and petioles. Samples were dried at 60 degrees C until completely dry, and dry weight measurements were collected immediately after removal from the dryer ovens. R8 plant samples were finely ground using a 2-mm screen in a Wiley mill (Thomas Scientific) (Figure 3.1). Carbon and nitrogen content were then determined by submission of 0.01 g subsamples to the Iowa State University Plant and Soils Analysis lab for combustion analysis (TruSpec CN, LECO Corp., St. Joseph, MO). Approximately 300 g of whole soybean seed was used to quantify seed protein and oil contents using near-infrared reflectance (NIR) spectroscopy (Infratec™ 1241 Grain Analyzer, FOSS).

2.4 Data Analysis

Outliers were detected and eliminated using interquartile range for each plant component, location, and growth stage independently. Biomass percentage for each vegetative plant component was calculated as the ratio of the individual component over total biomass for each sample. The residue biomass percentage was calculated similarly, but without the inclusion of the final seed weight in the final biomass. C:N ratios were calculated as the percentage carbon content divided by the percentage nitrogen content.

$$C:N \text{ ratio} = \frac{\text{carbon \%}}{\text{nitrogen \%}}$$

A mixed ANOVA was used to determine the fixed treatment effect of biomass weight, biomass percentage, and C:N ratio. Separate ANOVA were performed for each growth stage and vegetative component. The sole fixed effect was genotype and environment and the environment x genotype interaction were tested as random effects with blocks nested within environment. Environment was defined as the location and year combined. All other effects were tested against the general error term. Response variables included each plant component percentage by growth stage, and C:N ratio. Data were analyzed with JMP v. 12 (SAS Institute, Cary, NC). Harvest index (HI) was calculated using by the weight of the total seed mass divided by the total mass of each plot at growth stage R8.

$$HI = \frac{\text{Seed (g)}}{\text{Total biomass (g)}}$$

Broad-sense heritability (H^2) for each plant vegetative component was estimated on an entry mean basis following (Nyquist and Baker, 1991):

$$H^2 = (\sigma^2_G) / [\sigma^2_G + \left(\frac{\sigma^2_{GE}}{e}\right) + \left(\frac{\sigma^2_e}{re}\right)]$$

in which σ^2_G is genetic variance, σ^2_{GE} is genotype x environment (location-year) variance, σ^2_e is error variance, r is number of replications, and e is the number of environments.

Nitrogen harvest index (NHI) has been summarized algebraically in formula proposed by (Sinclair, 1998):

$$NHI = \frac{[\%N_{seed} \times HI]}{[HI \times (\%N_{seed} - \%N_{residue}) + \%N_{residue}]}$$

where seed N concentration is represented as $\%N_{seed}$ and $\%N$ present in the residue fraction as $\%N_{residue}$. Correlations with plant maturity and additional agronomic traits were determined using the multivariate procedure of JMP.

3. Results

3.1 Biomass partitioning

Biomass accumulation was not equal between growing season years 2014 and 2015 (Figure 3.2A) but was similar between the two locations in 2015. Biomass accumulated over the growing period as expected over the three selected reproductive growth stages: R1, R4, and R8 (Figure 3.3). Total biomass dry weight (g) among genotypes per plot section and per plant component was significantly different between the diverse genotypes at each reproductive growth stages (Table 3.2).

Dry weights of each plant section were divided by the total biomass weight to determine whether plant section proportions were significantly different between genotypes at each growth stage between, mirroring the results for biomass dry weight. (Hanway and Weber, 1971) reported approximately 55% leaves, 31% stems, and 14% petioles at R2 and 29% seed, 17% stems, 11% pods, and 43% leaves and petioles at R8. Mean percentages of 52.6% leaves, 30% stems, and 17.3% petioles at R1 found within this study (Table 3.2) were equivalent to the former cited percentages but not at final maturity (R8) with 37.9% seed,

24.4% stems, 15.2% pods, and 37.9% leaves and petioles. Analogous to significant differences found in all biomass weights, significant proportion differences were detected for each plant section at growth stages R4 and R8 (Table 3.2), except at R1. At beginning bloom (R1), plant proportions were not different between genotypes.

The relationship between biomass percentages and weights were determined using Pearson's correlation (Table 3.3). The relationship between seed weight and harvest index was positive, as expected ($r=0.56$). Maturity had a significant negative relationship with harvest index ($r = -0.38$) (Table 3.4). Significant correlations of biomass proportions were found at R4 with seed weight (Table 3.4), specifically petioles ($r = 0.18$) and stem ($r = -0.19$).

3.2 Residue quality

C:N ratios differed between years and location, with a lower C:N ratio in 2014 and higher C:N ratios in 2015 (Figure 3.2B). C:N ratios were significantly different between and among all partitioned components: stem, pod, fallen residue, and seed (Table 3.5). The highest mean C:N ratios were found in the stem (62.3), followed by leaves/petioles (24.5), pods (32.9), and seed (7.8) respectively (Table 3.5). This trend was consistently repeated across genotypes. High heritabilities for C:N were observed, and ranged from 0.78-0.96 and were ranked from highest to lowest: seed, leaf/petiole, pod, and stem.

Total C:N ratios of the final residue biomass for each genotype were calculated with a weighted mean using the genotype biomass proportions discovered in the first portion of this experiment (Figure 3.5). The range of residue C:N ratios was 27-53 between genotypes. Two genotypes, Prohio and U03-100612, were noted as elite cultivars that contrasted in total C:N ratios for residue biomass (Figure 3.5). Genotypes Prohio and U03-100612 had significantly different biomass amounts (Table 3.2), mostly attributed to the later maturity of Prohio and

longer growing season to accumulate biomass in comparison to U03-100612. The relationship with harvest index is dependent on the crop and genotypes. A curvilinear response was observed between NHI and HI, with NHI decreasing as higher HI was attained. (Figure 3.6)

4. Discussion

4.1 Biomass partitioning

Biomass accumulation differed between growing season years 2014 and 2015 and are attributed to differences in planting date and similarities for environments within years. In 2014, smaller total weights were observed for R1 and R4 because of a later planting date, but biomass size eventually exceeded 2015 weights by R8 because of a longer growing season that year. Total biomass was expected to be dissimilar especially evident between genotypes because of the design of the genetic panel. The lines of diverse ancestry and plant introductions were not expected to be vigorous, or as well-adapted, as the public elite accessions.

No significant difference was detected in biomass proportions at beginning bloom (R1) and implies that proportionally, soybeans are very similar at R1, even if biomass size is significantly different. Removing reproductive organs pod and seed at R4 and R8 as a percentage revealed stem mass and leaf/petiole mass resulted in significantly different proportions (Table 3.2). This indicates biomass proportions can be detected at R4 with initial divergence of plant sections beginning at R1. Average harvest index measured as seed percentage at 0.38 (Table 3.2) was not expected to reach the optimal harvest index of 0.56-0.58 (De Bruin and Pedersen, 2009) because the population panel was comprised of low-yielding genotypes, but the upper range 0.54 (Table 3.2) was expected to overlap with optimal harvest index. Our biomass collection method at R8 could capture fallen residue at

the end of the season, which would account for lower harvest index values. However, this infers soybeans, particularly elite lines, may not have attained optimal harvest index if all above-ground biomass is accounted for. Continued improvement in soybean harvest index is to be expected, where Morrison et al. (1999) found after seven decades of breeding and selection (1934–1992) that seed yield and harvest index were increased 0.5% per year, and as our evidence shows, still ongoing to the present.

Maturity was negatively correlated with harvest index, suggesting later-maturing genotypes are not maximizing theoretical biomass partitioning into seed, even though later-maturing genotypes had a positive relationship with total seed weight. The correlations of petiole and stem proportions at R4 in combination with high heritabilities (Table 3.2) may implicate using biomass proportions for soybean genotype improvement.

The relationship between seed weight and harvest index was positive, but not as strongly correlated as predicted, implying harvest index may not be good indicator of grain yield in a diverse soybean genotype panel. The association between harvest index and yield has been contradictory in soybean, where (Schapaugh and Wilcox, 1980) found no correlation between harvest index and yield, but (Frederick et al., 1991; Pedersen and Lauer, 2004) found a relationship between increased harvest index and improved yield potential

4.2 Residue quality

The recorded C:N ratios of soybean stem encompassed previous reported values (Prior et al., 2006) and expanded the range, possibly due the inclusion of a diverse panel of soybean lines compared to previous literature's focus on elite cultivars. Limited literature reported expected C:N ratios of leaves/petioles or pod due to difficulty in collecting fallen residue, in conjunction with stem. The range of overall residue C:N ratios, 27-53, extended

the previous reported mean C:N ratios found within eight soybean varieties, 31.85-51.27 (Prior et al., 2006).

High heritabilities of C:N per component confirms C:N ratios are genetically controlled and wide phenotypic variation (Table 3.5) indicates C:N ratios can be targets for soybean cultivar quality improvement in combination with biomass partitioning percentages. A concern for selecting favorable C:N ratios is whether it negatively correlates with grain yield, as after the commencement of seed fill, Gaspar et al. (2017) demonstrated high-yielding soybean uses both greater vegetative nitrogen remobilization and nitrogen uptake after R5. Yet, no relationship between grain yield and residue C:N in this study was observed, appearing to be independent (Figure 3.5B).

The curvilinear relationship observed for the NHI also indicates that more translocation of dry mass to the seeds is not proportional to nitrogen translocation to the seeds, and replicates a previous finding (Tamagno et al., 2017). Hence, it seems that soybeans are limited on nitrogen partitioning from residue to seeds at high HI. Selecting for high yield and high harvest index may indirectly increase nitrogen content of the residue, increasing C:N content and residue quality. Future studies with larger experimental plot sizes will be necessary in order to determine how soybean biomass quantity and quality affects the soil nitrogen mineralization, and importantly, if there is a measurable positive gain on the following season's crop yield.

5. Conclusion

This study characterized the genetic variation in biomass partitioning strategies in a 32-parent subset of the SoyNAM soybean panel. Three reproductive stages (R1, R4, R8) were destructively harvested and partitioned into stems, petioles, leaves, pod, seed, and/or fallen residue (petioles and leaves) component dry weights were calculated as proportions of

final biomass. Dry biomass weights were significantly different for genotypes at each growth stage, but significant differences in biomass proportions were only detected in the later reproductive growth stages, R4 and R8. Understandably, higher biomass percentages in reproductive organs at R4 and R8 correlated positively with grain yield; however, a higher percentage of petioles at R4 correlated positively seed weight at R8. The strongest relationship with grain yield was final biomass weight, but harvest index was negatively correlated with final biomass weight, indicating larger canopies may not be achieving their theoretical maximum yield. Our study additionally demonstrated genetic variation in carbon:nitrogen (C:N) residue quality for each of the partitioned residue components at physiological maturity (R8). The lack of a negative relationship between yield and C:N ratio and high heritability suggests this trait can be selected in breeding programs to improve soybean residue quality. Lower C:N ratios in soybean residue are hypothesized to increase soil mineralization, and, therefore, increase nitrogen availability for the succeeding year's crop. Future studies are needed to determine the genetic control of biomass partitioning strategies and C:N ratio in soybean in order to effectively utilize in breeding programs, likely through genome-wide association mapping or linkage mapping in new populations.

Limitations in this study included subsamples instead of whole plots and limited environments restricted to Central Iowa. Future studies may be limited by labor-intensive nature of biomass partitioning and residue collection. Looking forward, economic return in soybeans can be improved by both increasing yield through targeting biomass partitioning strategies and decreasing nitrogen inputs in the following crop season through lower soybean residue C:N ratios.

Acknowledgments

We sincerely appreciate the members of Singh Soybean Group at ISU, particularly Brian Scott and Jae Brungardt, for assistance with phenotyping; the ISU Biomass and Forage Quality lab and the ISU Soil and Plant Analysis lab for use of their facilities; Mike Fiscus and staff of the Iowa State Agronomy Research Farm for help throughout the growing season. We thank Dr. Mike Castellano for research discussions. We are thankful to Iowa State University (ISU), the Iowa Soybean Association, Monsanto Chair in Soybean Breeding at ISU, and the R. F. Baker Center for Plant Breeding at ISU for financial support. Race Higgins was supported by the Monsanto Chair in Soybean Breeding, Iowa Soybean association and R. F. Baker Center for Plant Breeding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Cotrufo, M.F., Wallenstein, M.D., Boot, C.M., Deneff, K., Paul, E., 2013. The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Glob. Change Biol.* 19, 988–995.
<https://doi.org/10.1111/gcb.12113>
- De Bruin, J.L., Pedersen, P., 2009. Growth, Yield, and Yield Component Changes among Old and New Soybean Cultivars. *Agron. J.* 101, 124.
<https://doi.org/10.2134/agronj2008.0187>
- Dhanapal, A.P., Ray, J.D., Singh, S.K., Hoyos-Villegas, V., Smith, J.R., Purcell, L.C., King, C.A., Fritschi, F.B., 2015. Genome-Wide Association Analysis of Diverse Soybean Genotypes Reveals Novel Markers for Nitrogen Traits. *Plant Genome* 8, 0.
<https://doi.org/10.3835/plantgenome2014.11.0086>
- Fehr, W.R., Caviness, C.E., Burmood, D.T., Pennington, J.S., 1971. Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. *Crop Sci.* 11, 929–931.
- Frederick, J.R., Woolley, J.T., Hesketh, J.D., Peters, D.B., 1991. Seed yield and agronomic traits of old and modern soybean cultivars under irrigation and soil water-deficit. *Field Crops Res.* 27, 71–82.

- Fritschi, F.B., Ray, J.D., Purcell, L.C., King, C.A., Smith, J.R., Charlson, D.V., 2013. Diversity and Implications of Soybean Stem Nitrogen Concentration. *J. Plant Nutr.* 36, 2111–2131. <https://doi.org/10.1080/01904167.2012.748800>
- García-Palacios, P., Milla, R., Delgado-Baquerizo, M., Martín-Robles, N., Álvaro-Sánchez, M., Wall, D.H., 2013. Side-effects of plant domestication: ecosystem impacts of changes in litter quality. *New Phytol.* 198, 504–513. <https://doi.org/10.1111/nph.12127>
- Gaspar, A.P., Laboski, C.A.M., Naeve, S.L., Conley, S.P., 2017. Dry Matter and Nitrogen Uptake, Partitioning, and Removal across a Wide Range of Soybean Seed Yield Levels. *Crop Sci.* 57, 2170. <https://doi.org/10.2135/cropsci2016.05.0322>
- Gentry, L.E., Below, F.E., David, M.B., Bergerou, J.A., 2001. Source of the soybean N credit in maize production. *Plant Soil* 236, 175–184.
- Gentry, L.F., Ruffo, M.L., Below, F.E., 2013. Identifying Factors Controlling the Continuous Corn Yield Penalty. *Agron. J.* 105, 295. <https://doi.org/10.2134/agronj2012.0246>
- Green, C.J., Blackmer, A.M., 1995. Residue Decomposition Effects on Nitrogen Availability to Corn following Corn or Soybean. *Soil Sci. Soc. Am. J.* 59, 1065–1070. <https://doi.org/10.2136/sssaj1995.03615995005900040016x>
- Hanway, J.J., Weber, C.R., 1971. Dry matter accumulation in eight soybean (*Glycine max* (L.) Merrill) varieties. *Agron. J.* 63, 227–230.
- Jin, J., Liu, X., Wang, G., Mi, L., Shen, Z., Chen, X., Herbert, S.J., 2010. Agronomic and physiological contributions to the yield improvement of soybean cultivars released from 1950 to 2006 in Northeast China. *Field Crops Res.* 115, 116–123. <https://doi.org/10.1016/j.fcr.2009.10.016>
- Kumudini, S., Hume, D.J., Chu, G., 2001. Genetic Improvement in Short Season Soybeans. *Crop Sci.* 41, 391. <https://doi.org/10.2135/cropsci2001.412391x>
- Lawlor, D.W., 2002. Carbon and nitrogen assimilation in relation to yield: mechanisms are the key to understanding production systems. *J. Exp. Bot.* 53, 773–787. <https://doi.org/10.1093/jexbot/53.370.773>
- Morrison, M.J., Voldeng, H.D., Cober, E.R., 2000. Agronomic changes from 58 years of genetic improvement of short-season soybean cultivars in Canada. *Agron. J.* 92, 780–784.
- Morrison, M.J., Voldeng, H.D., Cober, E.R., 1999. Physiological Changes from 58 Years of Genetic Improvement of Short-Season Soybean Cultivars in Canada. *Agron. J.* 91, 685. <https://doi.org/10.2134/agronj1999.914685x>

- Nyquist, W.E., Baker, R.J., 1991. Estimation of heritability and prediction of selection response in plant populations. *Crit. Rev. Plant Sci.* 10, 235–322. <https://doi.org/10.1080/07352689109382313>
- Pedersen, P., Lauer, J.G., 2004. Response of soybean yield components to management system and planting date. *Agron. J.* 96, 1372–1381.
- Prior, S.A., Torbert, H.A., Runion, G.B., Rogers, H.H., Ort, D.R., Nelson, R.L., 2006. Free-Air Carbon Dioxide Enrichment of Soybean. *J. Environ. Qual.* 35, 1470. <https://doi.org/10.2134/jeq2005.0163>
- Schapaugh, W.T., Wilcox, J.R., 1980. Relationship Between Harvest Indices and other Plant Characteristics in Soybean. *Crop Sci.* 20, 529–533. <https://doi.org/10.2135/cropsci1980.0011183X002000040028x>
- Sinclair, T.R., 1998. Historical changes in harvest index and crop nitrogen accumulation. *Crop Sci.* 38, 638–643.
- Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture, 2016. Web Soil Survey [WWW Document]. URL <http://websoilsurvey.nrcs.usda.gov/app/HomePage.htm> (accessed 12.19.16).
- Song, Q., Yan, L., Quigley, C., Jordan, B.D., Fickus, E., Schroeder, S., Song, B.-H., Charles An, Y.-Q., Hyten, D., Nelson, R., Rainey, K., Beavis, W.D., Specht, J., Diers, B., Cregan, P., 2017. Genetic Characterization of the Soybean Nested Association Mapping Population. *Plant Genome* 0, 0. <https://doi.org/10.3835/plantgenome2016.10.0109>
- Srinivasan, V., Kumar, P., Long, S.P., 2016. Decreasing, not increasing, leaf area will raise crop yields under global atmospheric change. *Glob. Change Biol.* n/a-n/a. <https://doi.org/10.1111/gcb.13526>
- Tamagno, S., Balboa, G.R., Assefa, Y., Kovács, P., Casteel, S.N., Salvagiotti, F., García, F.O., Stewart, W.M., Ciampitti, I.A., 2017. Nutrient partitioning and stoichiometry in soybean: A synthesis-analysis. *Field Crops Res.* 200, 18–27. <https://doi.org/10.1016/j.fcr.2016.09.019>
- USDA-NASS, 2016. Crop production 2016 Summary. USDA-NASS, Washington, DC.
- Zhao, X., Zheng, S.-H., Fatichin, Suzuki, A., Arima, S., 2014. Varietal Difference in Nitrogen Redistribution from Leaves and Its Contribution to Seed Yield in Soybean. *Plant Prod. Sci.* 17, 103–108. <https://doi.org/10.1626/pps.17.103>
- Zhu, X.-G., Long, S.P., Ort, D.R., 2010. Improving Photosynthetic Efficiency for Greater Yield. *Annu. Rev. Plant Biol.* 61, 235–261. <https://doi.org/10.1146/annurev-arplant-042809-112206>

Tables and figures

Table 3.1 Origin, population number, and maturity grouping of the soybean NAM parental genotype subset assessed in this study.

