

IMPROVEMENT OF COMMON BEAN (*PHASEOLUS VULGARIS* L.) FOR  
REGIONAL FOOD SYSTEMS IN THE NORTHEASTERN UNITED STATES

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Kristen A. Loria

May 2021

© 2021 Kristen A. Loria

## ABSTRACT

Dry common bean (*Phaseolus vulgaris* L.) is an important but underutilized staple crop for regional food systems in the Northeastern United States. Plant breeding can help to develop improved varieties that meet regional farmer needs while providing a compelling product for modern consumers. One strategy is the expansion of disease resistance and improved agronomic performance to non-commodity, heirloom-like market classes that hold potential for increased consumer engagement and higher grower returns. To this end, Chapter II describes efforts to combine important disease resistance for three major pathogens in the northeastern United States as well as favorable agronomic traits such as yield and upright growth habit into two heirloom varieties with unique seed coat phenotypes, using a marker-assisted selection model. The four targeted diseases were common bacterial blight, caused by *Xanthomonas campestris* pv. *phaseoli*, anthracnose, caused by *Colletotrichum lindemuthianum*, and *Bean common mosaic virus*. In the interest of future food security, in the Northeast and elsewhere, access to genetically diverse germplasm with locally or regionally adaptive traits is crucial. Decentralized networks of seed savers and regional seed companies in North America are an underutilized resource for important and useful genetic material. Chapter III analyzes genotypic and phenotypic diversity of 18 seed sources of ‘Jacob’s Cattle’ dry bean across seed saver networks. Significant genotypic and phenotypic divergence was identified, indicating that North American seed saver networks are a significant repository of *in-situ* germplasm conservation.

## BIOGRAPHICAL SKETCH

Kristen was born and raised in Ithaca, NY, and cut her teeth in the plant sciences at the age of ten, filling pipet tip boxes for a dime per box. Kristen studied Politics, Environmental Studies and Spanish at New York University for two years before returning to Ithaca to pursue a B.S. in Natural Resources at Cornell University. As an undergraduate research assistant in the Broussard lab, Kristen studied the community impacts of urban trees in the MillionTrees NYC project and began her food systems education via a teaching internship with the Groundswell Center for Local Food and Farming, independent study of community garden communication networks and studies in food and agriculture policy in Washington, DC.

After graduation, Kristen worked in school garden education and became involved with Greenhorns, a grassroots young farmer organization. Eventually Kristen moved to the Hudson Valley to work for the Greenhorns as program coordinator for the Farm Hack project. After a year of organizing everything from a 400-person land access symposium to a farmer “weed dating” mixer at the Common Ground Fair, Kristen ultimately transitioned to full-time farming, dabbling in dairy before settling into organic vegetable and field crop production.

While working at Sparrowbush Farm in Livingston, NY for several years, Kristen developed a particular interest in dry beans and grew her first seed crop, Vermont Cranberry dry beans. These experiences fostered an interest in seed systems and plant breeding that served the diverse agricultural community of the Northeast. This spark set her on a winding path back home to Ithaca to pursue a Master’s in Plant Breeding and Genetics, with the unique opportunity to work in the Mazourek lab managing participatory vegetable variety trials with the Northern Vegetable Improvement Collaborative as well as breeding dry beans for regional food systems.

This work is dedicated to anyone  
who has ever saved a bean seed –  
we stand on your shoulders!

## ACKNOWLEDGMENTS

I am incredibly grateful to everyone who patiently and generously shared their knowledge and time with me these past few years. First, my advisor Michael Mazourek for his enthusiastic support in allowing me to pursue my dry bean dreams here at Cornell, and for connecting me with a wonderful network of researchers and farmers through the NOVIC project. I am also indebted to my fellow lab members Lauren Brzozowski, Jack Fabrizio, Chris Hernandez, and Kyle LaPlant for helping guide me through the wild world of genomics and graduate school generally. Many thanks to Greg Inzinna, John Kurtz, and Maddie Halpert for their support with field research, and Paige Reeves for her mentorship at the lab bench. I'm also so grateful to my committee members Chris Smart and Margaret Smith for their generosity of time, thoughtful reflection and mentorship.

Thanks to Heron Breen of Fedco and Kathryn Gilbery of Seed Saver's Exchange for their collaboration on the heirloom bean diversity project, and thanks as well to the incredible public bean breeding community who thoughtfully answered my inquiries, generously shared resources and even seeds – among them Travis Parker, Paul Gepts, Jim Myers, Jim Kelly and Evan Wright. Their spirit of collaboration and exchange is inspiring. I'm also grateful to Heather Darby, Joe Bossen, June Russell and Jay Goldmark for their work supporting dry bean production for the Northeast and their willingness to share ideas and inspiration.

I am forever grateful to my mother, Rosemary Loria, who first shared with me the beauty and wonder of the plant world and has continued to support and inspire me along the way. I'm not crying, you're crying!

Finally, I want to acknowledge the humbling fact that every seed I planted, pollinated and harvested over the past three years finds its way to me thanks to many generations of farmers, gardeners, and seed savers that have come before me. I am thrilled to follow this ever-unwinding thread and hope to fruitfully and respectfully steward the gifts they have provided to us, for those that come after.

## TABLE OF CONTENTS

Abstract .....	iii
Biographical Sketch .....	iv
List of Tables .....	ix
List of Figures .....	x
Chapter 1: Introduction .....	1
Dry Common Bean Production in The United States.....	1
Genetic Improvement of Common Bean for Commodity Production Systems .....	3
Towards a Model for Improvement of Common Dry Bean for Regional Food Systems in the Northeast.....	5
Conclusion .....	22
References.....	24
Chapter 2: Improvement of Disease Resistance in Specialty Dry Bean Market Classes for Regional Markets .....	32
Significance of Seedborne Pathogens in Common Bean.....	32
Plant-Pathogen Interactions in the Incidence of Seedborne Disease .....	33
Selected Seedborne Pathogens of Common Bean in the Northeast.....	34
Improvement of Specialty Dry Bean Germplasm for Northeastern Regional Food Systems ..	43
References .....	59

Chapter 3: Exploration of Intra-Varietal Diversity of a Traditional North American Bean within Decentralized Seed Saving Networks .....	64
Introduction.....	64
Materials and Methods.....	64
Results.....	74
Discussion.....	87
References.....	91
Appendix I: NOVIC 2020 Squash Variety Trial Report .....	88
Appendix II: Cornell Small Farms Program Small Farms Quarterly Articles .....	95
Appendix III: Dry Bean Grower Needs Assessment .....	107

## LIST OF TABLES

Table 1.1 Models for variety development based on end users and use-values. ....	9
Table 1.2 Summary of priority traits in common bean for regional production.....	12
Table 2.1 Seedborne diseases and causal pathogens included.....	49
Table 2.2 Primers and annealing temperatures used in a High-Resolution Melt (HRM) assay. ..	50
Table 3.1 Passport data for all seed sources. ....	73
Table 3.2 Total SNP counts within varieties and seed sources.....	75