NAM parent	Origin	NAM population	Ancestry	Growth habit	Maturity (DAP)
4J105-3-4	Purdue Univ.	NAM 03	Elite	Indeterminate	137
5M20-2-5-2	Purdue Univ.	NAM 04	Elite	Indeterminate	142
CL0J095-4-6	Purdue Univ.	NAM 05	Elite	Indeterminate	133
CL0J173-6-8	Purdue Univ.	NAM 06	Elite	Indeterminate	136
HS6-3976	Ohio State	NAM 08	Elite	Indeterminate	137
IA3023	Iowa State Univ.	Universal Parent	Elite	Indeterminate	139
LD01-5907	Univ. of Illinois	NAM 11	Elite	Indeterminate	141
LD02-4485	Univ. of Illinois	NAM 12	Elite	Indeterminate	130
LG00-3372	USDA-ARS	NAM 38	Diverse	Indeterminate	135
LG03-2979	USDA-ARS	NAM 24	Diverse	Indeterminate	135
LG04-4717	USDA-ARS	NAM 26	Diverse	Indeterminate	136
LG05-4464	USDA-ARS	NAM 29	Diverse	Indeterminate	141
LG05-4832	USDA-ARS	NAM 30	Diverse	Indeterminate	138
LG90-2550	USDA-ARS	NAM 31	Diverse	Semi-determinate	133
LG92-1255	USDA-ARS	NAM 32	Diverse	Indeterminate	127
LG94-1128	USDA-ARS	NAM 33	Diverse	Indeterminate	128
LG94-1906	USDA-ARS	NAM 34	Diverse	Indeterminate	132
LG97-7012	USDA-ARS	NAM 36	Diverse	Indeterminate	134
LG98-1605	USDA-ARS	NAM 37	Diverse	Indeterminate	123
Maverick	Univ. of Missouri	NAM 15	Elite	Indeterminate	138
NE3001	Univ. of Nebraska	NAM 18	Elite	Semi-determinate	133
PI 398.881	South Korea	NAM 40	Plant introduction	Indeterminate	131
PI 404.188A	China	NAM 54	Plant introduction	Indeterminate	131
PI 427.136	South Korea	NAM 41	Plant introduction	Indeterminate	134
PI 437.169B	Russia	NAM 42	Plant introduction	Indeterminate	129
PI 507.681B	-	NAM 46	Plant introduction	Indeterminate	131
PI 518.751	Serbia	NAM 48	Plant introduction	Indeterminate	131
PI 561.370	China	NAM 50	Plant introduction	Indeterminate	136
PI 574.486	China	NAM 64	Plant introduction	Indeterminate	141
Prohio	Ohio State Univ.	NAM 09	Elite	Indeterminate	141
Skylla	Mich. State Univ.	NAM 22	Elite	Indeterminate	124
U03-100612	Univ. of Nebraska	NAM 23	Elite	Indeterminate	120

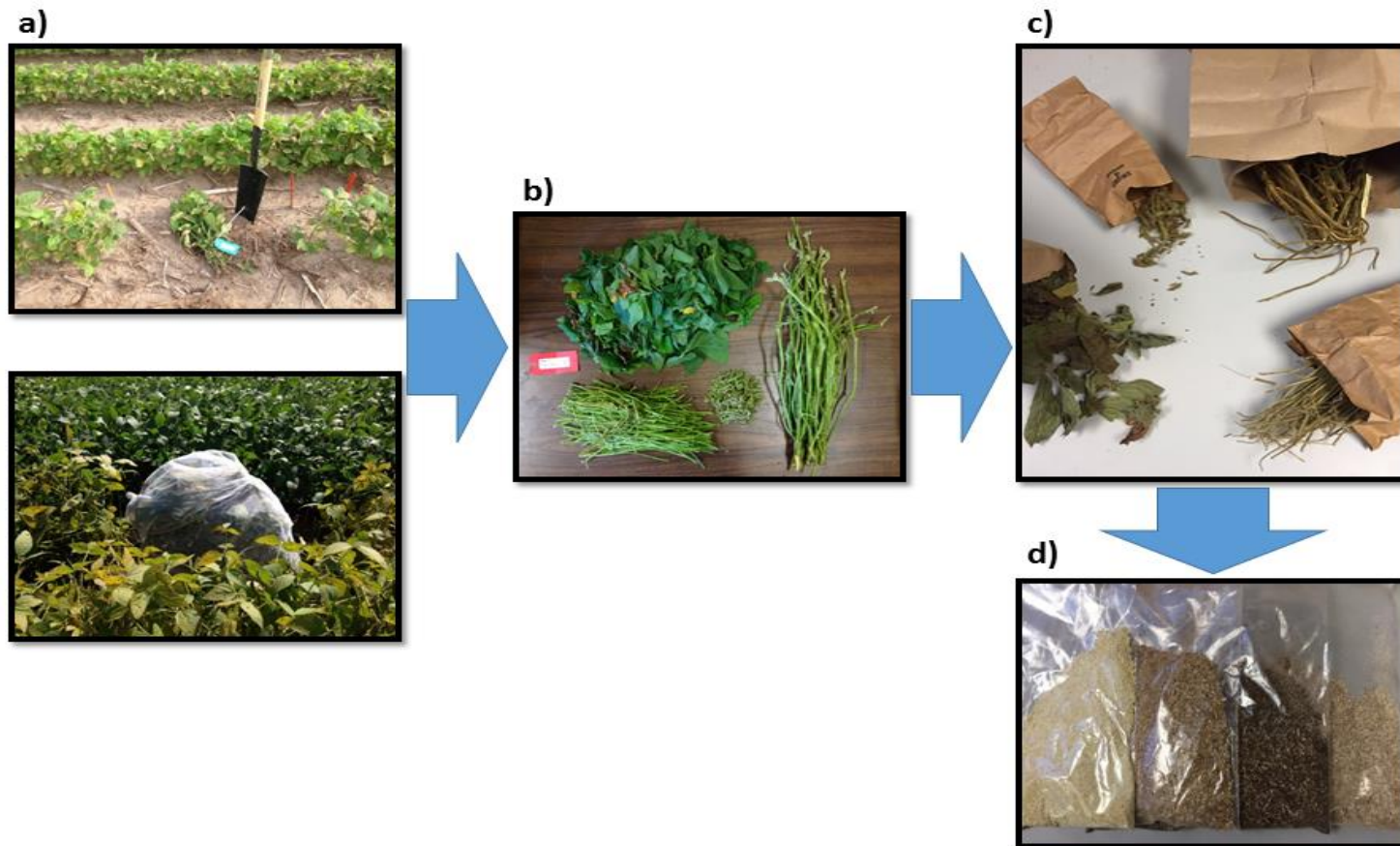


Figure 3.1 Flow diagram of biomass sample collection and partitioning . a) 3ft sections of soybeans destructively harvested above ground. R8 samples encased in insect mesh bags at R7 for fallen residue collection. b) Whole plant samples partitioned into plant organ components. c) Samples dried at 60 degrees C and weighed. d) R8 plant samples finely ground to evaluate carbon and nitrogen content through combustion analysis.

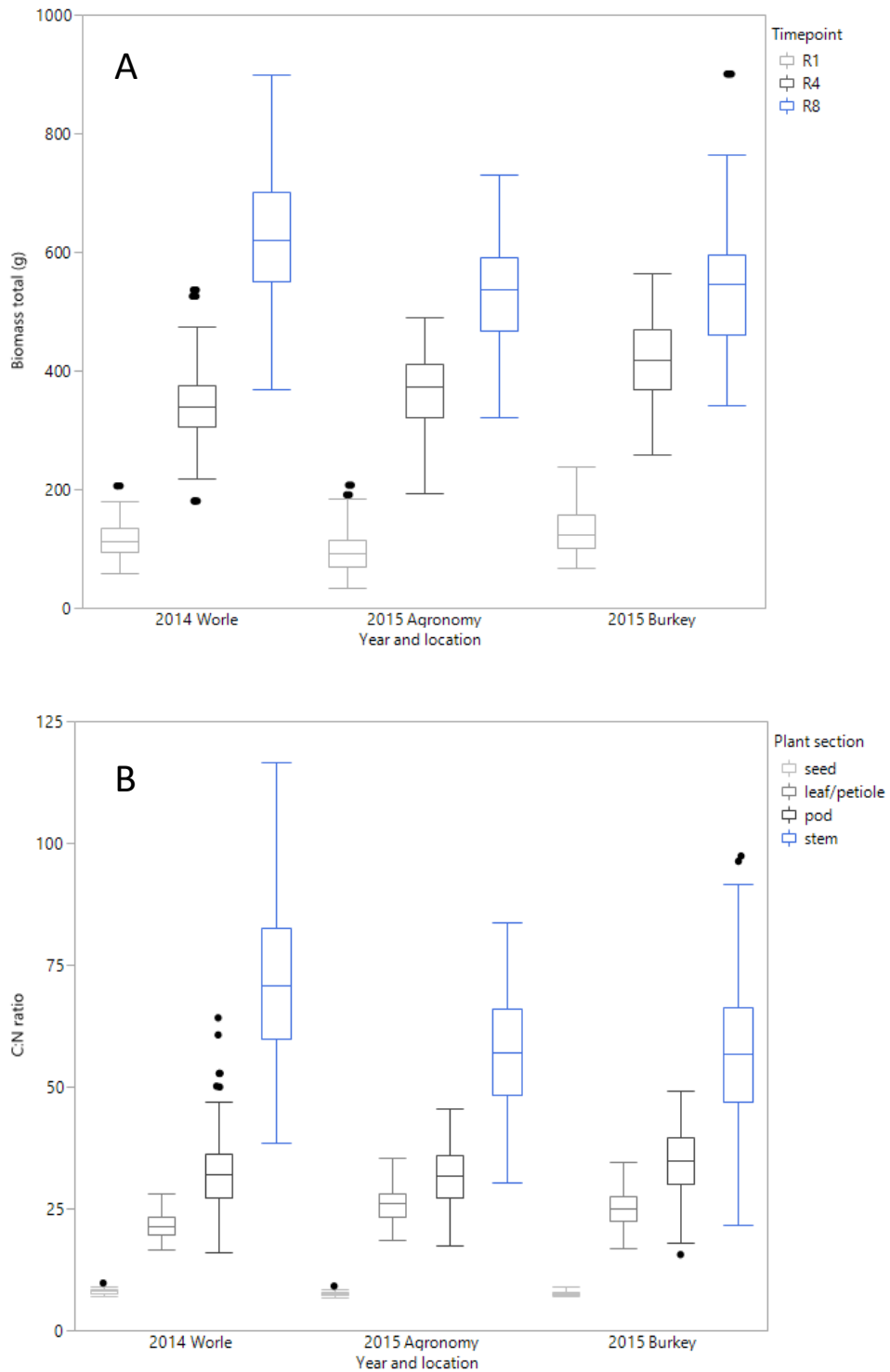


Figure 3.2 Boxplots of a) biomass weight (g) by experiment year-location and growth stage and b) average C:N ratio by experiment year-location. Differences between experiment years were significant for biomass and C:N ratio.

Table 3.2 Mean, range, standard deviation, ANOVA genotype factor significance, and heritability for biomass weights and percentage for three growth stages (R1, R4, R8) averaged over 32 genotypes, three locations, and three blocks in Central Iowa years 2014-2015.

Growth stage	Plant section	Biomass weight (g)					Biomass percentage (%)					
		Mean	Range	SD	P-value	H ^{2†}	Mean	Range	SD	P-value	H ^{2†}	P-value ^c
R1	leaf	59.2	21.7-127.7	20.4	<.0001	0.72	52.6	32.8-66	5.2	0.8682	-	-
	petiole	19.7	4.9-43.2	7.1	<.0001	0.80	17.3	11.8-23.6	1.8	0.1939	-	-
	stem	34.1	6.3-73.1	12.1	<.0001	0.74	30.1	18.9-50.4	4.9	0.9631	-	-
R4	leaf	121.7	43.8-202.9	22.9	0.0524	0.39	33.0	19.5-43.9	5.1	0.0088	0.51	<.0001
	petiole	62.7	13-105.3	17.2	<.0001	0.82	16.7	6.7-23.3	2.9	<.0001	0.84	<.0001
	pod	45.8	10.5-137.1	27.7	<.0001	0.81	12.0	2.9-32.7	6.3	<.0001	0.89	-
	stem	143.2	65.6-240.4	32.0	<.0001	0.78	38.4	25.1-49.8	4.5	<.0001	0.84	<.0001
R8 [‡]	leaf/petiole	127.2	34.7-237.7	37.1	<.0001	0.75	37.9	8.6-34	4.4	0.0064	0.52	0.0205
	pod	87.2	29.6-195	33.7	0.0042	0.56	15.2	7.7-27.5	4.2	<.0001	0.83	-
	stem	137.0	60.8-278.8	32.7	<.0001	0.89	24.4	14.7-37	4.4	<.0001	0.86	0.0205
	seed	213.1	68.9-336.5	49.5	0.0002	0.66	37.9	17.7-53.5	6.0	<.0001	0.83	-

[†]H², broad-sense heritability, calculated on an entry mean basis

[‡]R8 plant sections as percentage of total residue excluding seed. Harvest index as seed weight percentage of total biomass.

[§]Non-reproductive organ percentage (%)

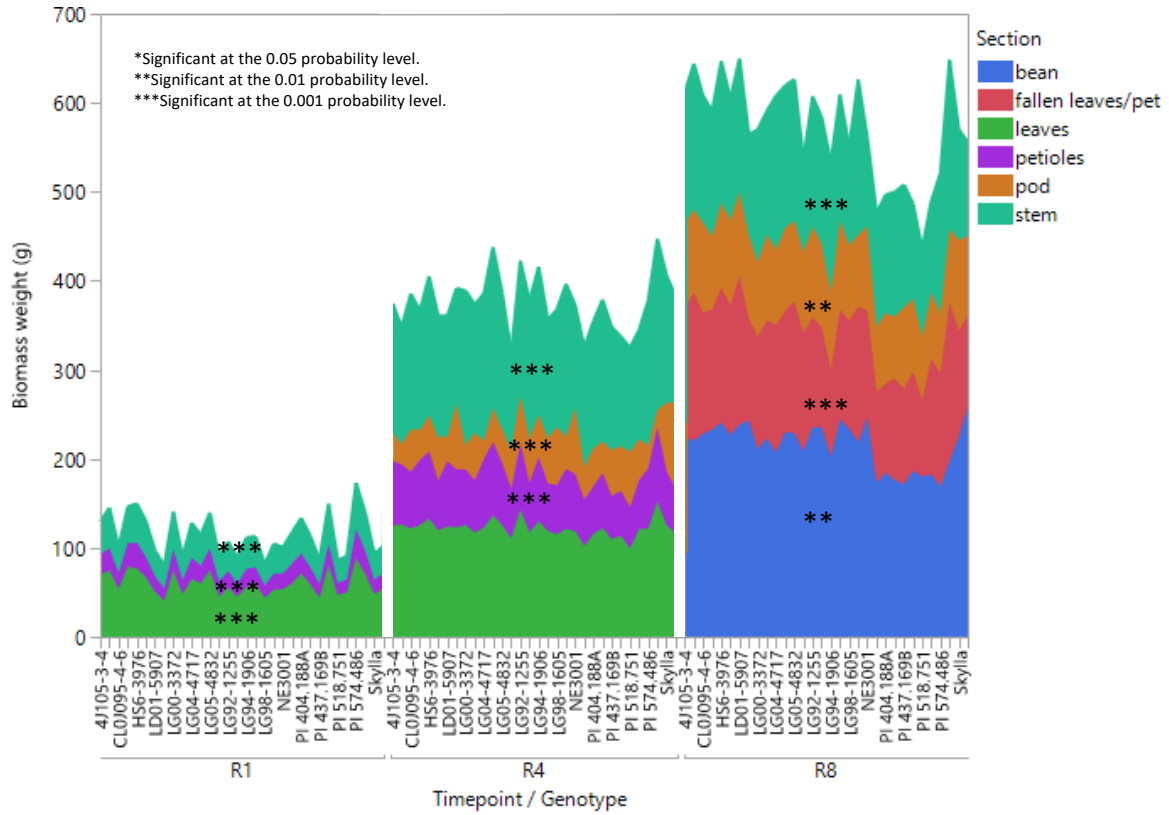


Figure 3.3 Biomass component dry weights by reproductive growth stage and genotype. Biomass weights were significantly different for all components and growth stages.

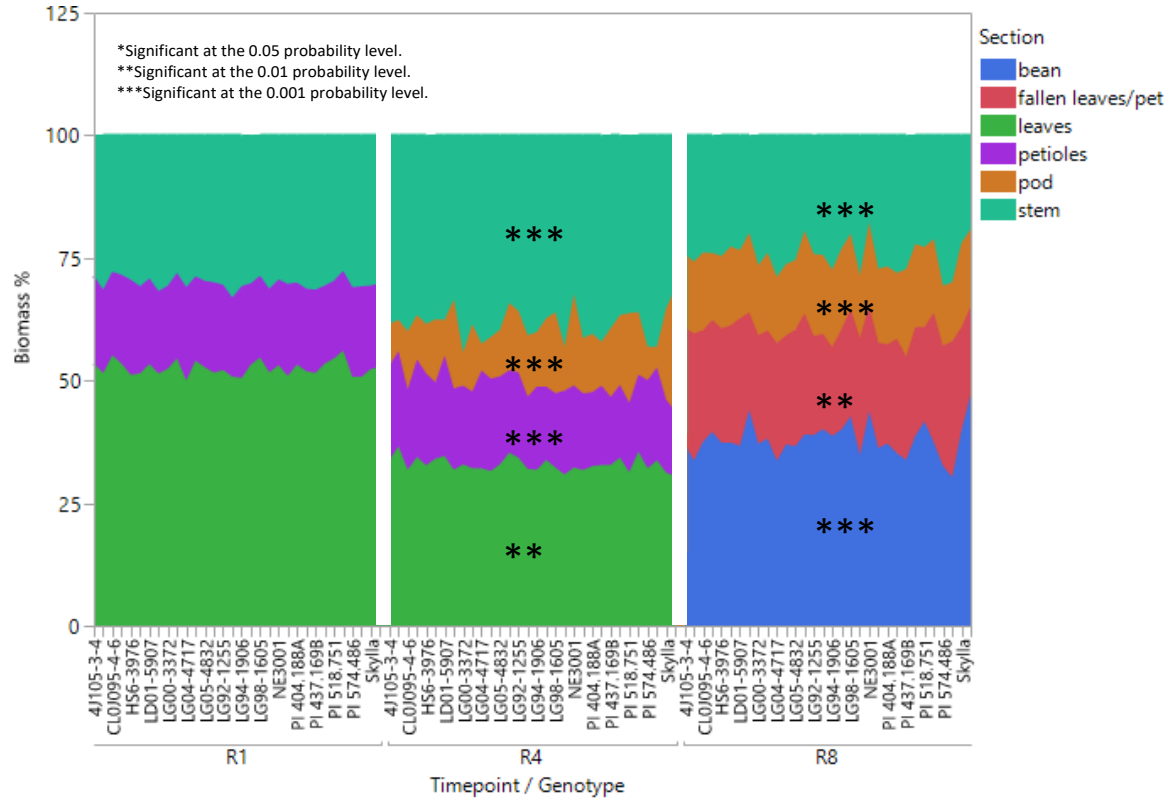


Figure 3.4 Biomass component percentages by reproductive growth stage and genotype. Biomass percentages were only significantly different after flowering (R1).

Table 3.3 Pearson's correlations and significance for biomass percentages, total biomass weights, and agronomic traits at growth stages R1, R4, and R8.