## LIST OF FIGURES

Figure 1.1 A visualization of cross-over Genotype X Environment (G X E) interaction .....	15
Figure 2.2 Percent disease ratings for common bacterial blight of parent and bulked F <sub>4</sub> family genotypes pooled by resistance genotype .....	54
Figure 3.3. Results from Discriminant Analysis of Principal Components (DAPC) .....	79
Figure 3.4 a-b. Distribution of minor allele genotype at SNPs identified across chromosomes ..	81
Figure 3.5. Boxplots of phenotypic trait values of ‘Jacob’s Cattle’ seed sources. ....	84
Figure 3.6 Seed phenotypes of seed sources of ‘Jacob’s Cattle’ .....	85
Figure 3.7 PC1 values for all 18 ‘Jacob’s Cattle’ seed sources plotted against measurements for 100-seed weight phenotypic trait. ....	86

## CHAPTER I

### INTRODUCTION

#### *Dry common bean production in the United States*

Approximately 1.6 million acres of dry beans are harvested in the United States annually (USDA-NASS). This ranks the US as #6 in world production of dry beans, behind Brazil, India, China, Burma and Mexico. Common bean has been grown in the Northeast by indigenous communities since the 12<sup>th</sup> century C.E. and has been a staple of indigenous foodways in the region ever since (Mann and Fields 1997). After European colonization, New York state was the first to commercially produce dry beans; Stephen Coe in 1836 planted and sold white pea beans in Orleans County, becoming the first non-indigenous commercial dry bean grower on record (Harveson and Urrea 2016). According to written records, which likely omit native producers of dry beans, New York State was the leading producer of dry beans in the country prior to 1899 (Stone 1903). In 1909, 115,000 acres or 1,681,000 bushels of dry beans were grown in New York state. Beans were grown as an affordable source of protein compared to meat, which was much more expensive. At the time, mostly white (navy) beans, kidney and yellow eye types were cultivated (Hardenburg 1914).

Several factors contributed to the spread of bean production westward during colonial expansion. First, the US government bought large quantities of dry beans to feed troops in the Civil War, increasing prices as well as the culinary popularity of dry beans among settlers establishing their farms in newly occupied Western territories. Several years of poor wheat production in the mid 19<sup>th</sup> century due to a new pest, *Sitophilus granarius*, also encouraged more dry bean cultivation in the West (Stone 1903). Due to several wet seasons and recurring

problems of mosaic virus and root rots, bean production in New York decreased sharply in the early 20<sup>th</sup> century and never recovered. Notably, with increasing knowledge of crop rotation and cultivars resistant to *Bean common mosaic virus* (BCMV), these obstacles have in part been overcome.

Today the five largest dry bean producing states are North Dakota (32%), Michigan (17%), Nebraska (11%), Minnesota (9%) and Idaho (8%). Commodity dry bean production is organized into standard market classes, which are pinto, light red kidney, dark red kidney, black, navy, great northern, small white, pink, small red and cranberry, with pinto, black and navy comprising the vast majority of production (USDA-AMS 2019). In New York state 12,619 acres were grown in 2017, up from 9,642 acres in 2012 (USDA-NASS 2017). Most of this production is shipped to Midwestern processors. In 2019, 26,192 acres of organic dry beans were grown in the United States, a figure that is growing in both acreage and proportion of total production (USDA NASS 2019). Dry bean consumer demand in the United States has also increased in recent years, reaching 9.6 pounds per person per year in 2018, up from 7.6 pounds in 2016 (USDA-AMS 2019). The majority of production is processed into canned products, with other end uses including dry packaged beans, brined bean packets, and bean flour (AgMRC 2017). Bean processing today is highly centralized with the number of processors and canneries having decreased over time. Canneries have made more stringent requirements for processing facilities that serve them, most recently requiring optical color sorters that require an investment of hundreds of thousands of dollars. These factors have made it more difficult for smaller processors to remain competitive (M. Callan, personal communication, 2020).

## ***Genetic Improvement of Common Bean for Commodity Production Systems***

Mainstream dry bean breeding efforts in the public and private sectors focus on several major market classes and prioritize the development of high-yielding cultivars with specific quality traits required by canneries (B. O. Brouwer, Murphy, and Jones 2015; Kelly 2010; Posa-Macalincag et al. 2002). Processors sell to canneries in New York or the Midwest, to dry packaged markets in the Northeast and beyond, or process and market dry products themselves (M. Callan, personal communication). Other priority traits include upright growth habit needed for direct harvest, days to maturity, and biotic and abiotic stress resistance or tolerance (P. Miklas et al. 2006). In the United States, priority traits with respect to biotic and abiotic stress vary by region. The majority of these traits are quantitative except for disease resistance - which can be either quantitative, qualitative or both – as well as some quality traits such as seed color, pattern or shape. As a self-pollinating crop, dry bean cultivars are typically released as pure lines resulting from single or several plant selections. In later stages of the breeding pipeline, advanced lines are tested in multiple locations within the target growing region with the goal of identifying a cultivar with stable, high yields across environments. Only lines that perform favorably in canning tests are advanced, a phenotyping process that requires significant time and labor (Kelly 2010).

Pedigree selection is the most common method for bean cultivar development, in which an initial cross is made between two or more parents. For  $F_2$  through  $F_5$  generations, single plant selection, with evaluation for traits such as maturity time, seed shape and color, and growth habit, is typically practiced. In generations  $F_6$  to  $F_8$ , full-sib families are evaluated for quantitative traits such as yield and disease resistance. Marker-assisted breeding can be utilized to cull undesirable segregating individuals in earlier generations, increasing efficiency of

selection. Marker-assisted breeding is especially useful in selecting for qualitative disease resistance, with numerous molecular markers being developed for major resistance genes. (Assefa et al. 2019; Kelly 2010; P. Miklas et al. 2006). Subsequent generations are evaluated in plant or family rows for quantitative traits such as yield and canning quality. The deployment of genomic selection, in which quantitative traits are predicted using genomic data rather than direct phenotyping, has been limited in common bean (Keller et al. 2020). Single-seed-descent is another method increasingly favored in “speed-breeding” programs, in which a single seed from each plant from each generation is advanced using glasshouse and winter nursery cultivation to achieve multiple generations per year (Watson et al. 2018).

Mainstream plant breeding has proven effective at developing annual grain crop varieties that respond favorably to inputs and perform well in predominant agricultural-producing regions. Two types of adaptation are targeted to achieve genetic gain in the face of genotype x environment (G x E) interaction: 1) broad adaptation for uniformly high performance across environments, or 2) specific adaptation to local or regional environments, used to select for genotypes that perform best in local environment and production contexts (Messmer et al. 2011). The vast majority of mainstream plant breeding programs employ a strategy of broad adaptation, using multi-environment trials to select cultivars that show the most uniform genetic gain across environments (B. O. Brouwer, Murphy, and Jones 2015). Different statistical models can be employed to control for G x E interaction in breeding programs. Among the most promising is the additive main effects and multiplicative interaction (AMMI model), as described by the following equation:

$$y_{ij} = \mu + G_i + E_j + \sum_{k=1}^K a'_{ki} b'_{kj} + e_{ij}$$

where  $y_{ij}$  is phenotypic response of the  $i^{th}$  genotype in the  $j^{th}$  environment,  $\mu$  is the mean response,  $G_i$  is the genotypic effect of  $i$ ,  $E_j$  is the environmental effect of  $j$ ,  $\sum_{k=1}^K a'_{ki} b'_{kj}$  is the partitioned effects of  $a_{ki}$  genotypic sensitivity of genotype  $I$  to hypothetical environmental variable  $b_k$  in environment  $j$  for the  $k$ th term (Bustos-Korts et al. 2018; de Leon et al. 2016).

In high-input nurseries and on-farm trials, G x E interaction is reduced by the replacement of ecosystem services with synthetic fertilizer, pesticides, and irrigation (Salvatore Ceccarelli 1996; Cooper and Blyth 1996). This means that a cultivar that is adapted to these artificially optimized growing conditions in many places may not perform well in sub-optimal conditions, leading Ceccarelli (1994) to hypothesize that such cultivars may actually be specifically adapted to optimal crop growth conditions rather than truly “widely adapted”. This reduces the likelihood that a variety will perform well in production systems that seek to reduce external inputs, such as agroecological production systems (Ponisio et al. 2014). For regions such as the Northeast where growing conditions are highly heterogeneous, average farm sizes are smaller and more variation in growing practices exists, there is a higher risk that a “one size fits all” approach will result in low or unstable productivity as well as the omission of region-specific traits that may be important to farmers and consumers (Atlin et al. 2001).

### ***Towards a model for improvement of common dry bean for regional food systems in the Northeast***

There is a strong demand for regionally and locally produced foods in the Northeast, with four of the ten top states producing for local markets (Massachusetts, New York, Pennsylvania and Vermont) located in the region (USDA NASS 2016). However, these markets have largely focused on vegetable, fruit and meat producers and regional supply chains for staple crops such

as small grains, pulses, and oilseeds have lagged behind. In order to first define the agricultural systems in the Northeast that plant breeding efforts should seek to support, we find the Northeast Sustainable Agriculture and Research and Education’s outcome statement (SARE n.d.) to be a fruitful starting point:

*Agriculture in the Northeast will be diversified and profitable, providing healthful products to customers. Farmers and the people they work with will steward resources to ensure sustainability and resilience, and foster conditions where farmers have high quality of life and communities can thrive.*

In the context of plant breeding, we can use this vision to define a target growing environment in which use of external inputs are minimized and crop health is rooted in the health of soil ecosystems and diverse cropping rotations, a production system that we will henceforth refer to as “agroecological” production systems in the tradition of Altieri (Altieri 1996). We contrast agroecological production systems with “high-input” systems heavily reliant on synthetic fertilizers, pesticides and, in some climates, supplemental irrigation. We recognize that farms exist within a dynamic array of production practices and management decisions, not within a binary, but for the purposes of discussion will contrast “agroecological” with “high-input” production systems in this paper. We seek to breed varieties not only for the agricultural systems we have today, but more importantly for the agricultural systems necessary for the future wellbeing of our region. In this definition we also seek to move the agricultural sustainability conversation in our region beyond an “organic vs. conventional” paradigm.

Regional and local food marketplaces operate via a values-based supply chain that allows growers to distinguish themselves, their products and their production methods to consumers from mainstream supply chains. Consumers choose to participate in regional or local food marketplaces due to a myriad of motivations, including support of local economies, organic or agroecological production methods, unique crop varieties or breeds, and more (Clancy and Ruhf 2010). This model embodies Porter and Kramer’s concept of “shared value,” such that regional farmers can access higher-value based on a perception of external value realized by a purchase (Berti and Mulligan 2016). This shared value in the context of Northeast food systems could be economic wellbeing of rural communities as well as stewardship of land and water (Berti and Mulligan 2016). Inevitably this shared value must be communicated to the consumer in some way. In commodity supply chains, labeling systems such as USDA Certified Organic are necessary to communicate shared value. Within regional food systems, shorter supply chains and identity-preserved or direct marketing strategies offer more opportunities to communicate shared value to consumers.

Increasing the value of grains and legumes via regional marketplaces is a key strategy for maintaining and expanding viability of regional farms in the face of competing land uses and low commodity prices (Berti and Mulligan 2016). The predominant grain legume grown in the Northeast is soybean (*Glycine max*), which has comparable yields per acre to dry bean, but typically lower gross returns due to significantly lower price per bushel on the commodity market (USDA-NASS n.d.). As a food crop, dry bean prices do not typically follow corn and soy prices, which, when grown in rotation systems, can help growers to mitigate risk from price fluctuations (Myers 2002). Interestingly, Crowder and Reganold (2015) found that the integration of more diverse crops into a rotational system also increased profitability cost/benefit

ratios, and use of crop rotation reduced yield gaps between conventional and organic production. This indicates that robust regional markets that support diversification of cropping systems may also help to support farm profitability over time by encouraging sound agroecological production methods. As with consumer preference based on regional production, higher profitability for organically labeled products mitigates economies of scale and allows small and mid-scale farms to remain profitable. Farms that employ agroecological production methods can use organic certification to capture value for their investments in sustainable production systems, or find alternative ways to convey this form of shared value.

The potential for expansion of dry bean cultivation in the Northeast is especially timely given the recent progress on developing specialty small grain varieties as well as strengthening supply chains for value-added production and direct-to-consumer markets in the Northeast. Small grains breeding work has focused on the use of traditional germplasm with unique quality traits in targeted crosses to modern germplasm, as well as participatory farmer-breeder-chef collaboration (Kissing-Kucek 2017b, 2017a; Sorrells et al. 2015). Policy and supply chain efforts of the past decade that incentivize use of regionally grown grains by value-added businesses, such as New York's Farm Brewery Law and GrowNYC's Regional Grains Project, have created a sharp increase in demand for high quality small grains in the Northeast, supporting a community of growers honed in on regional market supply chains. Such a landscape is ripe for the addition of other crops, such as dry beans, suitable to these regional supply chains that generally offer higher returns for growers (B. O. Brouwer, Murphy, and Jones 2015).

### **Collaborative development of regional plant breeding models**

Beyond optimization of a crop's agronomic performance, plant breeding can support regional food systems by facilitating the development of varieties that excel within economic and

cultural parameters of the region, thereby adding value to the farmer’s crop and meeting regional needs (B. O. Brouwer, Murphy, and Jones 2015). For example, improvement of crop pest resistance is a common plant breeding goal and can support regional food systems by increasing yields, decreasing input costs and mitigating ecological effects of control strategies. An equally important breeding goal for regional food systems, however, is the use-value of a variety, defined by a region’s food culture and economy. Brouwer et al. 2015 contrasts four models for plant breeding, based on differing “users” that are prioritized (Table 1.1)

Table 1.1 Models for variety development based on end users and use-values that can be prioritized within a complex agri-food system (Adapted from Brouwer et al. 2015). Different models result in distinct breeding goals based on these priorities.