	R1				R4					R8					†R4			†R8		Agronomic traits				
	leaf (%)	petiole (%)	stem (%)	total (g)	leaf (%)	petiole (%)	pod (%)	stem (%)	total (g)	leaf/ petiole		HI seed			leaf (%)	petiole (%)	stem (%)	leaf (%)	stem (%)	seed protein (%)	seed oil (%)	maturity		
										(%)	pod (%)	stem (%)	(%)	total (g)								(%)	(%)	(%)
R1		-0.34	-0.94	-0.12	-0.41	0.22	0.46	-0.32	0.18	-0.12	-0.47	-0.08	0.48	-0.26	-0.24	0.4	-0.05	-0.03	0.03	0.1	0.29	-0.11	0.1	
	petiole (%)	<.0001***	0	0.23	-0.06	0.16	-0.18	0.23	0.11	0.13	-0.1	0.19	-0.16	0.09	-0.19	0.08	0.15	-0.02	0.02	0.05	-0.09	0.26	-0.02	
	stem (%)	<.0001***	0.9701	0.04	0.46	-0.29	-0.42	0.25	-0.23	0.08	0.54	0.02	-0.45	0.24	0.32	-0.45	0	0.04	-0.04	-0.12	-0.27	0.03	-0.1	
	total (g)	0.0521	0.0001*	0.4939	-0.06	0.1	-0.22	0.31	0.26	0	-0.12	0.36	-0.18	0.07	-0.2	0.01	0.22	-0.22	0.22	0.14	-0.17	0.26	-0.06	
R4	leaf (%)	<.0001***	0.304	<.0001***	0.3363	-0.28	-0.6	-0.11	-0.41	0.22	0.69	-0.17	-0.52	0.33	0.89	-0.52	-0.59	0.24	-0.24	-0.11	-0.38	0.06	-0.1	
	petiole (%)	0.0004*	0.011*	<.0001***	0.0974	<.0001***	-0.19	-0.06	0.25	0.2	-0.53	0.21	0.07	-0.01	-0.45	0.91	-0.22	0	0	0.16	0.15	0.46	0.05	
	pod (%)	<.0001***	0.0032*	<.0001***	0.0003*	<.0001***	0.0016*	-0.6	0.23	-0.41	-0.26	-0.4	0.77	-0.27	-0.17	0.22	0.01	-0.03	0.03	-0.14	0.4	-0.56	0.3	
	stem (%)	<.0001***	0.0002*	<.0001***	<.0001***	0.0655	0.3516	<.0001***	-0.02	0.18	-0.07	0.6	-0.53	0.02	-0.48	-0.31	0.8	-0.23	0.23	0.22	-0.23	0.42	-0.33	
	total (g)	0.0035*	0.066	0.0001*	<.0001***	<.0001***	<.0001***	0.0001*	0.7116	-0.06	-0.39	0.17	0.19	0.11	-0.37	0.34	0.14	-0.14	0.14	0.06	0.19	0.15	0.22	
R8†	leaf/petiole (%)	0.0452*	0.0327*	0.1787	0.9537	0.0002*	0.0011*	<.0001***	0.0025*	0.3062														
	pod (%)	<.0001***	0.0885	<.0001***	0.0535	<.0001***	<.0001***	<.0001***	0.2653	<.0001***	0.1046	-0.1	-0.13	-0.57	0.11	0.04	0.03	-0.07	0.8	-0.8	0.1	-0.17	0.36	-0.27
	stem (%)	0.1795	0.002*	0.7496	<.0001***	0.0059*	0.0005*	<.0001***	<.0001***	0.0061*	0.032*	<.0001***	-0.41	-0.21	-0.43	0.04	0.45	-0.69	0.69	0.27	-0.15	0.38	-0.44	
	seed (%)	<.0001***	0.0072*	<.0001***	0.0029*	<.0001***	0.2528	<.0001***	<.0001***	0.0015*	<.0001***	<.0001***	<.0001***	<.0001***	-0.12	-0.19	0.38	-0.09	-0.16	0.16	-0.1	0.41	-0.38	0.56
	total (g)	<.0001***	0.1459	<.0001***	0.2252	<.0001***	0.864	<.0001***	0.7431	0.0848	0.0658	<.0001***	0.0007*	0.0483*	0.24	-0.11	-0.17	0.2	-0.2	-0.13	-0.07	0.29	0.75	
†R4	leaf (%)	<.0001***	0.0021*	<.0001***	0.0013*	<.0001***	<.0001***	0.0067*	<.0001***	<.0001***	0.4907	<.0001***	<.0001***	0.0017*	<.0001***	-0.5	-0.72	0.28	-0.28	-0.21	-0.23	-0.26	0.05	
	petiole (%)	<.0001***	0.2017	<.0001***	0.9348	<.0001***	<.0001***	0.0003*	<.0001***	<.0001***	0.6079	<.0001***	0.5571	<.0001***	0.065	<.0001***		-0.23	0	0.09	0.31	0.22	0.18	
	stem (%)	0.4262	0.0154*	0.9651	0.0003*	<.0001***	0.0002*	0.8592	<.0001***	0.0214*	0.2468	<.0001***	<.0001***	0.1451	0.0049*	<.0001***	0.0002*	-0.31	0.31	0.17	0.01	0.11	-0.19	
†R8	leaf (%)	0.5846	0.7165	0.4806	0.0003*	<.0001***	0.9465	0.6117	0.0002*	0.0255*	<.0001***	0.0549	<.0001***	0.0088*	0.0011*	<.0001***	0.9973	<.0001***	-1	-0.11	-0.02	0.03	0.07	
	stem (%)	0.5846	0.7165	0.4806	0.0003*	<.0001***	0.9465	0.6117	0.0002*	0.0255*	<.0001***	0.0549	<.0001***	0.0088*	0.0011*	<.0001***	0.9973	<.0001***	<.0001***	0.11	0.02	-0.03	-0.07	
Agronomic traits	seed protein (%)	0.1221	0.4304	0.0528	0.0189*	0.0795	0.0079*	0.0187*	0.0003*	0.3605	0.0924	<.0001***	<.0001***	0.1173	0.0414*	0.0007*	0.1666	0.0065*	0.0806	0.0806		-0.53	0.12	-0.17
	seed oil (%)	<.0001***	0.1666	<.0001***	0.0057*	<.0001***	0.0132*	<.0001***	0.0001*	0.0023*	0.0059*	<.0001***	0.0128*	<.0001***	0.2908	0.0001*	<.0001***	0.8788	0.7403	0.7403	<.0001***		-0.05	0.23
	maturity (DAP)	0.0627	<.0001***	0.6698	<.0001***	0.3527	<.0001***	<.0001***	<.0001***	0.0164*	<.0001***	0.0002*	<.0001***	<.0001***	<.0001***	<.0001***	0.0002*	0.0767	0.5959	0.5959	0.0547	0.4461		0.01
	seed (g)	0.1119	0.7856	0.1098	0.3248	0.1127	0.4059	<.0001***	<.0001***	0.0004*	<.0001***	0.3824	<.0001***	<.0001***	<.0001***	0.4165	0.0038*	0.0015*	0.2656	0.2656	0.0062*	0.0002*	0.8961	

[†]H², broad-sense heritability, calculated on an entry mean basis

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

Table 3.4 Pearson's correlations and significance for seed weight and harvest index with biomass percentages at growth stages R1, R4, and R8 and agronomic traits.

		Seed (g)		Harvest index (seed %)	
		<i>r</i>	<i>P</i> -value	<i>r</i>	<i>P</i> -value
R1	leaf (%)	-	0.1119	0.48	<.0001***
	petiole (%)	-	0.7856	-0.16	0.0072*
	stem (%)	-	0.1098	-0.45	<.0001***
	total (g)	-	0.3248	-0.18	0.0029*
R4	leaf (%)	-	0.1127	-0.52	<.0001***
	petiole (%)	-	0.4059	-	0.2528
	pod (%)	0.30	<.0001***	0.77	<.0001***
	stem (%)	-0.33	<.0001***	-0.53	<.0001***
	total (g)	0.22	0.0004*	0.19	0.0015*
R8	leaf/petiole (%)	-0.27	<.0001***	-0.57	<.0001***
	pod (%)	-	0.3824	-0.39	<.0001***
	stem (%)	-0.44	<.0001***	-0.41	<.0001***
	seed (%)	0.56	<.0001***	-	-
	total (g)	0.75	<.0001***	-0.12	0.0007*
†R4	leaf (%)	-	0.4165	-0.19	<.0001***
	petiole (%)	0.18	0.0038*	0.38	0.5571
	stem (%)	-0.19	0.0015*	-0.09	<.0001***
†R8	leaf (%)	-	0.2656	-0.16	<.0001***
	stem (%)	-	0.2656	0.16	<.0001***
Agronomic traits	seed protein (%)	-0.17	0.0062*	-0.10	<.0001***
	seed oil (%)	0.23	0.0002*	0.41	0.0128*
	maturity (DAP)	-	0.8961	-0.38	<.0001***
	seed (g)	-	-	0.56	<.0001***

†Non-reproductive organ percentage (%)

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

Table 3.5 Mean, range, standard deviation, ANOVA genotype factor significance, and heritability for C:N ratios for three growth stages (R1, R4, R8) averaged over 32 genotypes, three locations, and three blocks in Central Iowa years 2014-2015.

Plant section	Biomass weight (g)				
	Mean	Range	SD	<i>P</i> -value	$H^{2†}$
leaf/petiole	24.5	16.5-38	4.3	<.0001	0.90
pod	32.9	15.6-64.1	7.4	<.0001	0.86
seed	7.8	6.6-9.7	0.5	<.0001	0.96
stem	62.3	21.7-116.5	15.7	0.0002	0.78

† H^2 , broad-sense heritability, calculated on an entry mean basis

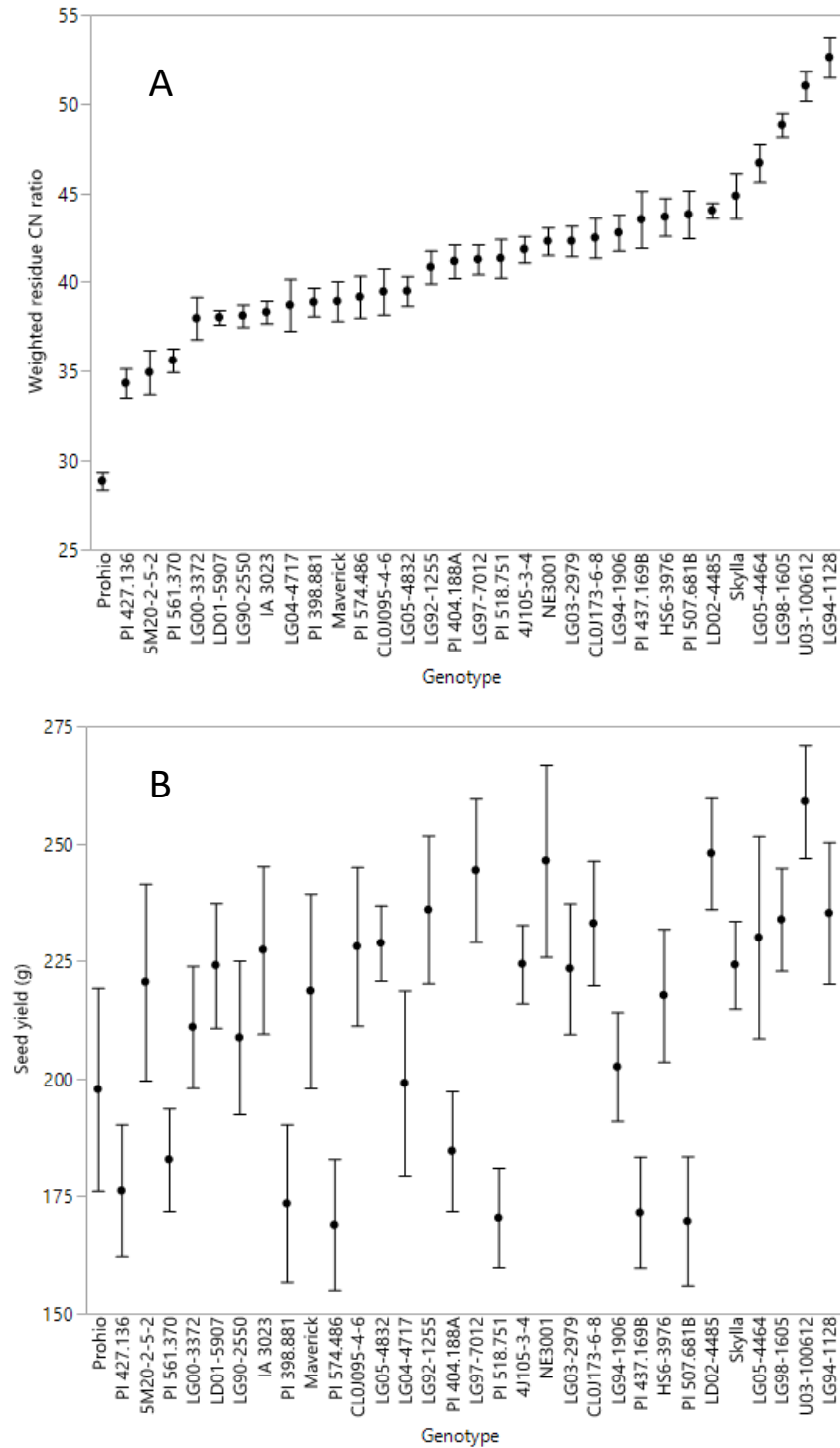


Figure 3.5 Plot of a) weighted C:N ratios by biomass composition for all soybean genotypes of the final residue biomass excluding seed and b) mean seed yield of the previously ordered genotypes by C:N ratio, displaying no apparent relationship between seed yield and C:N ratio of final residue biomass for this genotype panel. Error bars are constructed using one standard error from the mean.

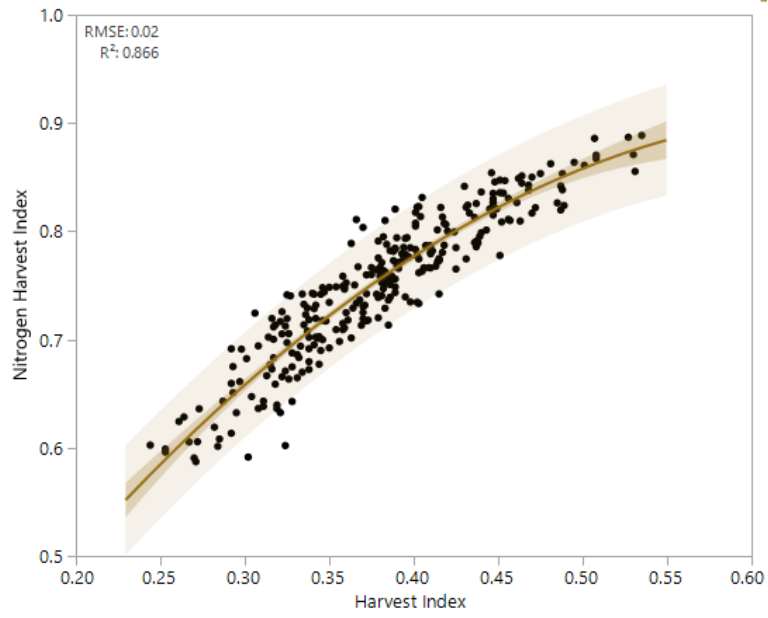


Figure 3.6 Curvilinear relationship between nitrogen harvest index and harvest index.

CHAPTER 4. LINKAGE MAPPING OF QTL FOR YIELD-RELATED VEGETATION INDICES IN A SOYBEAN NESTED ASSOCIATION MAPPING POPULATION

Race Higgins, Asheesh K Singh

Abstract

Physiological traits estimated through remote sensing beneficially supplement plant breeding programs to accelerate yield gain. Further identification of QTL and linked molecular markers for yield-related vegetative indices will enhance marker-assisted selection (MAS) in crop breeding. In the present study, a mapping population consisting of 535 F₅ derived RILs from the cross of universal hub parent IA3023 and four parent genotypes HS6-3976, NE3001, LG90-2550, and LG94-1128 of the soybean nested association mapping (SoyNAM) population were studied. Absolute reflectance was collected at two soybean growth stages, beginning bloom (R1) and beginning seed (R5), from where 14 vegetative indices representing physiological trait categories of carotenoids, chlorophyll, dry matter, vegetation, and water content were calculated. Field trials were performed in Central Iowa at Iowa State University's farm experiment stations during the 2015 and 2016 cropping seasons, providing data for three environments. Analysis of variance (ANOVA) did not reveal significant differences ($P < 0.01$) among RILs for all vegetative indices, and Broad-sense heritabilities were low. However, maximum correlations were identified for vegetative indices VREI2 ($r = -0.42$) and PRI ($r = 0.42$) at R5. A linkage map spanning 1,750.1 cM was constructed using 561 polymorphic SNP markers, with an average marker density of 3.12 cM/marker. Five QTL were detected for grain yield and indices NDVI, NMDI, NWIB, PSRI, and VREI2 measured at R5, spanning chromosomes 1, 3, 10 and 18 and explaining 2.89-5.29% of the phenotypic variance. These QTL can serve as aides to MAS in soybean breeding and inform future studies aimed at dissecting the physiology of soybean grain yield.

Abbreviations

CAI, cellulose absorption index; DMCI, dry matter content index; GY, grain yield; LAI, leaf area index; MAS, marker-assisted selection; NAM, nested association mapping; NDLI, normalized difference lignin index ; NDVI, normalized difference vegetation index; NMDI, normalized multi-band drought index; NWIB, normalized water index 2; PRI, photochemical reflectance index; PSRI, plant senescence reflectance index; QTL, quantitative trait locus; RARSa, ratio analysis of reflectance spectra chlorophyll A; RARSb, ratio analysis of reflectance spectra chlorophyll B; RIL, recombinant inbred line; SNP, single nucleotide polymorphism; VREI2, Vogelmann red edge index 2.

1. Introduction

Soybean [*Glycine max* (L.) Merrill] is a highly valued crop grown throughout much of the world for many purposes. The primary components are seed protein and oil, with processed soybeans serving as the world's largest source of animal protein feed and the second largest source of vegetable oil (USDA-NASS, 2016). Food security is and will continue to be a grave concern for the future due to growing global population, changes in available arable land, increased input costs, and predicted climate change impacts on crop yield. Therefore, it is very important to increase the yields of major commodities and all food crops to avert predicted food security crises (Lipper et al., 2014).

Soybean grain yield is a complex quantitative trait, and realization of the maximum yield potential is influenced by physiological and agronomic traits including light interception, photosynthetic capacity, and biomass partitioning (Monteith and Moss, 1977). Historically, the steady increase in soybean grain yield has been attained through empirical selection for grain yield over the past century. However, there is evidence that phenotyping for physiological traits,

as a complement to agronomic traits, may help in identifying selectable features that accelerate breeding for yield potential (Araus and Cairns, 2014; Keep et al., 2016). Currently, the soybean genetic base is narrow with low diversity, due to a genetic bottleneck after introduction to the US (Rincker et al., 2014). Introgressing exotic germplasm into cultivars to increase genetic diversity within domesticated crops has been used to enhance complex traits such as yield (Tanksley and McCouch, 1997) and may have unknowingly introduced novel genetic variation for yield-related physiological traits. In soybean, Thompson and Nelson (1998) tested experimental lines derived from crossing North American cultivars with several plant introductions, and several of these lines were incorporated into the soybean nested association mapping (SoyNAM) parent panel, including LG90-2550 and LG94-1128 of this study. An experimental population of high-yielding elite lines is enriched by including lines of diverse ancestry because it increases morphological and genetic diversity.