Model	Goal	Example
1. Commodity (dominant model)	<ul style="list-style-type: none"> <li>• Broad-adaptation across environments</li> <li>• Location and variety-neutral quality traits</li> </ul>	Dry bean canning quality as primary end-use; conformity to major market class designations prioritized
2. Consumer-based	<ul style="list-style-type: none"> <li>• Conservation of historic varieties</li> <li>• High-dollar/culinary end-use prioritized</li> </ul>	Promotion of heirloom varieties due to perceived biodiversity and culinary value
3. Farmer-based	<ul style="list-style-type: none"> <li>• Local agronomic adaptation through informal farmer selection</li> <li>• “Folk varieties”</li> </ul>	Development of farm-specific strains based on selection of initial variety or population
4. Farmer-breeder-consumer collaboration	<ul style="list-style-type: none"> <li>• Local or regional adaptation and end-use value prioritized</li> <li>• Leverage complementary expertise</li> </ul>	Creation of a cultivar with superior agronomic traits as well as appropriateness/value within a regional food system and cuisine

Local and regional demand for heirloom dry bean varieties, classifiable by the “Consumer-based” model above, has arisen in contrast to commodity supply chains (B. Brouwer

et al. 2016). “Heirloom” is a complex term but will be used here to describe population-varieties that have been cultivated for many generations, and have their origins in either landraces, formally released cultivars or both (Paul Gepts 1998; Silbernagel and Hannan 1988). In the United States and Canada, landraces cultivated by Native Americans as well as germplasm brought by European settlers, with likely origins in South American indigenous communities, formed the basis of modern bean market classes and many continue to be cultivated as traditional varieties at both subsistence and commercial scales (Kelly 2010). Heirloom varieties of many crop types are increasingly in demand by consumers due to a range of motivations including perceived cultural significance or “story,” culinary value or contribution to biodiversity preservation. This commercial demand can be framed as a direct response to the perceived homogenization of crop varieties in the marketplace by commodity-oriented production (B. Brouwer et al. 2016; Jordan 2007). Often crop landraces are cultivated due to a failure of modern breeding efforts to successfully meet farmer needs (Witcombe et al. 1996).

In visioning a third model that seeks to develop dry bean varieties for regional food systems in the Northeast, one must incorporate the modes of production, supply chains and consumer outlets that would engage in such a system. This visioning needs to be conducted by a stakeholder group that extends beyond the plant breeding community, but would potentially take into account 1) target production environment/s including scale of farm, geographic range, cropping systems, production equipment and post-harvest infrastructure; 2) target end-use incorporating preferences of processors, marketers, chefs and consumers within a region; and 3) collectively-held goals of agricultural sustainability, as defined above. These parameters are drawn by actors as part of a complex network rather than a centralized commodity supply chain, necessitating more dynamic engagement of actors within the breeding process. Plant breeders

play a crucial role in accessing suitable genetic material and providing technical expertise and science-based oversight of the breeding process (B. O. Brouwer, Murphy, and Jones 2015; J. C. Dawson et al. 2011). Many of these strategies utilize participatory and decentralized plant breeding methods that have been deployed in regionally-focused plant breeding efforts in small grains and vegetable crops (Brouwer et al. 2015; Dawson et al. 2011; Dawson et al. 2008; Kissing-Kucek 2017a; Lammerts van Bueren and Myers 2012).

Table 1.2 summarizes potential priority traits for dry beans in the northeast, categorized by type of actor. This summary is the result of grower surveys (see Appendix III) and extensive personal communications with marketers and retailers in the Northeastern regional dry bean community, as well as peer-reviewed sources.

Table 1.2 Summary of priority traits in common bean for regional production in the Northeast United States, with potential genetic resources for deployment. Results are based on a literature review. Distinct stakeholders have differing priority traits, though some traits overlap between stakeholders.

Stakeholder	Priority Traits	Citation
Farmer	<ul style="list-style-type: none"> <li>• Stable and high yield</li> <li>• Resource use efficiency</li> <li>• Weed competitiveness</li> <li>• Disease/pest resistance</li> <li>• Appropriate maturity time</li> <li>• Growth habit/ability to direct harvest</li> </ul>	(Miklas et al. 2006; Renato Corte et al. 2002; Schwartz et al. 2010)
Consumer	<ul style="list-style-type: none"> <li>• Eating quality</li> <li>• Diverse culinary uses</li> <li>• Nutrition</li> <li>• Productivity/affordability</li> </ul>	(Cichy et al. 2019; Klaedtke et al. 2017; M. Santalla et al. 2005; Marta Santalla et al. 1999)
Plant Breeder/Seed Producer	<ul style="list-style-type: none"> <li>• Genetically diverse parent lines</li> <li>• Resistance to seedborne disease</li> </ul>	(B. Brouwer et al. 2016; Cichy et al. 2015; Rivera et al. 2018; S. Singh 2001)
Food System Sustainability	<ul style="list-style-type: none"> <li>• Resource-use efficiency</li> <li>• Optimization of agro-ecological niche partitioning</li> <li>• Tolerance to biotic and abiotic stress</li> </ul>	(Miklas et al. 2006; Wilker et al. 2019, Altieri 1996)

### **Importance of regional and local adaptation**

The genetic basis of common bean commercial cultivars is extremely narrow, representing less than 5% of available genetic diversity within the *Phaseolus* gene pool. This is widely recognized as a result of narrow crosses within elite market class gene pools, and poses a challenge to future improvement of the crop unless a concerted effort to introduce more

genetically diverse breeding materials is made (Kelly et al. 2012; S. Singh 2001). A persistent challenge to dry bean performance includes spatial and temporal yield stability, often attributable to abiotic or biotic stress across growing seasons and environments (S. Singh 1999). Ceccarelli (1994) defines yield stability “perceived by farmers as minimization of crop failures, [as] the most important socio-economic breeding objective in...[a] marginal environment.”

Most dry bean growers in New York state typically buy new seed each year of a recommended pure line cultivar from seed producers in more arid western regions. This system arose due to chronic problems with seedborne pathogen infection of bean seed grown and replanted in the Northeast. Phytosanitary inspection programs, furrow irrigation and plant breeding for improved disease resistance has resulted in the drastic reduction of seedborne disease in bean crops, an important success (Cafati and Saettler 1980; Darby 2016). However, adaptation to a specific farm environment is limited if seed is purchased off farm each season, as the process of adaptation requires annual shifts in allele frequencies within a population in direct response to the target environment.

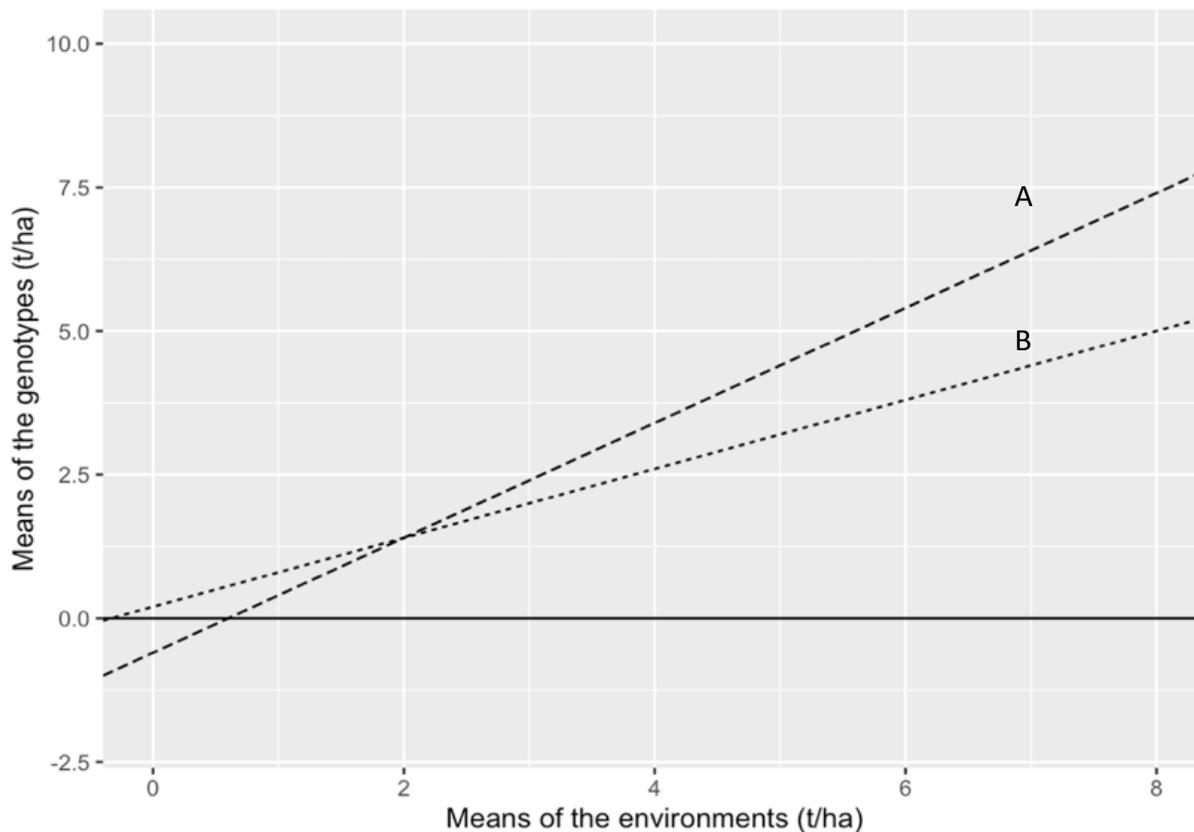
In studies of landrace wheat and dry bean varieties in Europe, measurable shifts in allele frequencies as well as observable phenotypic differences occurred as a result of cultivation in different environments in as few as three growing seasons (Klaedtke et al. 2017; Thomas et al. 2012). At the epigenetic level, studies show that the environment in which the mother plant is grown has measurable effects on seed traits such as dormancy and germination, and that maternal as well as embryonic genotype interact with these environmental effects (Penfield and MacGregor 2016; Donohue 2009). These maternal effects may serve to provide offspring with phenotypic adaptation to local environments, but this adaptation can only function if maternal and progeny environments match (Galloway 2005). Longer and more uneven germination time

in dry bean seed that was produced in western regions has been anecdotally reported by bean growers in the Northeast (Breen, personal communication, 2019). This indicates that once a regionally adapted, genetically diverse variety is developed, seed production in the target environment may facilitate the maintenance of desired traits. In order to grow dry bean seed in the Northeast while minimizing transmission of seedborne pathogens, a combined approach of resistance breeding, management practices, seed sterilization treatments and seed testing must be used. However, common bean seed production will not be the focus of this document.

G x E effect is composed of interaction between a given genotype and 1) predictable location effects, 2) unpredictable year effects and 3) the interaction between year and location, all of which comprise the environmental term of the interaction. Higher environmental heterogeneity results in increased G x E interaction as well as decreased yield stability (Vollmann and Menken 2011). Agroecological production systems are likely to exhibit higher environmental heterogeneity due to a lack of external inputs and an increased reliance on soil biology for crop nutrition (K. Murphy et al. 2005).

Breeding program design for agroecological production systems in a highly heterogeneous production region such as the Northeastern US can borrow from lessons of international development programs working with smallholder and subsistence farms in the Global South (Salvatore Ceccarelli 1994; J. Dawson and Goldringer 2012). These breeding programs observed a high incidence of crossover effect between research stations and farmer fields; that the best-performing cultivars under research station growing conditions were not the best performers in marginal growing conditions (Figure 1.1) (S. Ceccarelli 1989; Lawn 1988; Virk and Mangat 1991).

Figure 1.1 A visualization of cross-over Genotype X Environment (G X E) interaction. A and B are typical genotypes selected in high and low yielding environments, respectively. On either side of the crossover point, where lines A and B intersect, different genotypes will perform better, necessitating alignment of selection environment with the target environment (redrawn from Ceccarelli 1996).



Studies comparing performance of genotypes in contrasting production systems in the United States and Europe have yielded a multitude of results, indicating that the likelihood of crossover effect in non-target environments varies widely based on the crop and degree of divergence between environments (S. Ceccarelli 1989; Salvatore Ceccarelli 1994; Heilig and Kelly 2012; Messmer et al. 2011; K. M. Murphy et al. 2007). The presence of a crossover effect

would depend on the formula for correlated response to selection ( $CR_T$ ) in target environment T for a given trait A:

$$CR_T = i_S r_G \sqrt{H_S H_T} \sigma_{PT}$$

where  $CR_T$  is the correlated response in the target environment to selection in the breeding nursery,  $i_S$  is the selection intensity applied at the breeding nursery,  $r_G$  is the genotypic correlation between phenotype in selection and target environments,  $H_S$  and  $H_T$  are broad-sense heritabilities in the selection and target environments, respectively and  $\sigma_{PT}$  is the phenotypic standard deviation in the target environment (Atlin, Cooper, and Bjørnstad 2001). It can be seen that heritability of the given trait within both environments as well as the genetic correlation of the trait between the two environments will determine the efficiency of indirect selection and therefore performance in target environments.

Heritability will be lowest near the crossover point of performance between any two genotypes, whereas selection at either high or low ends of the stress level spectrum will benefit from the highest heritability. The position of the target environment on this spectrum relative to the crossover point can therefore be used to determine the optimal selection environment; if the target environment is more extreme than the point at which the crossover effect occurs, selection should be conducted in a high stress environment (Salvatore Ceccarelli 1996). Participatory plant breeding (PPB) programs seek to maximize response to selection by increasing  $r_G$ , while formal plant breeding (FPB) programs seek to maximize  $H_S$  (Atlin, Cooper, and Bjørnstad 2001).

In summary, mainstream breeding efforts have developed pure line cultivars and breeding lines with agronomic and disease resistance traits that will continue to be important genetic resources for many Northeastern dry bean growers. However, in order to ensure successful production in a changing climate and consumer market, as well as address needs of growers in

diverse environments and production systems in our region, a breeding program that integrates stakeholder priorities and provides an avenue for local adaptation is needed. Distinct breeding methods are needed to leverage these gains and can be seen as complementary to progress already made to improve broad adaptation by mainstream breeding programs (Cooper and Blyth 1996; J. Dawson and Goldringer 2012; Messmer et al. 2011). These methods will now be discussed.

### **Decentralized selection for local adaptation**

Public breeding programs have long been essential in furthering breeding work within crop types, locations and production systems that are not profitable enough to attract private industry investment, a role that is endangered as public cultivar development declines. Examples of crops susceptible to this “market failure” include cover crops, perennial crops and self-pollinated species with seed that can easily be saved, such as dry beans. With a decreasing number of public plant breeding programs across the country, continued release of dry bean varieties that perform well in non-target regions such as the Northeast may be the result of happenstance rather than intent (Shelton and Tracy 2017). Public breeding programs are often housed in land-grant universities with explicit missions of farmer education as well as well-developed extension networks for farmer outreach that can be utilized for participatory strategies and coalition building.