Many changes in morphological and physiological traits in soybean have accompanied changes in grain yield. Potential soybean yield is closely associated with plant photosynthesis (Slattery et al., 2017) and chlorophyll concentration is a robust indicator of photosynthetic capacity and primary production, the rate at which a crop can capture and store chemical energy (Gitelson et al., 2003; Koester et al., 2016). Changes in leaf relative water content affect total water potential, osmotic potential, and turgor pressure, and therefore influence whole-plant physiology. Only when there is sufficient turgor pressure can cells expand for vegetative growth and stomata to open to incorporate carbon dioxide to be used in the Calvin cycle. (Zygielbaum et al., 2012; Gray et al., 2016). Canopy water content is indicative of canopy transpiration and determines radiation use efficiency and biomass accumulation in soybean (Saryoko et al., 2018). Biomass accumulation has long been established as important driver of potential yield, and is

often estimated at leaf area index, or LAI (Ma et al., 1995; Board and Harville, 1996). Although genetic improvement of physiological traits can certainly lead to increased grain yield (GY), high-throughput, nondestructive measurements are necessary to rapidly collect many phenotypes for large mapping populations.

Remote sensing is a promising tool that rapidly and non-destructively collects vegetative indices related to chlorophyll content, carotenoids, vegetation, water content, and dry matter content that are used in soybean to predict yield (Ma et al., 2001; Bolton and Friedl, 2013; Johnson, 2014) and measure plant response to stress (Carter, 1994; Nutter Jr et al., 2002; Huang et al., 2016). An enormous number of spectral reflectance indices have been created to monitor vegetation health and productivity (Heinrich et al., 2011). Some indices have served as the industry standard for analyzing canopy “greenness” and detection of vegetation, such as the normalized difference vegetation index (Rouse Jr et al., 1974). However, many different indices have been developed depending on the specific trait to be monitored, and great advances in remote and proximal sensing technologies are currently underway. One advance has been the development of hyperspectral reflectance instruments (Haboudane et al., 2004). The major advantage of hyperspectral reflectance is that it allows users to calculate any number of desired spectral reflectance indices pertinent to a trait of interest (Heinrich et al., 2011).

Quantitative trait loci (QTL) mapping is a key approach for understanding the genetic architecture of yield components and physiological traits in crops. However, pinpointing QTL can be hampered by relatively large QTL intervals due to the limited number of markers. Nested association mapping is an alternative population design that was proposed to increase the resolution of QTL mapping (Yu and Buckler, 2006). Nested association mapping populations are developed by crossing multiple diverse founders to a common parent followed by the

development of recombinant inbred lines (RILs) or progenies in each family. In comparison to traditional QTL mapping, which only uses limited genetic information from two parents, NAM can increase genetic variation across contributing parental lines, increase genetic resolution, reduce linkage disequilibrium, and control population structure through design (Rafalski, 2010). The NAM design has been used successfully in soybean to map QTL controlling a number of traits such as grain yield stability (Xavier et al., 2018) and canopy coverage (Xavier et al., 2017).

The objective of this study was to identify QTL for yield-related physiological traits estimated through remote sensing. To accomplish this, nondestructive hyperspectral reflectance measurements were obtained on a set of four RIL populations derived from a subset of the soybean NAM population over two reproductive growth stages, beginning bloom (R1) and seed fill (R5). These measurements were used to calculate 14 vegetative indices that represented physiological traits of interest. Results from this study will provide knowledge of the relationship of physiological traits with grain yield and identify linked SNP markers for marker-assisted selection (MAS) in soybean breeding. These data will further suggest which reproductive stage any particular vegetative index may have higher QTL detection and inform future studies aimed at dissecting the physiology of soybean grain yield.

2. Materials and Methods

2.1 Plant Materials and Field Trials

A total of 560 F₅ RILs derived from the cross of universal hub parent IA 3023 and NAM parent genotypes HS6-3976, NE3001, LG90-2550, and LG94-1128 were phenotyped in this study (Table 4.1). RIL populations consisting of 140 RILs per population were previously developed and genomic and multi-environmental data for the soybean nested association mapping dataset is public (Xavier et al., 2015). Field trials were performed in Central Iowa at

Iowa State University's farm experiment stations during the 2015 and 2016 cropping seasons., providing data for three environments. It should be noted that only two environments of data in 2016 were collected for the flowering growth stage R1. The RILs were planted in randomized complete blocks, single replication, with RIL families nested within block. Parent genotypes served as checks within each block and family and IA3023 as a common check between blocks. Plots consisted of two 2.1 m rows with 76.2 cm between rows. The target seeding rate was 57k plants ha⁻¹ near the commercial standard. The field trials were managed following local normal practice and weeds were chemically controlled with a single application of Cobra® herbicide before beginning bloom (R1) and hand removed during crop growth.

2.2 Phenotyping

Absolute reflectance remote sensing data were measured using with a FieldSpec® 4 Hi-Res (ASD Inc., Boulder, CO), which ranges from 350 to 2500 nm with a single nanometer resolution. The portable spectroradiometer was mounted as a backpack. A white reference panel (Specralon® Labsphere Inc., North Dutton, NH) reading for remote-sensing was collected at the beginning of each field block to serve as a control. All light measurements were performed at 1000 h to 1400 on clear days. Measurements were collected directly above one of the two canopy rows in each plot. Canopy reflectance was captured as a single timepoint at two soybean reproductive growth stages: beginning bloom (R1) and seed fill (R5) per (Fehr et al., 1971) when approximately 50% of the plots were rated as the target growth stage. Grain yield (GY) was collected at each location for the whole plot with a two-row ALMACO = combine. GY was determined as weight of grain harvested per unit area (kg/ha).

2.3 Phenotypic Data Analysis

ASD Spectra were processed using ViewSpec pro software, and single and multiple wavebands of absolute reflectance were extracted with R software to calculate vegetative indices. An infinite number of indices are available with many either highly correlated or near-identical. A subset of indices related to traits of interest were selected to reduce redundancy and were summarized (Table 4.2). For each vegetative index and grain yield, outliers were detected and eliminated using interquartile range at each location and growth stage independently. Grain yield was additionally spatially adjusted for environmental variation using a moving grid of six adjacent plots total with R package “mvngGrAd” (Technow, 2011).

All statistical analyses were performed using R software v.3.4.0 with REML estimation method under R package “lme4” (Bates et al., 2014). An ANOVA was conducted to assess the significance of genotype, environment, genotype x environment interaction, and check x block interaction nested within environment on yield and vegetative indices in a fixed effects model. For each trait, the “Best linear unbiased predictions” (BLUPs) of the genotype effects were extracted from completely random effects models. Trait BLUPs were used for QTL mapping when significant, i.e. not estimated as a single factor for genotype. Broad-sense heritability (H^2) for each vegetative index and grain yield was estimated on an entry mean basis following (Nyquist and Baker, 1991):

$$H^2 = (\sigma^2_G) / [\sigma^2_G + \left(\frac{\sigma^2_{GE}}{e}\right) + \left(\frac{\sigma^2_e}{re}\right)]$$

in which σ^2_G is genetic variance, σ^2_{GE} is genotype x environment (location and year) variance, σ^2_e is error variance, r is number of replications, and e is the number of environments when genotype was significant.

2.4 Linkage Map Construction and QTL Analysis

The 560 RILs in this study are a subset of the larger soybean nested association mapping (SoyNAM) population that contains 5,555 RILs. Lines were previously genotyped in the F₅ generation with the SoyNAM6K BeadChip (Xavier et al., 2016). The chip was designed using SNPs discovered after complete sequencing of the DNA of all 41 parental lines to minimize bias by sampling issues associated with rare variants. Non-segregating SNPs, variants with a minor allele frequency (MAF) lower than 0.15, and redundant markers were removed in the original dataset (Xavier et al., 2016). Quality assured genotype information was retrieved from SoyBase for all SoyNAM parents and progeny using WM82.a2 coordinates (Grant et al., 2009). A small number of RILs from each population did not have genotype information provided (Table 4.1). SNP markers were anchored by chromosome and filtered using IciMapping 4.1 software (Meng et al., 2015). Among the initial 4,273 polymorphic SNP markers available, 561 SNPs were retained after filtering SNPs with large numbers of missing values (15% or more). 12 markers were discarded for having no chromosome position.

QTL analysis was performed using inclusive composite interval mapping (ICIM) with IciMapping 4.1 software. Map distances between markers were calculated with the Kosambi mapping function. The walking speed chosen for all QTL was 1.0 cM, with $P = 0.001$ in stepwise regression. Each trait by timepoint was filtered by a threshold determined by 1,000 permutations at a probability level of $\alpha=0.05$. Maximum trait LOD scores ranged from 5.6 to 11.0, justifying individual thresholds by trait, although most traits were similar in LOD threshold. Each QTL was represented by a 20 cM interval with the LOD maximum as center. The phenotypic variance explained (PVE) was estimated through stepwise regression.

3. Results

3.1 Phenotypic Evaluation

Vegetative indices and yield were strongly influenced by environment and vegetative indices by growth stage (Figure 4.1). Across traits, environments 2 and 3 from 2016 were slightly more similar to each other than environment 1 from 2015, although this observation was not consistent. Distributions of traits between environments tended to be more similar within R5 than R1 for many traits. Overall yield was lower in 2015 than both environments in 2016 (Figure 4.1). Standard deviations for traits on average were lower for R5 than R1 (Table 4.3). The lower count of vegetative indices at R1 is from a missing environment (Table 4.3).

ANOVA were conducted and Broad-sense heritabilities calculated for each vegetative trait by growth stage and yield (Table 4.4). There were no significant differences among the 560 RILs for LAI, NMDI, RARSb at R1 and DMCI and Green at R5. The RedEdge summary index was not significant for genotype at either growth stage. However, significant differences for genotype were found among most traits. Environment was significant for every trait at each growth stage, confirming a strong observed location effect. The interaction effects $check \times block(environment)$ and $genotype \times environment$ were significant for some traits with no clear pattern. Neither interactions were significant for grain yield. The highest heritability observed was for grain yield at 0.45 (Table 4.4) and may be due to the spatial adjustment yield received in comparison to the vegetative indices. Vegetative heritabilities were low and ranged from 0-0.30. Of this range, VREI2, PRI, and NDVI notably had higher heritabilities between both reproductive growth stages, ranging from 0.18-0.25, and heritabilities for NMDI and PSRI fit within this range exclusively at R5 (Table 4.4).

3.2 Correlations Between Traits

Pearson's coefficients of correlation were calculated for BLUPs based on data from either two environments for R1 vegetative indices or three for R5 vegetative indices and yield (Table 4.5). The maximum correlation was between VREI2 and RARSb ($r = -0.72$) at R5, and the maximum correlation at R1 between NDVI and RARSa ($r = 0.65$) (Table 4.5). Greater correlations were observed within growth stage and almost absent between growth stage (Figure 4.2). Greater vegetative index correlations were found within R5 than R1. In relationship to yield, the maximum correlation was for VREI2 ($r = -0.42$) and PRI ($r = 0.42$) each at R5 (Figure 4.3). Other vegetative indices with significant negative correlations with yield included NWIB at R5 ($r = -0.33$), PSRI at R5 ($r = -0.31$), and VREI2 at R1 ($r = -0.15$) and significant positive correlations NDVI at R5 ($r = 0.34$), RARSb at R5 ($r = 0.25$), NMDI at R5 ($r = 0.22$), LAI at R5 ($r = 0.21$), NDLI at R5 ($r = 0.14$), and PRI at R1 ($r = 0.11$).

3.3 Linkage Map Construction

Twenty linkage groups corresponding to the 20 haploid soybean chromosomes were constructed from the 561 polymorphic SNP markers with a total length of 1750.11 cM, smaller than the 2291.64 cM reported for the soybean consensus map 3.0 (Grant et al., 2009) (Table 4.6), but very similar to the composite genetic linkage map of 1736 cM based on all of the RILs for the soybean NAM population. (Song et al., 2017). The number of SNP markers in each soybean chromosome ranged from 11 mapped on chromosome 9 to 69 on chromosome 2. The SNP markers were moderately well distributed throughout the genome, although chromosomes 4 and 18 exhibited lower marker densities (Figure 4.4). The overall SNP density was 3.12 cM, with the highest density of 0.86 cM on chromosome 1, and the lowest density of 10.93 cM on chromosome 18 (Table 4.6).

3.4 QTL Analysis of Grain Yield and Vegetative Indices

ICIM identified five distinct QTL for BLUPs of grain yield and the following five vegetative indices at seed fill (R5): NDVI, NMDI, NWIB, PSRI, and VREI2 (Table 4.7). No QTL were detected for vegetative indices at beginning bloom (R1). QTL were detected on four soybean chromosomes: 1, 3, 10, and 18 (Figure 4.4). QTL by trait are denoted as *QNDVI.R5-chr1*, *QNDVI.R5-chr10*, *QNMDI.R5-chr18*, *QNWIB.R5-chr1*, *QPSRI.R5-chr3*, *QVREI2.R5-chr1*, and *QGY-chr18* and together ranged from 2.89-5.29% phenotypic variance explained (PVE) (Table 4.7). Positive additive effects were desirable for grain yield, NDVI, and NMDI, and negative additive effects desirable for NWIB, PSRI, and VREI2 (Figure 4.3). All four NAM families contributed to QTL detection (Table 4.7).

Two distinct QTL for NDVI were identified on chromosomes 1 and 10, with *QNDVI.R5-chr1* explaining 3.08% and *QNDVI.R5-chr10* explaining 2.89% of the phenotypic variance. The positive alleles for *QNDVI.R5-chr1* with the greatest additive effect derived from NAM 31, genotype LG90-2550, and *QNDVI.R5-chr10* the greatest additive effect derived from NAM 33, genotype NE3001. One distinct QTL for PSRI was identified on chromosome 3. *QPSRI.R5-chr3* explained 5.29% of the phenotypic variance, and alleles from NAM 31 and 18 contributed equally with the favored negative additive effect (Table 4.7).

Two QTL with pleiotropic effects were detected on chromosomes 1 and 18. On chromosome 1 at 12 cM, the single QTL for NWIB, *QNWIB.R5-chr1*, explained 5.37% of the phenotypic variance and alleles from NAM 31 and 18 contributed to the wanted negative additive effect. At the same position, *QVREI2.R5-chr1* explained 4.74% of the phenotypic variance and alleles from NAM 18, 31, and 33 contributed to the negative additive effect. On chromosome 18 at 185 cM the single QTL for NMDI and grain yield, *QNMDI.R5-chr18* and *QGY-chr18*, explained 4.73% and 5.18% of the phenotypic variance respectively. Alleles from

all NAM families contributed to the positive additive effect, primarily NAM 18 and 33 for both traits (Table 4.7).

4. Discussion

4.1 Correlations Between Traits and Yield

Relatively few studies have reported an extensive list of vegetative index relationships with soybean yield. Of the five vegetative indices mapped in this study, the Normalized Difference Vegetation Index (NDVI) is the most commonly studied. NDVI is actively used at county-level corn and soybean yield forecasting efforts in the Corn Belt region (Johnson, 2014; Kross et al., 2015). However, vegetative indices displaying the greatest correlations with grain yield in this study were VREI2 ($r = -0.42$) and PRI ($r = 0.42$) at R5. Luetchens and Lorenz, (2018) determined the photochemical reflectance index (PRI) out of several indices was most strongly correlated to relative water content in maize. Feng et al. (2015) compared 15 indices for nitrogen uptake, often an indicator of chlorophyll content, in wheat where VREI2 proved to be both powerful and robust for monitoring above ground nitrogen uptake. In addition to utilizing indices more sensitive to water content, emphasis should be directed towards using indices sensitive to foliar chlorophyll content to provide useful information on leaf photosynthetic capacity and correlation to grain yield. Vegetation indices consistently exhibited greater correlation with yield at reproductive growth stage R5 than R1. This may be due to a missing environment for R1, but several studies in wheat have confirmed NDVI is more predictive of yield at later growth stages (Sembiring et al., 2000; Moges et al., 2005; Crain et al., 2017) and Moges et al.(2005).attributed this largely due to an increased percentage of soil covered by vegetation at later stages. Although we reduced soil exposure by collecting absolute reflectance directly above the canopy, minimizing soil reflectance should be taken into consideration.

Reflectance of a canopy is strongly influenced by the soil reflectance especially when plant density is low and may limit vegetative indices to growth stages at full canopy coverage.

4.2 Yield and Vegetative Index QTL

Environmental effects were large and Broad-sense heritabilities low for vegetative indices, however, five vegetative trait and grain yield BLUPs mapped to a total of five QTL. *QNMDI.R5-chr18* and *QGY-chr18*, NMDI and grain yield respectively, mapped to an identical position on chromosome 18 at 185 cM. The Normalized Multi-band Drought Index (NMDI) was initially proposed for monitoring soil and vegetation moisture from space and enhanced the sensitivity to drought severity in comparison to the Normalized Difference Water Index (NDWI), related to the Normalized Water Index 2 (NWIB) used in this study (Wang and Qu, 2007). The additive effect for both traits was greater from NAM populations 18 and 33, derived from parent genotypes that segregate for the semideterminate growth habit. Based on the timing of the termination of apical stem growth, soybean cultivars are classified between two categories of stem architecture, determinate and indeterminate, with a continuum in between. A previous linkage analysis with 20 F₂ plants demonstrated that *Dt2* was located at the distal end of the short arm of chromosome 18 (Muehlbauer et al., 1989), co-localizing in position of the detected NMDI and grain yield QTL. Ping et al. (2014) further fine-mapped *Dt2* between markers *SSR_18_1821* and *SSR_18_1825* in a mapping population between two NAM parents included in this study, NE3001 and IA3023, and determined *Dt2* is a gain-of-function MADS-domain factor gene that causes semideterminancy. Soybean stem growth habit is a key adaptation and agronomic trait that directly affects plant height, flowering time and duration, canopy size, maturity, and markedly, water use efficiency and soybean yield (Specht et al., 2001) which may explain the co-localization of the NMDI and grain yield QTL with *Dt2*.

The Plant Senescence Reflectance Index (PSRI) is used to estimate carotenoid content, the characteristic feature accompanying natural and possibly stress-induced leaf senescence (Merzlyak et al., 1999). Dhanapal et al. (2015) identified 28 putative candidate SNPs for measured and estimated carotenoids in a soybean GWAS study, including a SNP for extractable carotenoid content on chromosome 3 at 47,434,930 base pairs. This SNP is approximately 4.6MB from the right flanking marker of *QPSRI.R5-chr3*, potentially identifying the same genetic region. NDVI and the Vogelmann Red Index are both sensitive to chlorophyll content, and Li et al. (2010) mapped QTL on chromosomes 1 and 10 for chlorophyll content measured from seedling to blooming stage in the 244 F_{2:3} and F_{2:4} families in 2 environments, published on SoyBase as leaflet chlorophyll QTL 1-4, 1-6, and 1-8 (Table 4.8). Directly comparing genetic positions is hampered by reduced genetic distances in our assembled linkage map, whereas chromosome 1 and 10 are approximately 74 and 50 cM shorter than the soybean consensus map (Table 4.6). *QNWIB.R5-chr1* co-localized with *QVREI2.R5-chr1* and Mian et al. (1996) reported a trait QTL related to the Normalized Water Index 2 (NWIB), water use efficiency (WUE), an important trait that has been associated with drought tolerance, on chromosome 1. This QTL was published in SoyBase as WUE QTL 1-1 (Table 4.8).

4.3 Potential Application of QTL in Soybean Breeding

Grain yield is highly affected by environments, and it is difficult to select high-yielding lines in smaller plots at the early stage of a breeding program. Trait-assisted selection can be an efficient strategy when correlated traits are obtained earlier or more inexpensively than a focal trait (Fernandes et al., 2017). Soybean yield was demonstrated to be significantly and positively correlated with vegetative indices at R5 and R1, and indices can be collected rapidly, non-destructively, and inexpensively than seed harvest at final plant maturity. Identifying the genetic control of correlated traits further enables a program to apply marker-assisted selection for yield

related traits in breeding programs. QTL were mapped from BLUPs estimated over three environments, suggesting these QTL and could be considered for selecting for yield across environments.

5. Conclusion

A linkage map was constructed from four RIL populations of the SoyNAM panel using the SoyNAM6K BeadChip; it was sufficient in mapping QTL for grain yield (GY) and several vegetative indices, specifically NDVI, NMDI, NWIB, PSRI, and VREI2 at seed fill (R5), with no QTL detected at beginning bloom (R1) in this study. Two pleiotropic QTL clusters for GY and NMDI, propositioned to be previously identified semideterminate growth habit gene *Dt2*, and NWIB and VREI2 were detected. Three individual QTL for NDVI (2) and PSRI (1) were identified, with an average confidence interval of 4 cM. Therefore, these QTL could serve as a reference for future QTL mapping studies for fine mapping, candidate gene discovery, and MAS in soybean breeding.

Acknowledgments

Race Higgins was supported by the Monsanto Chair in Soybean Breeding, Iowa Soybean Association (ISA) and R. F. Baker Center for Plant Breeding. We sincerely appreciate the members of Singh Soybean Group at Iowa State University (ISU), particularly Kyle Parmley and Matt Carroll for assistance with phenotyping. We are thankful to ISU, Monsanto Chair in Soybean Breeding at ISU, and the R. F. Baker Center for Plant Breeding at ISU for research experimentation financial support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Araus, J.L., and J.E. Cairns. 2014. Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19(1): 52–61. doi: 10.1016/j.tplants.2013.09.008.
- Bates, D., M. Maechler, B. Bolker, and S. Walker. 2014. lme4: Linear mixed-effects models using Eigen and S4. R Package Version 1(7): 1–23.
- Board, J.E., and B.G. Harville. 1996. Growth dynamics during the vegetative period affects yield of narrow-row, late-planted soybean. *Agron. J.* 88(4): 567–572.
- Boegh, E., H. Soegaard, N. Broge, C.B. Hasager, N.O. Jensen, K. Schelde, and A. Thomsen. 2002. Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. *Remote Sens. Environ.* 81(2): 179–193.
- Bolton, D.K., and M.A. Friedl. 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* 173: 74–84.
- Carter, G.A. 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Remote Sens.* 15(3): 697–703.
- Chappelle, E.W., M.S. Kim, and J.E. McMurtrey. 1992. Ratio analysis of reflectance spectra (RARS): an algorithm for the remote estimation of the concentrations of chlorophyll a, chlorophyll b, and carotenoids in soybean leaves. *Remote Sens. Environ.* 39(3): 239–247.
- Christenson, B.S., W.T. Schapaugh, N. An, K.P. Price, and A.K. Fritz. 2014. Characterizing Changes in Soybean Spectral Response Curves with Breeding Advancements. *Crop Sci.* 54(4): 1585. doi: 10.2135/cropsci2013.08.0575.
- Crain, J., M. Reynolds, and J. Poland. 2017. Utilizing High-Throughput Phenotypic Data for Improved Phenotypic Selection of Stress-Adaptive Traits in Wheat. *Crop Sci.* 57(2): 648–659. doi: 10.2135/cropsci2016.02.0135.
- Dhanapal, A.P., J.D. Ray, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, C.A. King, and F.B. Fritschi. 2015. Association Mapping of Total Carotenoids in Diverse Soybean Genotypes Based on Leaf Extracts and High-Throughput Canopy Spectral Reflectance Measurements (DA Lightfoot, Ed.). *PLOS ONE* 10(9): e0137213. doi: 10.1371/journal.pone.0137213.
- Fehr, W.R., C.E. Caviness, D.T. Burmood, and J.S. Pennington. 1971. Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. *Crop Sci.* 11(6): 929–931.
- Feng, W., B.-B. Guo, H.-Y. Zhang, L. He, Y.-S. Zhang, Y.-H. Wang, Y.-J. Zhu, and T.-C. Guo. 2015. Remote estimation of above ground nitrogen uptake during vegetative growth in winter wheat using hyperspectral red-edge ratio data. *Field Crops Res.* 180: 197–206. doi: 10.1016/j.fcr.2015.05.020.

- Fernandes, S.B., K.O.G. Dias, D.F. Ferreira, and P.J. Brown. 2017. Efficiency of multi-trait, indirect, and trait-assisted genomic selection for improvement of biomass sorghum. *Theor. Appl. Genet.*: 1–9. doi: 10.1007/s00122-017-3033-y.
- Gitelson, A.A., Y. Gritz, and M.N. Merzlyak. 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* 160(3): 271–282.
- Gitelson, A.A., Y.J. Kaufman, and M.N. Merzlyak. 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58(3): 289–298.
- Grant, D., R.T. Nelson, S.B. Cannon, and R.C. Shoemaker. 2009. SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* 38(suppl_1): D843–D846.
- Gray, S.B., O. Dermody, S.P. Klein, A.M. Locke, J.M. Mcgrath, R.E. Paul, D.M. Rosenthal, U.M. Ruiz-Vera, M.H. Siebers, and R. Strellner. 2016. Intensifying drought eliminates the expected benefits of elevated carbon dioxide for soybean. *Nat. Plants* 2(9): 16132.
- Haboudane, D., J.R. Miller, E. Pattey, P.J. Zarco-Tejada, and I.B. Strachan. 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* 90(3): 337–352. doi: 10.1016/j.rse.2003.12.013.
- Heinrich, V., G. Krauss, C. Gotze, and C. Sandow. 2011. IDB - Index DataBase. Index Database IDB Proj. <https://www.indexdatabase.de/> (accessed 21 March 2015).
- Huang, Y., L. Yuan, K.N. Reddy, and J. Zhang. 2016. In-situ plant hyperspectral sensing for early detection of soybean injury from dicamba. *Biosyst. Eng.* 149: 51–59. doi: 10.1016/j.biosystemseng.2016.06.013.
- Johnson, D.M. 2014. An assessment of pre-and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sens. Environ.* 141: 116–128.
- Keep, N.R., W.T. Schapaugh, P.V.V. Prasad, and J.E. Boyer. 2016. Changes in Physiological Traits in Soybean with Breeding Advancements. *Crop Sci.* 56(1): 122. doi: 10.2135/cropsci2013.07.0499.
- Koester, R.P., B.M. Nohl, B.W. Diers, and E.A. Ainsworth. 2016. Has photosynthetic capacity increased with 80 years of soybean breeding? An examination of historical soybean cultivars. *Plant Cell Environ.*: n/a-n/a. doi: 10.1111/pce.12675.
- Kross, A., H. McNairn, D. Lapen, M. Sunohara, and C. Champagne. 2015. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *Int. J. Appl. Earth Obs. Geoinformation* 34: 235–248. doi: 10.1016/j.jag.2014.08.002.

- Li, G.-J., H.-N. Li, L.-G. Cheng, and Y.-M. Zhang. 2010. QTL Analysis for Dynamic Expression of Chlorophyll Content in Soybean (*Glycine max* L. Merr.). *Acta Agron. Sin.* 36(2): 242–248. doi: 10.1016/S1875-2780(09)60033-X.
- Lipper, L., P. Thornton, B.M. Campbell, T. Baedeker, A. Braimoh, M. Bwalya, P. Caron, A. Cattaneo, D. Garrity, K. Henry, R. Hottle, L. Jackson, A. Jarvis, F. Kossam, W. Mann, N. McCarthy, A. Meybeck, H. Neufeldt, T. Remington, P.T. Sen, R. Sessa, R. Shula, A. Tibu, and E.F. Torquebiau. 2014. Climate-smart agriculture for food security. *Nat. Clim. Change* 4(12): 1068–1072. doi: 10.1038/nclimate2437.
- Luetchens, J., and A.J. Lorenz. 2018. Changes in Dynamic Leaf Traits in Maize Associated with Year of Hybrid Release. *Crop Sci.* 58(2): 551–563. doi: 10.2135/cropsci2017.04.0256.
- Ma, B.L., L.M. Dwyer, C. Costa, E.R. Cober, and M.J. Morrison. 2001. Early prediction of soybean yield from canopy reflectance measurements. *Agron. J.* 93(6): 1227–1234.
- Ma, B.L., M.J. Morrison, and H.D. Voldeng. 1995. Leaf greenness and photosynthetic rates in soybean. *Crop Sci.* 35(5): 1411–1414.
- Melillo, J.M., J.D. Aber, and J.F. Muratore. 1982. Nitrogen and Lignin Control of Hardwood Leaf Litter Decomposition Dynamics. *Ecology* 63(3): 621–626. doi: 10.2307/1936780.
- Meng, L., H. Li, L. Zhang, and J. Wang. 2015. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* 3(3): 269–283. doi: 10.1016/j.cj.2015.01.001.
- Merzlyak, M.N., A.A. Gitelson, O.B. Chivkunova, and V.Y. Rakitin. 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.* 106(1): 135–141.
- Mian, M.A.R., M.A. Bailey, D.A. Ashley, R. Wells, T.E. Carter, W.A. Parrott, and H.R. Boerma. 1996. Molecular Markers Associated with Water Use Efficiency and Leaf Ash in Soybean. *Crop Sci.* 36(5): 1252. doi: 10.2135/cropsci1996.0011183X003600050030x.
- Moges, S.M., W.R. Raun, R.W. Mullen, K.W. Freeman, G.V. Johnson, and J.B. Solie. 2005. Evaluation of green, red, and near infrared bands for predicting winter wheat biomass, nitrogen uptake, and final grain yield. *J. Plant Nutr.* 27(8): 1431–1441.
- Monteith, J.L., and C.J. Moss. 1977. Climate and the efficiency of crop production in Britain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 281(980): 277–294.
- Muehlbauer, G.J., J.E. Specht, P.E. Staswick, G.L. Graef, and M.A. Thomas-Compton. 1989. Application of the near-isogenic line gene mapping technique to isozyme markers. *Crop Sci.* 29(6): 1548–1553.

- Nagler, P.L., Y. Inoue, E. Glenn, A. Russ, and C.S. Daughtry. 2003. Cellulose absorption index (CAI) to quantify mixed soil–plant litter scenes. *Remote Sens. Environ.* 87(2–3): 310–325. doi: 10.1016/j.rse.2003.06.001.
- Nutter Jr, F.W., G.L. Tylka, J. Guan, A.J.D. Moreira, C.C. Marett, T.R. Rosburg, J.P. Basart, and C.S. Chong. 2002. Use of remote sensing to detect soybean cyst nematode-induced plant stress. *J. Nematol.* 34(3): 222.
- Nyquist, W.E., and R.J. Baker. 1991. Estimation of heritability and prediction of selection response in plant populations. *Crit. Rev. Plant Sci.* 10(3): 235–322. doi: 10.1080/07352689109382313.
- Peñuelas, J., I. Filella, C. Biel, L. Serrano, and R. Save. 1993. The reflectance at the 950–970 nm region as an indicator of plant water status. *Int. J. Remote Sens.* 14(10): 1887–1905.
- Peñuelas, J., I. Filella, and J.A. Gamon. 1995. Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytol.* 131(3): 291–296. doi: 10.1111/j.1469-8137.1995.tb03064.x.
- Ping, J., Y. Liu, L. Sun, M. Zhao, Y. Li, M. She, Y. Sui, F. Lin, X. Liu, Z. Tang, H. Nguyen, Z. Tian, L. Qiu, R.L. Nelson, T.E. Clemente, J.E. Specht, and J. Ma. 2014. Dt2 Is a Gain-of-Function MADS-Domain Factor Gene That Specifies Semideterminacy in Soybean. *Plant Cell* 26(7): 2831–2842. doi: 10.1105/tpc.114.126938.
- Rafalski, J.A. 2010. Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13(2): 174–180.
- Rincker, K., R. Nelson, J. Specht, D. Sleper, T. Cary, S.R. Cianzio, S. Casteel, S. Conley, P. Chen, V. Davis, C. Fox, G. Graef, C. Godsey, D. Holshouser, G.-L. Jiang, S.K. Kantartzi, W. Kenworthy, C. Lee, R. Mian, L. McHale, S. Naeve, J. Orf, V. Poysa, W. Schapaugh, G. Shannon, R. Uniatowski, D. Wang, and B. Diers. 2014. Genetic Improvement of U.S. Soybean in Maturity Groups II, III, and IV. *Crop Sci.* 0(0): 0. doi: 10.2135/cropsci2013.10.0665.
- Romero, A., I. Aguado, and M. Yebra. 2012. Estimation of dry matter content in leaves using normalized indexes and PROSPECT model inversion. *Int. J. Remote Sens.* 33(2): 396–414.
- Rouse Jr, J., R.H. Haas, J.A. Schell, and D.W. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTS.
- Saryoko, A., Y. Fukuda, I. Lubis, K. Homma, and T. Shiraiwa. 2018. Physiological activity and biomass production in crop canopy under a tropical environment in soybean cultivars with temperate and tropical origins. *Field Crops Res.* 216: 209–216. doi: 10.1016/j.fcr.2017.11.012.

- Sembiroing, H., H.L. Lees, W.R. Raun, G.V. Johnson, J.B. Solie, M.L. Stone, M.J. DeLeon, E.V. Lukina, D.A. Cossey, J.M. LaRuffa, C.W. Woolfolk, S.B. Phillips, and W.E. Thomason. 2000. Effect of growth stage and variety on spectral radiance in winter wheat. *J. Plant Nutr.* 23(1): 141–149. doi: 10.1080/01904160009382003.
- Slattery, R.A., A. VanLoocke, C.J. Bernacchi, X.-G. Zhu, and D.R. Ort. 2017. Photosynthesis, Light Use Efficiency, and Yield of Reduced-Chlorophyll Soybean Mutants in Field Conditions. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00549.
- Specht, J.E., K. Chase, M. Macrander, G.L. Graef, J. Chung, J.P. Markwell, M. Germann, J.H. Orf, and K.G. Lark. 2001. Soybean response to water. *Crop Sci.* 41(2): 493–509.
- Tanksley, S.D., and S.R. McCouch. 1997. Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. *Science* 277(5329): 1063–1066. doi: 10.1126/science.277.5329.1063.
- Technow, F. 2011. R package mvngGrAd: Moving grid adjustment in plant breeding field trials. (0.1.5).
- Thompson, J.A., and R.L. Nelson. 1998. Utilization of diverse germplasm for soybean yield improvement. *Crop Sci.* 38(5): 1362–1368.
- USDA-NASS. 2016. Crop production 2016 Summary. USDA-NASS, Washington, DC.
- Vogelmann, J.E., B.N. Rock, and D.M. Moss. 1993. Red edge spectral measurements from sugar maple leaves. *Int J Remote Sens* 14(8): 1563–1575.
- Wang, L., and J.J. Qu. 2007. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophys. Res. Lett.* 34(20).
- Xavier, A., W.D. Beavis, J.E. Specht, B. Diers, W.M. Muir, and K.M. Rainey. 2015. SoyNAM: Soybean Nested Association Mapping Dataset. R Package Version 1.
- Xavier, A., B. Hall, A.A. Hearst, K.A. Cherkauer, and K.M. Rainey. 2017. Genetic Architecture of Phenomic-Enabled Canopy Coverage in *Glycine max*. *Genetics* 206(2): 1081–1089. doi: 10.1534/genetics.116.198713.
- Xavier, A., D. Jarquin, R. Howard, V. Ramasubramanian, J.E. Specht, G.L. Graef, W.D. Beavis, B.W. Diers, Q. Song, P.B. Cregan, R. Nelson, R. Mian, J.G. Shannon, L. McHale, D. Wang, W. Schapaugh, A.J. Lorenz, S. Xu, W.M. Muir, and K.M. Rainey. 2018. Genome-Wide Analysis of Grain Yield Stability and Environmental Interactions in a Multiparental Soybean Population. *G3 Genes Genomes Genet.* 8(2): 519–529. doi: 10.1534/g3.117.300300.

Xavier, A., W.M. Muir, and K.M. Rainey. 2016. Assessing Predictive Properties of Genome-Wide Selection in Soybeans. *G3 Genes Genomes Genet.* 6(8): 2611–2616. doi: 10.1534/g3.116.032268.

Yu, J., and E.S. Buckler. 2006. Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* 17(2): 155–160.

Zygielbaum, A.I., T.J. Arkebauer, E.A. Walter-Shea, and D.L. Scoby. 2012. Detection and measurement of vegetation photoprotection stress response using PAR reflectance. *Isr. J. Plant Sci.* 60(1–2): 37–47.

Tables and figures

Table 4.1 Summary of parental genotype information including name, origin, NAM number, ancestry, growth habit, and maturity grouping followed by Recombinant Inbred Line (RIL) population size and lines with quality assured genotypic information

NAM parent	Origin	NAM population	Ancestry	Growth habit	Maturity	Quality assured	
						RIL genotypes	RIL genotypes
HS6-3976	Ohio State	NAM 08	Elite	Indeterminate	3	140	138
NE3001	Univ. of Nebraska	NAM 18	Elite	Semi-determinate	3	140	136
LG90-2550	USDA-ARS	NAM 31	Diverse	Semi-determinate	3	140	127
LG94-1128	USDA-ARS	NAM 33	Diverse	Indeterminate	2	140	134
IA3023	Iowa State Univ.	Universal Parent	Elite	Indeterminate	3	-	-

Table 4.2 Summary of selected vegetation indices with abbreviation, general class type, equation, and source.

Vegetation Index Description	Abbreviation	Type	Equation ^a	Source
Photochemical Reflectance Index	PRI	Carotenoids	$(\rho_{531} - \rho_{570}) / (\rho_{531} + \rho_{570})$	Peñuelas et al., 1995
Plant Senescence Reflectance Index	PSRI	Carotenoids	$(\rho_{680} - \rho_{500}) / \rho_{750}$	Merzlyak, et al., 1999
Ratio Analysis of Reflectance Spectra Chlorophyll A	RARSa	Chlorophyll	$(\rho_{675} / \rho_{700})$	Chappelle et al., 1992
Ratio Analysis of Reflectance Spectra Chlorophyll B	RARSb	Chlorophyll	$(\rho_{675} / (\rho_{650} \times \rho_{700}))$	Chappelle et al., 1992
Vogelmann Red Edge Index 2	VREI2	Chlorophyll	$(\rho_{734} - \rho_{747}) / (\rho_{715} + \rho_{726})$	Vogelmann, et al. 1993
Cellulose Absorption Index	CAI	Dry matter	$0.5 \times (\rho_{2000} - \rho_{2200}) - \rho_{2100}$	Nagler et al., 2003
Dry Matter Content Index	DMCI	Dry matter	$(\rho_{2305} - \rho_{1495}) / (\rho_{2305} + \rho_{1495})$	Romero et al., 2012
Normalized Difference Lignin Index	NDLI	Dry matter	$[\log(1/\rho_{1754}) - \log(1/\rho_{1680})] / [\log(1/\rho_{1754}) + \log(1/\rho_{1680})]$	Melillo et al., 1982
Green Average	Green	Vegetation	$\text{Avg}(\rho_{505} \text{ to } \rho_{595})$	Gitelson et al., 1996
Leaf Area Index	LAI	Vegetation	$3.618 * \text{EVI} - 0.118$	Boegh et al., 2002
Near Infrared	RedEdge	Vegetation	$\text{Avg}(\rho_{800} \text{ to } \rho_{850})$	Christenson et al., 2014
Normalized Difference Vegetation Index	NDVI	Vegetation	$(\rho_{780} - \rho_{670}) / (\rho_{780} + \rho_{670})$	Rouse Jr et al., 1973
Normalized Multi-Band Drought Index	NMDI	Water content	$(\rho_{860} - (\rho_{1640} - \rho_{2130})) / (\rho_{860} + (\rho_{1640} + \rho_{2130}))$	Wang and Qu, 2007
Normalized Water Index 2	NWIB	Water content	$(\rho_{970} - \rho_{850}) / (\rho_{970} + \rho_{850})$	Peñuelas et al., 1993

^a ρ is reflectance and the subscript is wavelength (nm).