Participatory plant breeding (PPB) involves farmers as researchers and directly engages them in priority setting, screening germplasm, making crosses and selection, as well as seed increases and maintenance. Mass selection or pedigree selection methods can be used, and on-station screening is still likely to be required at least in early generations to select for traits such as disease resistance that do not always occur on farm every year. In program design, a smaller

number of “farmer experts” may be more effective at making accurate selections than a larger number of unexperienced farmers, and this method can still be successful in diffusing information and seed sources through farmer-to-farmer exchange.

PPB has been shown to be successful in many different crops and locations globally, with farmers effectively increasing yields through selection on their farms, at times more effectively than professional plant breeders (S. Ceccarelli et al. 2001; Kissing-Kucek 2017b; Sorrells et al. 2015). PPB has also been shown to increase farmer adoption and long-term retention of varieties compared to conventional methods, to enhance crop biodiversity including in-situ genetic conservation, and to reduce research costs (Sperling et al. 1993; Witcombe et al. 1996). Participatory breeding generally shows the most advantage over on-station breeding when farm and station environments differ dramatically, but in any environment, ensures that the diversity of traits required by farmers are considered, whether professional breeders are aware of them or not (J. C. Dawson, Murphy, and Jones 2008).

The efficiency of on-farm selection is often challenged by limited farmer time to conduct selection and high in-field variation, with the result that selection often is based on a limited set of traits or traits that may have low heritability. However, experimental design to ameliorate these challenges can increase efficiency, for example using gridded selection for post-harvest ear traits (Smith et al. 2001). Due to high variation across sites and years, more accurate selection can be achieved by combining results of multiple farmers and using adequate replication (Atlin, Cooper, and Bjørnstad 2001). According to the formula for correlated response to selection given above (Atlin, Cooper, and Bjørnstad 2001) as well as a constant  $H_T$  and  $\sigma_{PT}$  in a particular target environment, the correlated response to selection can be simplified as follows:

$$CR_T \propto i_S r_G \sqrt{H_S}$$

In choosing a method for selection in this context,  $i_S$  can be increased by increasing population size and intensity of selection. In PPB, this is most likely to be achieved through mass selection due to the labor intensity of pedigree selection at a large scale. Genotypic correlation  $r_G$  can be increased by selecting in an environment that is highly representative of the target environment; and  $H_S$  can be increased by increasing precision of evaluation or increasing genetic variance in population (Atlin, Cooper, and Bjørnstad 2001).

The importance of replication in PPB methods is underscored by an analysis of predicted selection response in replicated trials compared to unreplicated trials, defined as  $\sqrt{\frac{H_{replicated}}{H_{unreplicated}}}$ .

Given levels of genotype, genotype-environment interaction and residual variances typical of cereal crops in a temperate climate, magnitude of response would increase by approximately 25% in each selection cycle given two replicate sites and 40% given four replicate sites (Atlin, Cooper, and Bjørnstad 2001).

In inbred crops such as dry bean, the initial population can be formed by 1) a single cross; 2) a top-cross of three parents; 3) bulked crosses of more than 4 parents, or 4) a composite cross of more than 4 parents in any configuration (K. Murphy et al. 2005). This population would then be bulked on-station until sufficient seed was in hand to begin on-farm selection. In a farm-as-incomplete-block (FAIB) design, a network of farmers each receiving a subset of the larger population each season could conduct mass or pedigree selection, with sub-populations either rotated each season or bulked and redistributed (Atlin, Cooper, and Bjørnstad 2001).

Evolutionary breeding can be integrated into a participatory mass selection scheme, taking advantage of natural selection that is occurring alongside artificial selection at each farmer-site, though it requires larger population sizes of at least 10,000 plants (K. Murphy et al. 2005). In self-pollinated crops, once homozygosity is achieved through several generations of selection,

individual plants can eventually be selected as pure lines. At this stage, end-use quality of promising lines could be evaluated through collaboration with chefs or consumer tasting panels. Final selections could be bulked to form a blend that conforms sufficiently to market standards while still maintaining genetic diversity for future adaptation (J. C. Dawson, Murphy, and Jones 2008).

Increased genetic diversity within crop fields is especially important in organic and low-input systems due to increased G x E interaction. Genetic heterogeneity at the field scale allows for population buffering in response to stress rather than just single genotype response to environmental variation, defined as individual buffering. In allogamous crop species, population buffering can be accomplished by maintaining genetic variation within populations. In autogamous species such as common bean, multi-lines or mixtures can introduce some population buffering (Cooper and Blyth 1996). Heterogeneous populations are an important tool in achieving yield stability across seasons, which is defined by farmers as the most important crop trait but in modern breeding programs has been superseded by broad adaptation as the most important breeding goal (S. Ceccarelli et al. 2001; J. C. Dawson, Murphy, and Jones 2008)

The utility of varietal mixtures and multiline cultivars in managing field-scale disease is not new, but has been increasingly recognized as an important tool in improving agricultural sustainability in the face of climate change (Mundt 2002). In a bulk or evolutionary selection effort, a subset of superior pure lines, presumably possessing an array of relevant alleles for resistance due to initial inclusion of disease resistant parents, can be bulked to form a cultivar mixture with greater disease resistance as well as greater yield stability than each individual pure line would demonstrate in a monoculture (K. Murphy et al. 2005). The cultivation of such a mixed population over time would continue to enrich the proportion of resistance alleles,

assuming that pathogen pressure was present and susceptibility to the pathogen reduced seed yield of an individual (K. Murphy et al. 2005). The bulk population selection method culminating in a variety mixture is also likely to result in pure lines that excel in a mixed population; to select genotypes that, as Allard and Adams (1969) noted, are “good competitors as well as good neighbors.”

Importantly, a decentralized and participatory plant breeding approach is also effective in generating and maintaining genetic diversity of the crop on a regional scale, increasing the likelihood of in-situ stewardship of genetic resources by farmers over the long term (Galluzzi et al. 2010; Thomas et al. 2015).

### **Identification of parental materials**

The public bean breeding community has done an outstanding job in evaluating, developing and exchanging materials for genetic improvement of common bean through collaborative diversity panels, regional trialing and breeding work. Germplasm for regional food systems can be assessed and stewarded by: 1) trialing of modern cultivars and breeding lines, regionally adapted heirlooms and landraces to determine suitability; 2) breeding using regionally important germplasm and selection for agronomic and quality traits in target environments; and 3) selection and genetic maintenance within existing germplasm that already meets grower needs. Parental selection should consider potential performance in the target agri-food system as well as the introduction and maintenance of sufficient genetic diversity into the program for continued genetic improvement. Professional plant breeders are often best positioned to increase genetic variance of the population by sourcing diverse breeding materials in consultation with farmers and end users (Atlin, Cooper, and Bjørnstad 2001). However, the direct use of unadapted parents should be minimized in the process of population development in PPB due to

the need for reasonable performance success in a commercial farm context (J. C. Dawson, Murphy, and Jones 2008). In addition to improved cultivars or breeding lines likely to possess broad adaptation, local or regional landraces can contribute specific adaptation to the population as well as high end-use and quality traits that are difficult for farmers to select for in the field (K. Murphy et al. 2005).