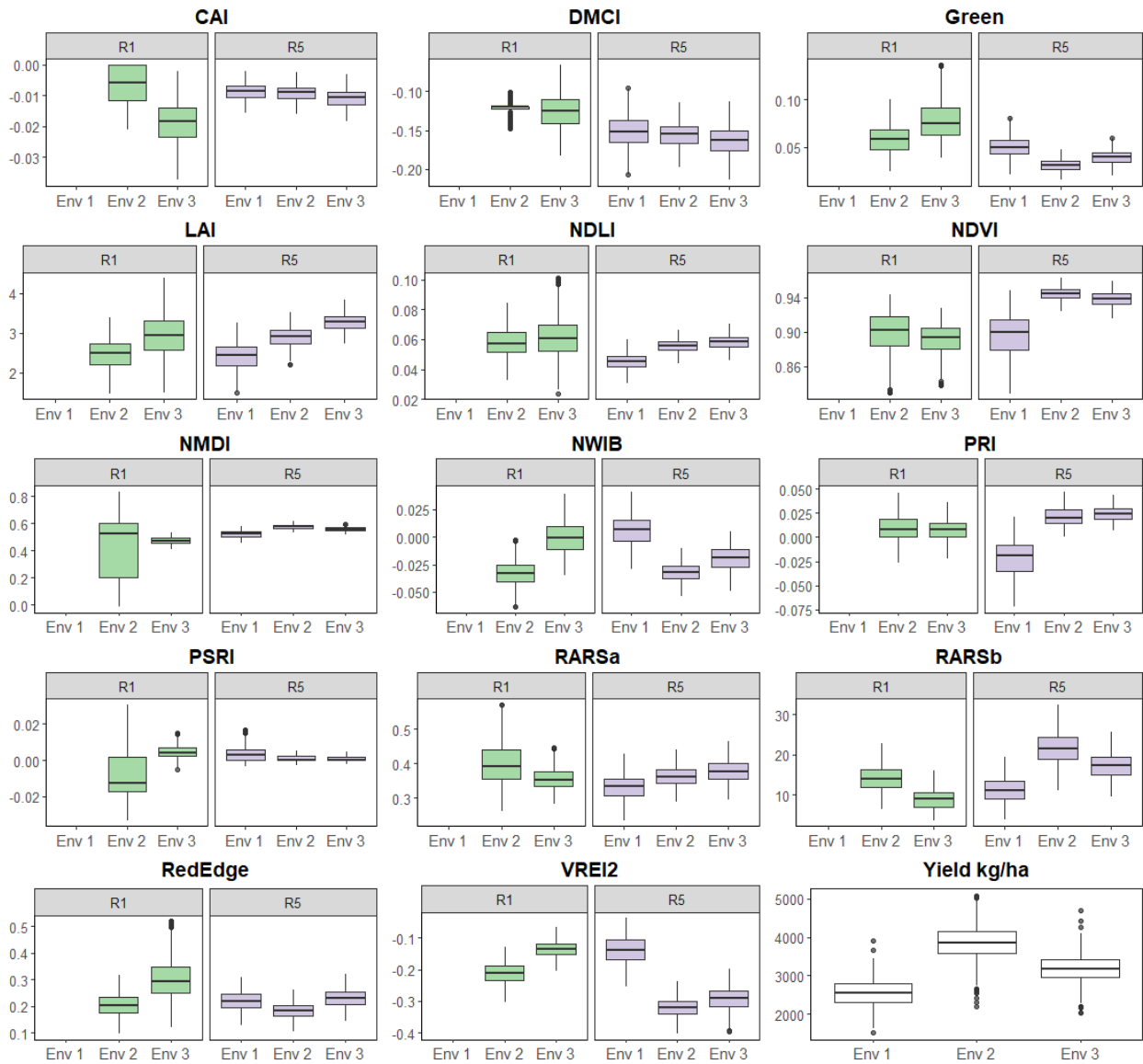


Figure 4.1 Boxplot distribution of vegetation indices and yield by environments ($n=2-3$) and soybean growth stages beginning bloom (R1), seed fill (R5), and final maturity for grain yield (R8). Strong environmental effects were observed for all traits

Table 4.3 Summary of vegetation indices and yield including count, mean, and standard deviation by growth stage post quantile outlier removal.

Trait	R1			R5		
	N	Mean	Std Dev	N	Mean	Std Dev
CAI	979	-0.0132	0.0092	1746	-0.0096	0.0029
DMCI	957	-0.1237	0.0183	1724	-0.1569	0.0193
Green	1157	0.0691	0.0213	1716	0.0408	0.0118
LAI	1170	2.7121	0.5674	1741	2.8544	0.4527
NDLI	1070	0.0595	0.0121	1744	0.0531	0.0076
NDVI	1057	0.8946	0.0226	1696	0.9266	0.0269
NMDI	1134	0.4573	0.1597	1736	0.5519	0.0285
NWIB	1096	-0.0159	0.0212	1755	-0.0149	0.0197
PRI	1156	0.0082	0.012	1741	0.0082	0.0249
PSRI	1065	-0.0026	0.0112	1663	0.0018	0.003
RARSa	1091	0.3818	0.0567	1744	0.3573	0.0369
RARSb	1169	11.552	3.8664	1704	16.7352	5.5616
RedEdge	1126	0.2545	0.0817	1731	0.2124	0.0393
VREI2	1176	-0.1735	0.0494	1737	-0.2489	0.0899
Yield (kg/ha)	1761	3218.438	735.2003	-	-	-

Table 4.4 Analysis of Variance and broad-sense heritabilities for vegetation indices and yield by growth stage. “NS” P > 0.05; “*” P ≤ 0.05; “**” P ≤ 0.01; “***” P ≤ 0.001.

Growth			CAI	DMCI	Green	LAI	NDLI	NDVI	NMDI	NWIB	PRI	PSRI	RARSa	RARSb	RedEdge	VREI2	Yield ^a
Stage	Factor	df															
R1	Env	1	***	**	***	***	***	***	***	***	*	***	***	***	***	***	***
	Check*Block(Env)	7	*	NS	***	*	NS	NS	NS	*	NS	*	NS	**	**	*	NS
	Genotype	564	***	*	**	NS	*	**	NS	***	*	**	*	NS	NS	*	***
	Genotype*Env	557	*	NS	***	NS	NS	NS	NS	*	NS	NS	NS	NS	NS	NS	NS
	Residuals	46															
	H^{2b}		0.07	0.12	0.14	-	0.07	0.23	-	0.16	0.18	0.07	0.10	-	-	0.22	0.45
R5	Env	2	***	***	***	***	***	***	***	***	***	***	***	***	***	***	-
	Check*Block(Env)	10	*	NS	NS	**	*	NS	NS	*	NS	NS	*	**	*	NS	-
	Genotype	564	*	NS	NS	***	*	**	**	***	*	**	**	***	NS	***	-
	Genotype*Env	1092	NS	NS	NS	*	NS	NS	NS	**	NS	NS	*	***	NS	*	-
	Residuals	68															-
	H^2		0.04	-	-	0.17	0.13	0.21	0.23	0.30	0.21	0.25	0.06	0.05	-	0.25	-

^aYield measured at physiological maturity (R8).

^b H^2 , broad-sense heritability, calculated on an entry mean basis.

Table 4.5 Pearson's correlations and p-values of vegetation index and yield BLUPs by growth stage, bottom left and upper right respectively.

		R1										R5										Yield	
		CAI	DMCI	Green	NDLI	NDVI	NWIB	PRI	PSRI	RARSa	VREI2	CAI	LAI	NDLI	NDVI	NMDI	NWIB	PRI	PSRI	RARSa	RARSb	VREI2	Yield
R1	CAI	-	0.00	0.59	0.81	0.52	0.81	0.00	0.00	0.67	0.59	0.36	0.28	0.08	0.17	0.63	0.77	0.75	0.97	0.02	0.72	0.32	0.27
	DMCI	0.42	-	0.68	0.87	0.05	0.87	0.11	0.36	0.78	0.10	0.63	0.11	0.03	0.03	0.20	0.09	0.04	0.92	0.59	0.30	0.01	0.11
	Green	0.02	0.02	-	0.37	0.18	0.12	0.00	0.00	0.33	0.00	0.14	0.67	0.14	0.13	0.06	0.00	0.07	0.47	0.11	0.04	0.00	0.25
	NDLI	-0.01	-0.01	0.04	-	0.91	0.97	0.48	0.73	0.94	0.15	0.47	0.85	0.41	0.35	0.70	0.89	0.78	0.13	0.29	0.35	0.65	0.95
	NDVI	-0.03	0.08	-0.06	0.00	-	0.91	0.00	0.02	0.00	0.00	0.15	0.64	0.68	0.16	0.87	0.53	0.46	0.17	0.09	0.18	0.04	0.46
	NWIB	-0.01	-0.01	-0.07	0.00	-0.01	-	0.60	0.87	0.94	0.78	0.57	0.83	0.81	0.31	0.72	0.97	0.18	0.76	0.98	0.13	0.38	0.81
	PRI	0.18	0.07	-0.39	0.03	0.25	-0.02	-	0.00	0.00	0.00	0.19	0.71	0.30	0.09	0.21	0.01	0.00	0.01	0.02	0.00	0.00	0.01
	PSRI	0.17	0.04	-0.22	0.02	-0.11	0.01	-0.12	-	0.20	0.00	0.26	0.49	0.16	0.57	0.29	0.37	0.33	0.23	0.02	0.59	0.90	0.82
	RARSa	-0.02	-0.01	-0.04	0.00	0.65	0.00	0.15	-0.06	-	0.18	0.16	0.88	0.51	0.66	0.88	0.62	0.38	0.82	0.91	0.52	0.55	0.65
	VREI2	-0.02	-0.07	0.49	-0.06	-0.17	-0.01	-0.61	0.22	-0.06	-	0.25	0.40	0.30	0.01	0.10	0.05	0.00	0.03	0.00	0.00	0.00	0.00
R5	CAI	-0.04	0.02	-0.06	-0.03	0.06	0.02	0.06	-0.05	0.06	-0.05	-	0.00	0.00	0.00	0.00	0.00	0.28	0.22	0.15	0.00	0.00	0.62
	LAI	0.05	0.07	-0.02	0.01	0.02	-0.01	0.02	0.03	-0.01	-0.04	-0.51	-	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NDLI	0.07	0.09	-0.06	0.04	0.02	-0.01	0.04	0.06	0.03	-0.05	-0.37	0.74	-	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NDVI	0.06	0.09	-0.07	0.04	0.06	-0.04	0.07	0.02	0.02	-0.11	0.18	0.33	0.24	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NMDI	0.02	0.06	-0.08	-0.02	-0.01	-0.02	0.05	-0.05	-0.01	-0.07	0.24	0.07	-0.06	0.28	-	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	NWIB	0.01	-0.07	0.14	-0.01	-0.03	0.00	-0.11	-0.04	-0.02	0.08	-0.35	-0.19	-0.15	-0.42	-0.54	-	0.00	0.00	0.00	0.00	0.00	0.00
	PRI	0.01	0.09	-0.08	0.01	0.03	-0.06	0.13	-0.04	0.04	-0.14	0.05	0.43	0.23	0.64	0.23	-0.48	-	0.00	0.00	0.00	0.00	0.00
	PSRI	0.00	0.00	0.03	-0.07	-0.06	0.01	-0.12	0.05	-0.01	0.09	-0.05	-0.36	-0.28	-0.56	-0.16	0.42	-0.51	-	0.00	0.00	0.00	0.00
	RARSa	-0.10	-0.02	-0.07	-0.05	0.07	0.00	0.10	-0.10	0.00	-0.13	0.06	-0.19	-0.13	-0.31	-0.12	0.15	-0.13	0.13	-	0.01	0.34	0.16
	RARSb	0.02	0.04	-0.09	-0.04	0.06	-0.07	0.13	-0.02	0.03	-0.16	0.44	-0.19	-0.17	0.49	0.19	-0.29	0.47	-0.20	0.12	-	0.00	0.00
VREI2	-0.04	-0.12	0.15	0.02	-0.09	0.04	-0.15	-0.01	-0.03	0.22	-0.34	-0.22	-0.14	-0.68	-0.39	0.53	-0.65	0.40	-0.04	-0.72	-	0.00	
Yield	0.05	0.07	-0.05	0.00	0.03	-0.01	0.11	-0.01	0.02	-0.15	0.02	0.21	0.14	0.34	0.22	-0.33	0.42	-0.31	-0.06	0.25	-0.42	-	

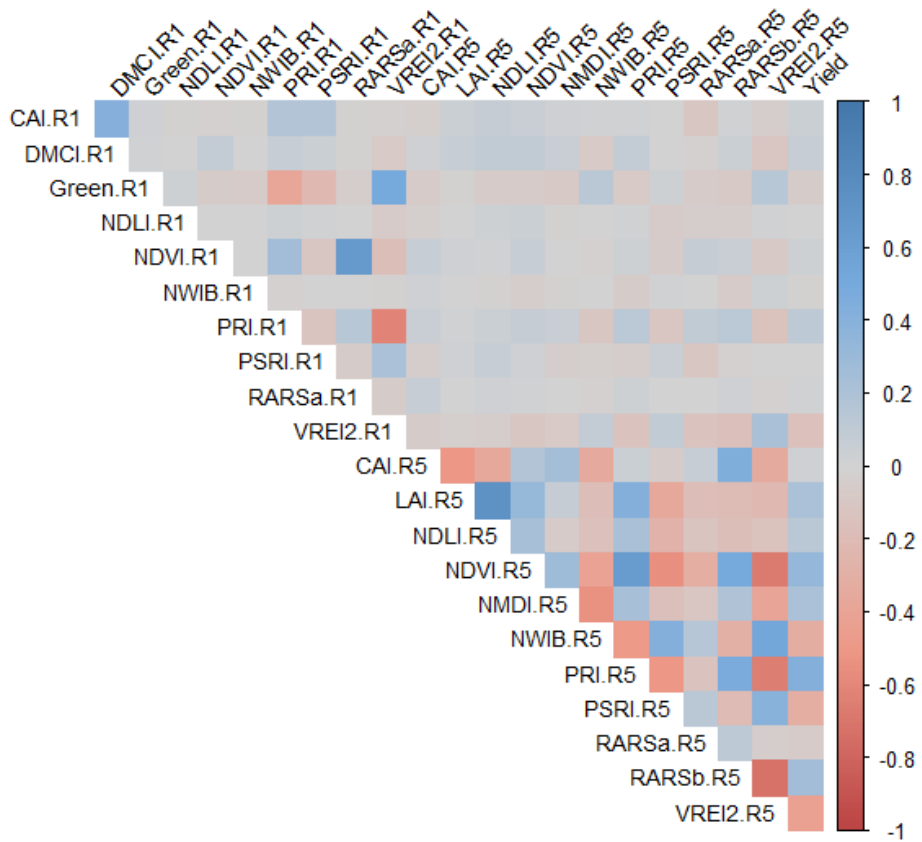


Figure 4.2 Heatmap of Pearson's correlations between vegetation index and yield BLUPs by growth stage. Greater correlations were observed within growth stage and within R5.

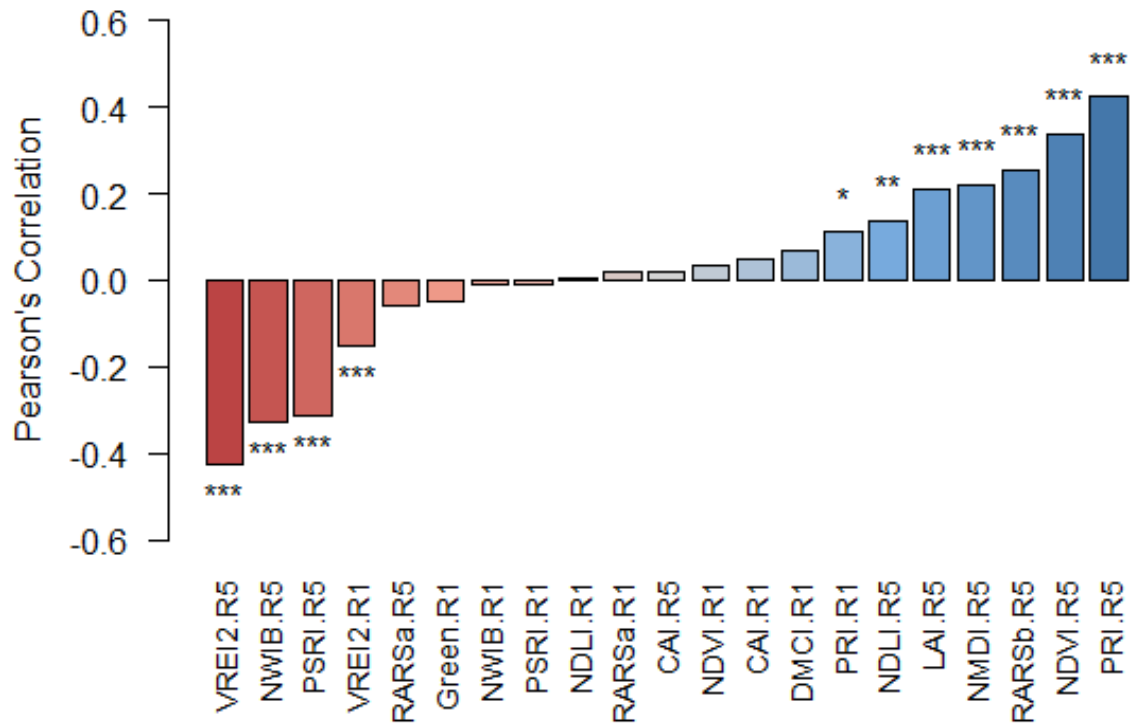


Figure 4.3 Bar chart of Pearson's correlations between vegetative index BLUPs and grain yield BLUPs in ascending order. Significant correlations are noted by "*" $P \leq 0.05$; "***" $P \leq 0.01$; "****" $P \leq 0.001$.

Table 4.6 Summary of soybean chromosomes, linkage group ID, marker number, observed linkage length, consensus linkage map for reference, and SNP marker density.

Chromosome Number	Linkage Group	Marker Number	Linkage length (cM)	Soybean Consensus Map 3.0 (cM)	SNP density (cM)
1	D1a	28	24.02	98.41	0.86
2	D1b	69	80.11	140.63	1.16
3	N	37	95.18	99.51	2.57
4	C1	14	138.80	112.32	9.91
5	A1	18	73.30	86.75	4.07
6	C2	19	52.56	136.51	2.77
7	M	12	92.10	135.15	7.68
8	A2	58	112.35	146.67	1.94
9	K	11	69.55	99.60	6.32
10	O	13	80.32	132.89	6.18
11	B1	52	110.59	124.24	2.13
12	H	24	49.38	120.50	2.06
13	F	53	112.44	120.03	2.12
14	B2	12	78.68	108.18	6.56
15	E	25	87.04	99.88	3.48
16	J	29	128.99	92.27	4.45
17	D2	15	16.65	119.19	1.11
18	G	17	185.89	105.00	10.93
19	L	13	63.65	101.14	4.9
20	I	42	98.51	112.77	2.35
sum		561	1750.11	2291.64	avg = 3.12

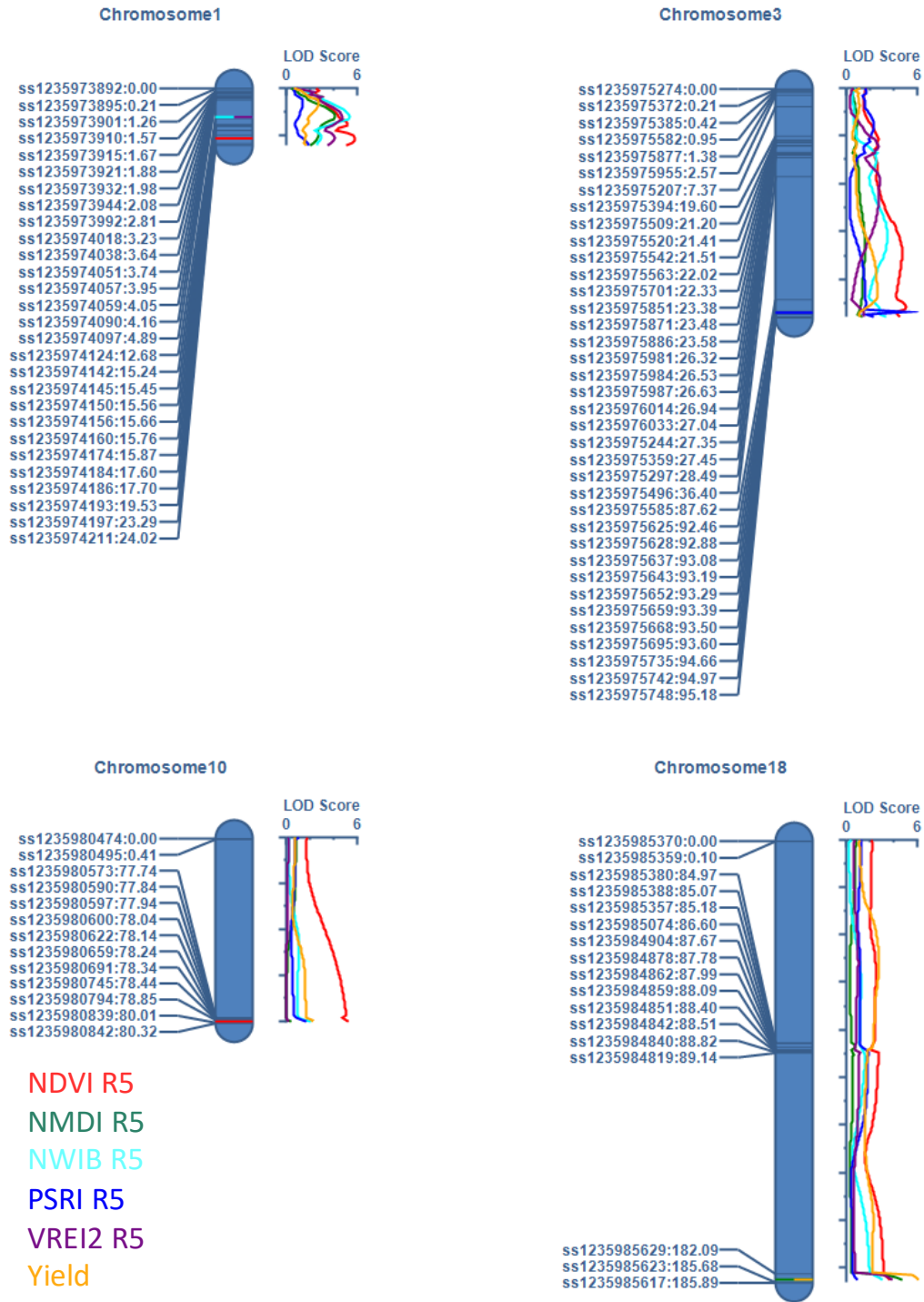


Figure 4.4 Genetic linkage map of chromosomes 1, 3, 10, and 18 with QTL for vegetation indices and yield , including marker name, genetic distance, and LOD score. Traits, QTL position, and LOD scores are marked by color.

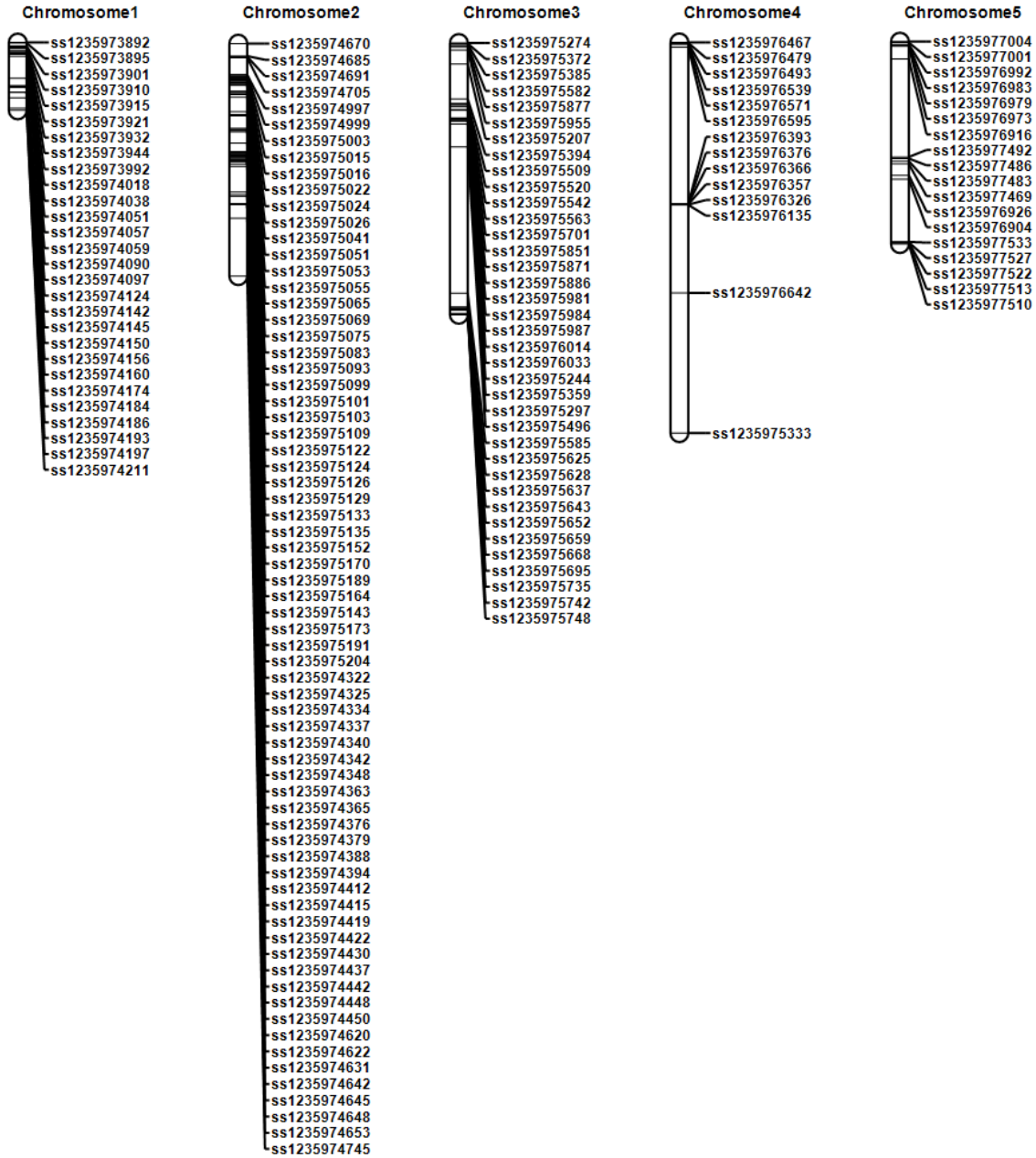


Figure 4.5 Genetic linkage map of soybean chromosomes 1-20 with marker names. The total length was 1750.11 cM consisting of 561 markers.



Figure 4.5 continued

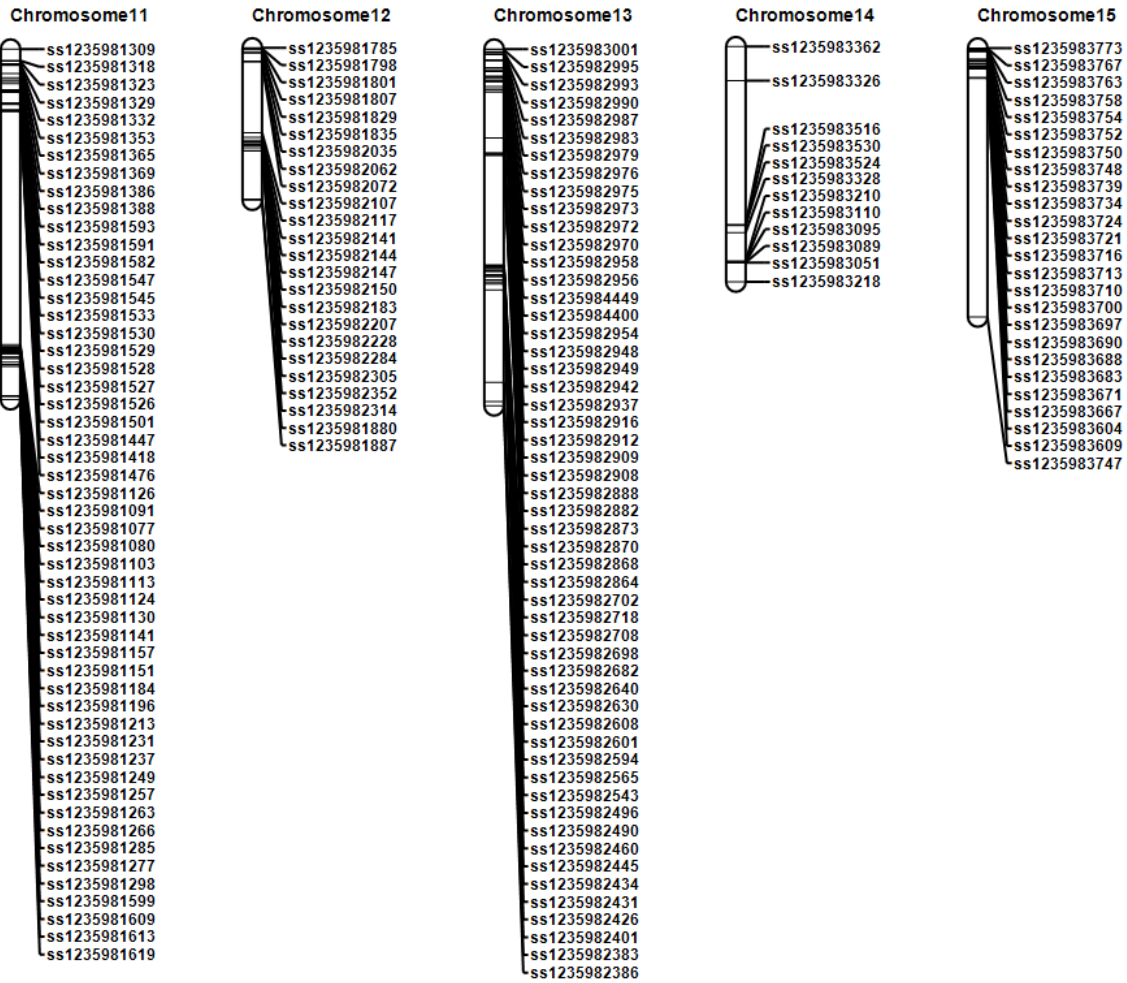


Figure 4.5 continued

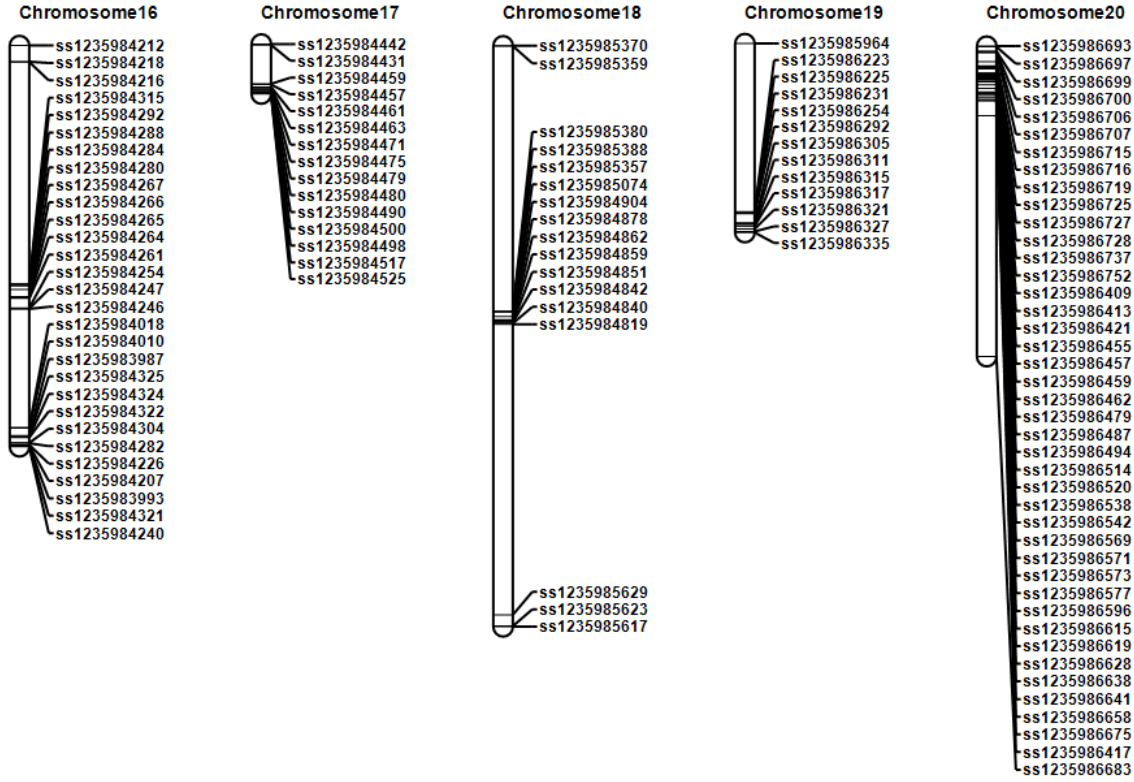


Figure 4.5 continued

Table 4.7 Summary of QTL identified for five vegetation indices and yield by growth stage, including genetic position, confidence interval, left and right flanking marker, LOD score, LOD score by family, and family additive effects. A total of 5 distinct QTL were identified in a four-RIL nested association mapping population on chromosomes 1, 3, 10, and 18.

Trait Name	Growth Stage	Chr. Number	Position (cM)	Left CI	Right CI	Left Marker	Right Marker	Cumulative LOD	Cumulative PVE (%)	Family LOD Score				Family Additive Effect			
										NAM 08	NAM 31	NAM 33	NAM 18	NAM 08	NAM 31	NAM 33	NAM 18
NDVI	R5	1	21	18.5	24.0	ss1235974193	ss1235974197	5.81	3.08	1.39	3.25	0.37	0.80	-0.0003	0.0006	0.0002	0.0002
NDVI	R5	10	80	78.5	80.0	ss1235980794	ss1235980839	5.24	2.89	0.99	0.31	3.45	0.49	0.0001	0.0001	0.0007	-0.0001
NMDI	R5	18	185	183.5	185.0	ss1235985629	ss1235985623	4.73	5.17	0.20	3.00	0.18	1.35	0.0002	0.0008	0.0002	0.0006
NWIB	R5	1	12	7.5	14.5	ss1235974097	ss1235974124	5.37	5.02	1.04	2.76	0.78	0.80	0.0003	-0.0005	0.0003	-0.0003
PSRI	R5	3	93	92.5	93.5	ss1235975628	ss1235975637	5.04	5.29	0.18	2.70	0.34	1.81	0	-0.0001	0	-0.0001
VREI2	R5	1	12	7.5	16.5	ss1235974097	ss1235974124	4.74	4.78	1.25	3.01	0.01	0.47	0.0012	-0.0017	-0.0001	-0.0007
Yield	R8	18	185	182.5	185.0	ss1235985629	ss1235985623	6.05	5.18	0.44	1.97	0.01	3.63	0.1800	0.4526	0.0264	0.6686

Table 4.8 List of 41 QTL reported for traits related to vegetation indices on chromosomes 1, 3, 10, and 18 from Soybase (www.soybase.org) with beginning and end genetic positions.

Trait	QTL Number	Linkage Group	Chr. Number	Start (cM)	End (cM)
Drought index	1-1	D1a	1	62.56	65.62
Drought tolerance	6-1	N	3	70.43	81.39
Drought tolerance	6-3	O	10	60.00	88.90
Leaflet area	1-6	O	10	7.30	9.30
Leaflet area	9-5	O	10	39.50	41.50
Leaflet area	2-3	O	10	58.43	60.43
Leaflet area	2-2	G	18	52.90	59.30
Leaflet area	1-4	G	18	78.00	80.00
Leaflet chlorophyll	1-6	D1a	1	62.36	77.48
Leaflet chlorophyll	1-4	D1a	1	104.27	108.88
Leaflet chlorophyll	1-14	N	3	74.80	76.48
Leaflet chlorophyll	1-3	N	3	84.54	102.05
Leaflet chlorophyll	2-1	O	10	9.53	129.30
Leaflet chlorophyll	1-8	O	10	56.93	59.43
Leaflet chlorophyll	1-2	G	18	4.53	12.54
Plant weight, dry	1-3	G	18	60.36	66.55
Seed yield	30-4	N	3	90.55	94.55
Seed yield	30-3	G	18	94.50	105.50
Seed yield	8-2	D1a	1	63.52	65.52
Seed yield	15-12	N	3	52.25	54.25
Seed yield	15-13	N	3	74.91	74.91
Seed yield	27-4	N	3	90.55	94.55
Seed yield	4-1	N	3	135.00	137.00
Seed yield	22-19	O	10	0.00	8.75
Seed yield	15-2	O	10	4.44	6.44
Seed yield	32-2	O	10	19.43	21.43
Seed yield	23-15	O	10	19.45	21.45
Seed yield	28-9	O	10	38.82	40.82
Seed yield	25-3	O	10	39.82	55.81
Seed yield	28-12	O	10	66.93	68.93
Seed yield	23-8	O	10	81.08	83.08
Seed yield	31-12	O	10	104.01	106.09
Seed yield	21-4	G	18	1.20	3.20
Seed yield	31-6	G	18	21.88	29.15
Seed yield	21-8	G	18	37.47	39.47
Seed yield	15-4	G	18	42.38	44.38
Seed yield	22-17	G	18	61.63	76.76
Seed yield	22-18	G	18	68.66	76.76
Seed yield	27-3	G	18	94.50	105.50
Seed yield	15-11	G	18	95.47	97.47
WUE	1-1	G	18	51.60	53.60

CHAPTER 5. GENERAL CONCLUSIONS

High-throughput phenotyping of physiological traits driving yield is an active area of investigation and application in breeding programs to increase the rate of genetic gain for crop production. This dissertation investigated areas related to high-throughput phenotyping and physiological traits driving soybean yield by expanding upon previous studies in elite germplasm to a panel of soybean of diverse ancestries. In Chapter 2, genotype and seeding rate interactions for yield, seed components, and agronomic traits were evaluated, and adaptive elastic net models identified the underlying physiological traits predicting yield response to three levels of seeding rate. A significant genotype x seeding rate interaction was only detected for lodging, and not for yield, seed weight, seed oil percentage, seed protein percentage, height, or maturity. These results suggest that current soybean germplasm and soybean of wide genetic ancestry respond similarly to seeding rate and implies introgression of diverse material may not detrimentally affect yield response to seeding rate. Physiological traits predicting the yield response within and across seeding rates were summarized, with chlorophyll traits determined as the leading predictors across seeding rates in this study. Our further characterization of diverging traits between the seeding rate yield models may provide the research community targets for soybean improvement for current and future seeding rate practices. Moreover, further characterizing genotype x seeding rate across diverse germplasm to understand the mechanisms underlying yield response to seeding rate is an important direction for future research and soybean improvement.