For quantitative traits such as yield and end-use quality, progeny phenotypes will likely fall along mid-parent means, so as many parents as possible should be selected with high values for these traits. At least one parent with a given desired resistance allele, or, in the case of some pathogen-plant systems, lack of a susceptibility allele, for pathogens present in the target environment also must be included, taking into account any diversity in pathovar populations. With persistent natural field pathogen pressure, resistance alleles should quickly be enriched through selection. In other cases, early generation selection for disease resistance using molecular markers or artificially inoculated trials can be used before populations are bulked and shared with farmers (J. C. Dawson, Murphy, and Jones 2008; K. Murphy et al. 2005).

### ***Conclusion***

Plant breeding is one of many tools in the toolbox for support of a regional food system that provides collectively held goals including 1) diverse, profitable agricultural enterprises 2) stewardship of land and water, and 3) healthy communities with high quality of life. Dry beans are a healthful, high-protein crop well-suited for diverse agroecological production systems in our region. Mainstream plant breeding for the United States has developed improved dry bean varieties that excel in high-input mechanized production systems designed to meet canning industry needs. These genotypes also excel on many farms (and in many growing seasons) in

heterogeneous growing regions such as the Northeast, but in order to meet the collectively held goals outlined above, particularly in the face of changing climate, a new approach to design and implementation of dry bean breeding is needed in the Northeast. This new model would employ a decentralized approach to setting breeding program priorities, evaluating potential parent germplasm, and selecting from diverse populations for performance on farms as well as in the marketplace and on the consumer's plate.

## REFERENCES

- Altieri, Miguel. 1996. *Agroecology: The Science of Sustainable Agriculture* (Second Edition). CRC Press.
- AgMRC. 2017. "Dry Edible Bean Profile | Agricultural Marketing Resource Center." Retrieved January 14, 2019 (<https://www.agmrc.org/commodities-products/grains-oilseeds/dry-edible-bean-profile>).
- Assefa, Teshale, A. Assibi Mahama, Anne V. Brown, Ethalinda K. S. Cannon, Jean Claude Rubyogo, Idupulapati M. Rao, Matthew W. Blair, and Steven B. Cannon. 2019. "A Review of Breeding Objectives, Genomic Resources, and Marker-Assisted Methods in Common Bean (*Phaseolus Vulgaris* L.)." *Molecular Breeding*.
- Atlin, G. N., M. Cooper, and Å. Bjørnstad. 2001. "A Comparison of Formal and Participatory Breeding Approaches Using Selection Theory." *Euphytica* 122: 463–75.
- Bashi, Zafer, Ryan Mccullough, Liane Ong, and Miguel Ramirez. 2019. *Alternative Proteins: The Race for Market Share Is On*.
- Berti, Giaime and Catherine Mulligan. 2016. "Competitiveness of Small Farms and Innovative Food Supply Chains: The Role of Food Hubs in Creating Sustainable Regional and Local Food Systems." *Sustainability* 8(7):616.
- Blair, Matthew W. and Sandra M. Lorigados. 2016. "Diversity of Common Bean Landraces, Breeding Lines, and Varieties from Cuba." *Crop Science* 56(1):322.
- Bliss, F. A. 1993. "Breeding Common Bean for Improved Biological Nitrogen Fixation". *Developments in Plant and Soil Sciences* 152.
- Brouwer, Brook O., Kevin M. Murphy, and Stephen S. Jones. 2015. "Plant Breeding for Local Food Systems: A Contextual Review of End-Use Selection for Small Grains and Dry Beans in Western Washington."
- Brouwer, Brook, Louisa Winkler, Kelly Atterberry, Stephen Jones, and Carol Miles. 2016. "Exploring the Role of Local Heirloom Germplasm in Expanding Western Washington Dry Bean Production." *Agroecology and Sustainable Food Systems* 40(4):319–32.
- Bustos-Korts, Daniela, Ignacio Romagosa, Gisela Borràs-Gelonch, Ana Maria Casas, Gustavo A. Slafer, and Fred van Eeuwijk. 2018. "Genotype by Environment Interaction and Adaptation." Pp. 1–44 in *Encyclopedia of Sustainability Science and Technology*. New York: Springer.
- Cafati, C. R. and A. W. Saettler. 1980. "Transmission of *Xanthomonas Phaseoli* in Seed of Resistant and Susceptible *Phaseolus* Genotypes". In *Ecology and Epidemiology*.

- Campa, Ana, Ester Murube, and Juan José Ferreira. 2018. "Genetic Diversity, Population Structure, and Linkage Disequilibrium in a Spanish Common Bean Diversity Panel Revealed through Genotyping-by-Sequencing." *Genes* 9(11).
- Ceccarelli, S. 1989. "Wide Adaptation: How Wide?" *Euphytica* 40(3):197–205.
- Ceccarelli, S., S. Grando, E. Bailey, A. Amri, M. El-Felah, F. Nassif, S. Rezgui, and A. Yahyaoui. 2001. "Farmer Participation in Barley Breeding in Syria, Morocco and Tunisia." *Euphytica* 122: 521–36.
- Ceccarelli, Salvatore. 1994. "Specific Adaptation and Breeding for Marginal Conditions." *Euphytica* 77(3):205–19.
- Ceccarelli, Salvatore. 1996. "Adaptation to Low/High Input Cultivation." *Euphytica* 92(1–2):203–14.
- Chávez-Servia, José Luis, Elena Heredia-García, Netzahualcoyotl Mayek-Pérez, Elia N. Aquino-Bolaños, Sanjuana Hernández-Delgado, José C. Carrillo-Rodríguez, Homar R. Gill-Langarica, and Araceli M. Vera-Guzmán. 2016. "Diversity of Common Bean (*Phaseolus Vulgaris* L.) Landraces and the Nutritional Value of Their Grains." in *Grain Legumes*.
- Cichy, Karen A., Timothy G. Porch, James S. Beaver, Perry Cregan, Deidre Fourie, Raymond P. Glahn, Michael A. Grusak, Kelvin Kamfwa, Dennis N. Katuuramu, Phil McClean, Eninka Mndolwa, Susan Nchimbi-Msolla, Marcial A. Pastor-Corrales, and Phillip N. Miklas. 2015. "A *Phaseolus Vulgaris* Diversity Panel for Andean Bean Improvement." *Crop Science* 55(5):2149–60.
- Cichy, Karen A., Jason A. Wiesinger, Matthew Berry, Susan Nchimbi-Msolla, Deidre Fourie, Timothy G. Porch, Daniel Ambechew, and Phillip N. Miklas. 2019. "The Role of Genotype and Production Environment in Determining the Cooking Time of Dry Beans (*Phaseolus Vulgaris*)." *Legume Science* 1(1).
- Clancy, Kate and Kathryn Ruhf. 2010. "Is Local Enough? Some Arguments for Regional Food Systems." *Choices*.
- Cooper, M. and D. E. Blyth. 1996. "Understanding Plant Adaptation to Achieve Systematic Applied Crop Improvement - A Fundamental Challenge" edited by M. Cooper and G. L. Hammer. *Plant Adaptation and Crop Improvement* 636.
- Crowder, David W. and John P. Reganold. 2015. "Financial Competitiveness of Organic Agriculture on a Global Scale." *Proceedings of the National Academy of Sciences of the United States of America* 112(24):7611–16.
- Darby, Heather. 2016. *Northeast Dry Bean Production Guide*. UVM Northwest Crops and Soils Program.