Chapter 3 characterized the genetic variation in biomass partitioning strategies in the 32-parent subset of the SoyNAM soybean panel also used in Chapter 2. Three reproductive stages (R1, R4, R8) were destructively harvested and partitioned into stems, petioles, leaves,

pod, seed, and/or fallen residue (petioles and leaves) component dry weights were calculated as proportions of final biomass. Dry biomass weights were significantly different for genotypes at each growth stage, but significant differences in biomass proportions were mainly detected in the later reproductive growth stages, R4 and R8. Overall, higher biomass percentages in reproductive organs at R4 and R8 correlated positively with grain yield; however, a higher percentage of leaves at R1 correlated positively with pod percentage at R4 and harvest index at R8. The strongest relationship with grain yield was final biomass weight, but harvest index was negatively correlated with final biomass weight, indicating larger canopies may not be achieving their theoretical maximum yield. Our study additionally demonstrated genetic variation in carbon:nitrogen (C:N) residue quality for each of the partitioned residue components at physiological maturity (R8). The lack of a negative relationship between yield and C:N ratio and high heritability suggests this trait can be selected in breeding programs to improve soybean residue quality. Lower C:N ratios in soybean ratio are hypothesized to increase soil mineralization, and, therefore, increase nitrogen availability for the succeeding year's crop. Further studies are needed to unravel the genetic control of biomass partitioning strategies and C:N ratio in soybean. Ultimately, economic return in soybeans can be improved by both increasing yield through targeting biomass partitioning strategies and decreasing nitrogen inputs in the following crop season through lower soybean residue C:N ratios.

In Chapter 4, a linkage map was constructed from four RIL populations of the SoyNAM panel using the SoyNAM6K BeadChip; it was sufficient in mapping QTL for grain yield (GY) and several vegetative indices, specifically NDVI, NMDI, NWIB, PSRI, and VREI2 at seed fill (R5), with no QTL detected at beginning bloom (R1) in this study. Two

pleiotropic QTL clusters for GY and NMDI, propositioned to be previously identified semideterminate growth habit gene *Dt2*, and NWIB and VREI2 were detected. Three individual QTL for NDVI (2) and PSRI (1) were identified, with an average confidence interval of 4 cM. Therefore, these QTL could serve as a reference for future QTL mapping studies for fine mapping, candidate gene discovery, and MAS in soybean breeding.

REFERENCES

- Agudamu, T. Yoshihira, and T. Shiraiwa. 2016. Branch development responses to planting density and yield stability in soybean cultivars. *Plant Prod. Sci.* 19(3): 331–339. doi: 10.1080/1343943X.2016.1157443.
- Araus, J.L., and J.E. Cairns. 2014. Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19(1): 52–61. doi: 10.1016/j.tplants.2013.09.008.
- Board, J. 2000. Light interception efficiency and light quality affect yield compensation of soybean at low plant populations. *Crop Sci.* 40(5): 1285–1294.
- Board, J.E., and B.G. Harville. 1996. Growth dynamics during the vegetative period affects yield of narrow-row, late-planted soybean. *Agron. J.* 88(4): 567–572.
- Bolton, D.K., and M.A. Friedl. 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* 173: 74–84.
- Carpenter, A.C., and J.E. Board. 1997. Branch yield components controlling soybean yield stability across plant populations. *Crop Sci.* 37(3): 885–891.
- Carter, G.A. 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Remote Sens.* 15(3): 697–703.
- Chang, W.-S., H.-I. Lee, and M. Hungria. 2015. Soybean Production in the Americas. p. 393–400. *In* Principles of Plant-Microbe Interactions. Springer, Cham.
- De Bruin, J.L., and P. Pedersen. 2009. New and Old Soybean Cultivar Responses to Plant Density and Intercepted Light. *Crop Sci.* 49(6): 2225. doi: 10.2135/cropsci2009.02.0063.
- Dhanapal, A.P., J.D. Ray, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, C.A. King, and F.B. Fritschi. 2015. Genome-Wide Association Analysis of Diverse Soybean Genotypes Reveals Novel Markers for Nitrogen Traits. *Plant Genome* 8(3): 0. doi: 10.3835/plantgenome2014.11.0086.
- Duvick, D.N. 1984. Genetic diversity in major farm crops on the farm and in reserve. *Econ. Bot.* 38(2): 161–178.
- Edwards, J.T., L.C. Purcell, and D.E. Karcher. 2005. Soybean Yield and Biomass Responses to Increasing Plant Population among Diverse Maturity Groups. *Crop Sci.* 45(5): 1778. doi: 10.2135/cropsci2004.0570.
- Egli, D.B. 1993. RELATIONSHIP OF UNIFORMITY OF SOYBEAN SEEDLING EMERGENCE TO YIELD. *J. Seed Technol.* 17(1): 22–28.

- Elmore, R.W. 1998. Soybean cultivar responses to row spacing and seeding rates in rainfed and irrigated environments. *J. Prod. Agric.* 11(3): 326–331.
- Fawcett, J., J. Sievers, J. DeJong, and L. Rossiter. 2017. On-Farm Corn and Soybean Planter Trials. *Farm Prog. Rep.* 2016(1): 112.
- Ferreira, A.S., B. Junior, A. Antonio, F. Werner, C. Zucareli, J.C. Franchini, H. Debiassi, A.S. Ferreira, B. Junior, A. Antonio, F. Werner, C. Zucareli, J.C. Franchini, and H. Debiassi. 2016. Plant density and mineral nitrogen fertilization influencing yield, yield components and concentration of oil and protein in soybean grains. *Bragantia* 75(3): 362–370. doi: 10.1590/1678-4499.479.
- Frederick, J.R., C.R. Camp, and P.J. Bauer. 2001. Drought-stress effects on branch and mainstem seed yield and yield components of determinate soybean. *Crop Sci.* 41(3): 759–763.
- Fritschi, F.B., J.D. Ray, L.C. Purcell, C.A. King, J.R. Smith, and D.V. Charlson. 2013. Diversity and Implications of Soybean Stem Nitrogen Concentration. *J. Plant Nutr.* 36(14): 2111–2131. doi: 10.1080/01904167.2012.748800.
- García-Palacios, P., R. Milla, M. Delgado-Baquerizo, N. Martín-Robles, M. Álvaro-Sánchez, and D.H. Wall. 2013. Side-effects of plant domestication: ecosystem impacts of changes in litter quality. *New Phytol.* 198(2): 504–513. doi: 10.1111/nph.12127.
- Gentry, L.E., F.E. Below, M.B. David, and J.A. Bergerou. 2001. Source of the soybean N credit in maize production. *Plant Soil* 236(2): 175–184.
- Gentry, L.F., M.L. Ruffo, and F.E. Below. 2013. Identifying Factors Controlling the Continuous Corn Yield Penalty. *Agron. J.* 105(2): 295. doi: 10.2134/agronj2012.0246.
- Gitelson, A.A., Y. Gritz, and M.N. Merzlyak. 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* 160(3): 271–282.
- Grassini, P., J.A. Torrión, H.S. Yang, J. Rees, D. Andersen, K.G. Cassman, and J.E. Specht. 2015. Soybean yield gaps and water productivity in the western US Corn Belt. *Field Crops Res.* 179: 150–163.
- Gray, S.B., O. Dermody, S.P. Klein, A.M. Locke, J.M. Mcgrath, R.E. Paul, D.M. Rosenthal, U.M. Ruiz-Vera, M.H. Siebers, and R. Strellner. 2016. Intensifying drought eliminates the expected benefits of elevated carbon dioxide for soybean. *Nat. Plants* 2(9): 16132.

- Green, C.J., and A.M. Blackmer. 1995. Residue Decomposition Effects on Nitrogen Availability to Corn following Corn or Soybean. *Soil Sci. Soc. Am. J.* 59(4): 1065–1070. doi: 10.2136/sssaj1995.03615995005900040016x.
- Green-Tracewicz, E., E.R. Page, and C.J. Swanton. 2011. Shade Avoidance in Soybean Reduces Branching and Increases Plant-to-Plant Variability in Biomass and Yield Per Plant. *Weed Sci.* 59(01): 43–49. doi: 10.1614/WS-D-10-00081.1.
- Haboudane, D., J.R. Miller, E. Pattey, P.J. Zarco-Tejada, and I.B. Strachan. 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* 90(3): 337–352. doi: 10.1016/j.rse.2003.12.013.
- Heinrich, V., G. Krauss, C. Gotze, and C. Sandow. 2011. IDB - Index DataBase. Index Database IDB Proj. <https://www.indexdatabase.de/> (accessed 21 March 2015).
- Huang, Y., L. Yuan, K.N. Reddy, and J. Zhang. 2016. In-situ plant hyperspectral sensing for early detection of soybean injury from dicamba. *Biosyst. Eng.* 149: 51–59. doi: 10.1016/j.biosystemseng.2016.06.013.
- Jin, J., X. Liu, G. Wang, L. Mi, Z. Shen, X. Chen, and S.J. Herbert. 2010. Agronomic and physiological contributions to the yield improvement of soybean cultivars released from 1950 to 2006 in Northeast China. *Field Crops Res.* 115(1): 116–123. doi: 10.1016/j.fcr.2009.10.016.
- Johnson, D.M. 2014. An assessment of pre-and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sens. Environ.* 141: 116–128.
- Keep, N.R., W.T. Schapaugh, P.V.V. Prasad, and J.E. Boyer. 2016. Changes in Physiological Traits in Soybean with Breeding Advancements. *Crop Sci.* 56(1): 122. doi: 10.2135/cropsci2013.07.0499.
- Koester, R.P., B.M. Nohl, B.W. Diers, and E.A. Ainsworth. 2016. Has photosynthetic capacity increased with 80 years of soybean breeding? An examination of historical soybean cultivars. *Plant Cell Environ.*: n/a-n/a. doi: 10.1111/pce.12675.
- Kumudini, S., D.J. Hume, and G. Chu. 2001. Genetic Improvement in Short Season Soybeans. *Crop Sci.* 41(2): 391. doi: 10.2135/cropsci2001.412391x.
- Lawlor, D.W. 2002. Carbon and nitrogen assimilation in relation to yield: mechanisms are the key to understanding production systems. *J. Exp. Bot.* 53(370): 773–787. doi: 10.1093/jexbot/53.370.773.

- Lee, C.D., D.B. Egli, and D.M. TeKrony. 2008. Soybean Response to Plant Population at Early and Late Planting Dates in the Mid-South. *Agron. J.* 100(4): 971. doi: 10.2134/agronj2007.0210.
- Ma, B.L., L.M. Dwyer, C. Costa, E.R. Cober, and M.J. Morrison. 2001. Early prediction of soybean yield from canopy reflectance measurements. *Agron. J.* 93(6): 1227–1234.
- Ma, B.L., M.J. Morrison, and H.D. Voldeng. 1995. Leaf greenness and photosynthetic rates in soybean. *Crop Sci.* 35(5): 1411–1414.
- Monteith, J.L., and C.J. Moss. 1977. Climate and the efficiency of crop production in Britain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 281(980): 277–294.
- Morrison, M.J., H.D. Voldeng, and E.R. Cober. 2000. Agronomic changes from 58 years of genetic improvement of short-season soybean cultivars in Canada. *Agron. J.* 92(4): 780–784.
- Nutter Jr, F.W., G.L. Tylka, J. Guan, A.J.D. Moreira, C.C. Marett, T.R. Rosburg, J.P. Basart, and C.S. Chong. 2002. Use of remote sensing to detect soybean cyst nematode-induced plant stress. *J. Nematol.* 34(3): 222.
- Pedersen, P., and J.G. Lauer. 2004. Response of soybean yield components to management system and planting date. *Agron. J.* 96(5): 1372–1381.
- Rafalski, J.A. 2010. Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13(2): 174–180.
- Rincker, K., R. Nelson, J. Specht, D. Sleper, T. Cary, S.R. Cianzio, S. Casteel, S. Conley, P. Chen, V. Davis, C. Fox, G. Graef, C. Godsey, D. Holshouser, G.-L. Jiang, S.K. Kantartzi, W. Kenworthy, C. Lee, R. Mian, L. McHale, S. Naeve, J. Orf, V. Poysa, W. Schapaugh, G. Shannon, R. Uniatowski, D. Wang, and B. Diers. 2014. Genetic Improvement of U.S. Soybean in Maturity Groups II, III, and IV. *Crop Sci.* 0(0): 0. doi: 10.2135/cropsci2013.10.0665.
- Rouse Jr, J., R.H. Haas, J.A. Schell, and D.W. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTS.
- Saryoko, A., Y. Fukuda, I. Lubis, K. Homma, and T. Shiraiwa. 2018. Physiological activity and biomass production in crop canopy under a tropical environment in soybean cultivars with temperate and tropical origins. *Field Crops Res.* 216: 209–216. doi: 10.1016/j.fcr.2017.11.012.
- Slattery, R.A., A. VanLoocke, C.J. Bernacchi, X.-G. Zhu, and D.R. Ort. 2017. Photosynthesis, Light Use Efficiency, and Yield of Reduced-Chlorophyll Soybean Mutants in Field Conditions. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00549.

- Srinivasan, V., P. Kumar, and S.P. Long. 2016. Decreasing, not increasing, leaf area will raise crop yields under global atmospheric change. *Glob. Change Biol.*: n/a-n/a. doi: 10.1111/gcb.13526.
- Suhre, J.J., N.H. Weidenbenner, S.C. Rowntree, E.W. Wilson, S.L. Naeve, S.P. Conley, S.N. Casteel, B.W. Diers, P.D. Esker, J.E. Specht, and V.M. Davis. 2014. Soybean Yield Partitioning Changes Revealed by Genetic Gain and Seeding Rate Interactions. *Agron. J.* 106(5): 1631. doi: 10.2134/agronj14.0003.
- Tanksley, S.D., and S.R. McCouch. 1997. Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. *Science* 277(5329): 1063–1066. doi: 10.1126/science.277.5329.1063.
- Thompson, J.A., and R.L. Nelson. 1998. Utilization of diverse germplasm for soybean yield improvement. *Crop Sci.* 38(5): 1362–1368.
- USDA-NASS. 2016. Crop production 2016 Summary. USDA-NASS, Washington, DC.
- Vega, C. 2000. Reproductive Allometry in Soybean, Maize and Sunflower. *Ann. Bot.* 85(4): 461–468. doi: 10.1006/anbo.1999.1084.
- Weiner, J., and S.C. Thomas. 1986. Size Variability and Competition in Plant Monocultures. *Oikos* 47(2): 211. doi: 10.2307/3566048.
- Wilson, E.W., S.C. Rowntree, J.J. Suhre, N.H. Weidenbenner, S.P. Conley, V.M. Davis, B.W. Diers, P.D. Esker, S.L. Naeve, and J.E. Specht. 2014. Genetic Gain \times Management Interactions in Soybean: II. Nitrogen Utilization. *Crop Sci.* 54(1): 340–348.
- Xavier, A., B. Hall, A.A. Hearst, K.A. Cherkauer, and K.M. Rainey. 2017. Genetic Architecture of Phenomic-Enabled Canopy Coverage in *Glycine max*. *Genetics* 206(2): 1081–1089. doi: 10.1534/genetics.116.198713.
- Xavier, A., D. Jarquin, R. Howard, V. Ramasubramanian, J.E. Specht, G.L. Graef, W.D. Beavis, B.W. Diers, Q. Song, P.B. Cregan, R. Nelson, R. Mian, J.G. Shannon, L. McHale, D. Wang, W. Schapaugh, A.J. Lorenz, S. Xu, W.M. Muir, and K.M. Rainey. 2018. Genome-Wide Analysis of Grain Yield Stability and Environmental Interactions in a Multiparental Soybean Population. *G3 Genes Genomes Genet.* 8(2): 519–529. doi: 10.1534/g3.117.300300.
- Yu, J., and E.S. Buckler. 2006. Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* 17(2): 155–160.
- Zanon, A.J., N.A. Streck, and P. Grassini. 2016. Climate and management factors influence soybean yield potential in a subtropical environment. *Agron. J.* 108(4): 1447–1454.

- Zhao, X., S.-H. Zheng, Fatichin, A. Suzuki, and S. Arima. 2014. Varietal Difference in Nitrogen Redistribution from Leaves and Its Contribution to Seed Yield in Soybean. *Plant Prod. Sci.* 17(1): 103–108. doi: 10.1626/pp.s.17.103.
- Zhu, X.-G., S.P. Long, and D.R. Ort. 2010. Improving Photosynthetic Efficiency for Greater Yield. *Annu. Rev. Plant Biol.* 61(1): 235–261. doi: 10.1146/annurev-arplant-042809-112206.
- Zygielbaum, A.I., T.J. Arkebauer, E.A. Walter-Shea, and D.L. Scoby. 2012. Detection and measurement of vegetation photoprotection stress response using PAR reflectance. *Isr. J. Plant Sci.* 60(1–2): 37–47.

ACKNOWLEDGMENTS

I would like to thank all the people who supported and made possible this research, and am humbled by their work ethic, passion, and integrity. I would like to thank my committee chair, Danny Singh, and my committee members Drs. Thomas Lubberstedt, Maria Salas Fernandez, Michael Castellano, and Andy VanLooche. Many thanks towards research staff Brian Scott, Jae Brungardt, Jennifer Hicks, and former Postdoc's Drs. Jiaoping Zhang, Chowda Reddy Rekalakunta, and Teshale Mamo. Fellow graduate students were invaluable in their efforts of data collection and analysis help, including Tara Moellers, Kyle Parmley, Kevin Falk, Matt Carroll, Sarah Jones, David Blystone, Sara Coser, Leonardo Azevedo, Martha Ibore, and John Shook. Numerous undergraduate research workers gave their time to data collection, and although not mentioned by name, their work is much appreciated.

Gratitude is also deserved toward Dr. Jim Rouse and staff of the Iowa Crop Improvement Association for initial assistance with field plot experiment planting and harvest, and Mike Fiscus and staff at the Iowa State Agronomy Farm for assistance with field operation and facilities. They were crucial in initially helping us with field experiments as a young research lab. Several other facilities and training within Iowa State University were generously shared by several people, including Kerry Culp of the Plant and Soil Analysis Lab, Danielle Wilson and Dr. Ken Moore of the Biomass and Forage Quality Lab, and Dr. Charles Hurburgh and Glenn Rippke of the Grain Quality Lab. Statistic suggestions and guidance were kindly provided by Dr. Somak Dutta and Adam Martin-Schwarze of the Iowa State University's Department of Statistics.

My sincere gratitude to the following organization for supporting me with a graduate research assistant stipend and tuition: R.F. Baker Center for Plant Breeding, Monsanto chair in soybean breeding, Iowa Soybean Association and Iowa State University. I sincerely appreciate their financial support for research experimentation. I am very grateful for the National Council of Commercial Plant Breeders for the graduate student award (2017), and ISU for travel and research awards.

Finally, I'd like to thank the many friends, colleagues, department faculty and staff I've encountered at Iowa State. Their presence, diverse backgrounds, and discussion of varied research topics made for a stimulating environment and provided fresh outlooks and approaches to this research. To my close friends and family, I cannot give enough gratitude for the encouragement and support shared throughout the duration of this dissertation.