- Darby, Heather and Erica Cummings. 2017. *2016 Heirloom Dry Bean Variety Trial*. UVM Northwest Crops and Soils Program.
- Darby, Heather and Erica Cummings. 2018. *2017 Heirloom Dry Bean Variety Trial*. UVM Northwest Crops and Soils Program.
- Darby, Heather, Erica Cummings, Lily Calderwood, Abha Gupta, Julian Post, and Sara Ziegler. 2016. *2015 Heirloom Dry Bean Variety Trial*. UVM Northwest Crops and Soils Program.
- Dawson, Julie C., Kevin M. Murphy, and Stephen S. Jones. 2008. "Decentralized Selection and Participatory Approaches in Plant Breeding for Low-Input Systems." *Euphytica* 160(2):143–54.
- Dawson, Julie C., Pierre Rivière, Jean-François Berthelot, Florent Mercier, Patrick de Kochko, Nathalie Galic, Sophie Pin, Estelle Serpolay, Mathieu Thomas, Simon Giuliano, Isabelle Goldringer, Julie C. Dawson, Pierre Rivière, Jean-François Berthelot, Florent Mercier, Patrick de Kochko, Nathalie Galic, Sophie Pin, Estelle Serpolay, Mathieu Thomas, Simon Giuliano, and Isabelle Goldringer. 2011. "Collaborative Plant Breeding for Organic Agricultural Systems in Developed Countries." *Sustainability* 3(8):1206–23.
- Dawson, Julie and Isabelle Goldringer. 2012. "Breeding for Genetically Diverse Populations: Variety Mixtures and Evolutionary Populations." Pp. 77–98 in *Organic Crop Breeding*, edited by E. Lammerts van Bueren and J. Myers. John Wiley & Sons.
- Döring, Thomas F., Samuel Knapp, Geza Kovacs, Kevin Murphy, and Martin S. Wolfe. 2011. "Evolutionary Plant Breeding in Cereals-Into a New Era." *sustainability* 3(10):1944-1971.
- Galluzzi, Gea, Pablo Eyzaguirre, Valeria Negri, G Galluzzi, and P Eyzaguirre. 2010. "Home Gardens: Neglected Hotspots of Agro-Biodiversity and Cultural Diversity." *Biodivers Conserv* 19:3635–54.
- Gepts, Paul. 1998. "Origin and Evolution of Common Bean: Past Events and Recent Trends." *HortScience* 33(7):1124–30.
- Harlan, J. R. 1975. "Our Vanishing Genetic Resources." *Science* 188(4188):617–21.
- Heilig, James A. and James D. Kelly. 2012. "Performance of Dry Bean Genotypes Grown under Organic and Conventional Production Systems in Michigan." *Agronomy* 104(5):1485.
- Hoppe, Robert, James MacDonald, and Penni Korb. 2010. *Small Farms in the United States: Persistence Under Pressure*.
- Jordan, Jennifer A. 2007. "The Heirloom Tomato as Cultural Object: Investigating Taste and Space." *Sociologia Ruralis* 47(1):20–41.

- Keller, Beat, Daniel Ariza-Suarez, Juan de la Hoz, Johan Steven Aparicio, Ana Elisabeth Portilla-Benavides, Hector Fabio Buendia, Victor Manuel Mayor, Bruno Studer, and Bodo Raatz. 2020. "Genomic Prediction of Agronomic Traits in Common Bean (*Phaseolus Vulgaris* L.) Under Environmental Stress." *Frontiers in Plant Science* 11:1001.
- Kelly, James D. 2010. "The Story of Bean Breeding: Dry Bean Production and Breeding Research in the U.S".  
([www.canr.msu.edu/beanbreeding/\\_pdf/Story\\_of\\_Bean\\_Breeding\\_in\\_the\\_US.pdf](http://www.canr.msu.edu/beanbreeding/_pdf/Story_of_Bean_Breeding_in_the_US.pdf))
- Kelly, James D. 2001. "Remaking bean plant architecture for efficient production". *Advances in Agronomy* 71: 109-143.
- Kelly, James D., Karen A. Cichy, Gregory V. Varner, and Evan M. Wright. 2012. "Registration of 'Snowdon' White Kidney Bean." *Journal of Plant Registrations*.
- Kissing-Kucek, Lisa. 2017a. *Final Report for GNE15-107: Participatory Breeding of High-Value Wheat for the Northeast*. Northeast Sustainable Agriculture Research and Education.
- Kissing-Kucek, Lisa. 2017b. "Participatory Breeding of Wheat for Organic Production" Cornell University.
- Klaedtke, Stephanie, Leonardo Caproni, Julia Klauck, Paul de la Grandville, Martin Dutartre, Pierre Stassart, Véronique Chable, Valeria Negri, Lorenzo Raggi, Stephanie M. Klaedtke, Leonardo Caproni, Julia Klauck, Paul De la Grandville, Martin Dutartre, Pierre M. Stassart, Véronique Chable, Valeria Negri, and Lorenzo Raggi. 2017. "Short-Term Local Adaptation of Historical Common Bean (*Phaseolus Vulgaris* L.) Varieties and Implications for In Situ Management of Bean Diversity." *International Journal of Molecular Sciences* 18(3):493.
- Lammerts van Bueren, E. T. 2012. "Organic Crop Breeding: Integrating Organic Agricultural Approaches and Traditional and Modern Plant Breeding Methods." Pp. 3–13 in *Organic Crop Breeding*, edited by E.T. Lammerts van Bueren and J.R. Myers. Oxford, UK: Wiley-Blackwell.
- Lawn, R. J. 1988. "Breeding for Improved Plant Performance in Drought-Prone Environments." Pp. 213–19 in *Drought research priorities for the dryland tropics*, edited by E. . Bidinger and C. Johansen.
- de Leon, Natalia, Jean-Luc Jannink, Jode W. Edwards, and Shawn M. Kaeppler. 2016. "Introduction to a Special Issue on Genotype by Environment Interaction." *Crop Science* 56(5):2081–89.
- Loria, Kristen, Michael Mazourek, and Gregory Inzinna. 2019. *2019 Dry Bean Variety Trial Report*. eOrganic Variety Trial Database. ([www.varietytrials.eorganic.info](http://www.varietytrials.eorganic.info)).
- Lyon, Alexandra, William Tracy, Micaela Colley, Patrick Culbert, Michael Mazourek, James Myers, Jared Zystro, and Erin M. Silva. 2019. "Adaptability Analysis in a Participatory

- Variety Trial of Organic Vegetable Crops.” *Renewable Agriculture and Food Systems* 1–17.
- Messmer, Monika, Isabell Hildermann, Kristian Thorup-Kristensen, and Zed Rengel. 2011. “Nutrient Management in Organic Farming and Consequences for Direct and Indirect Selection Strategies.” Pp. 15–38 in *Organic Crop Breeding*, edited by E.T. Lammerts van Bueren and J.R. Myers. Oxford, UK: Wiley-Blackwell.
- Michaels, Thomas, Craig Sheaffer, Hannah Swegarden, and Claire Flavin. 2013. “Organic Dry Bean Production Systems and Cultivar Choices.” eOrganic.
- Miklas, Phillip, James D. Kelly, Steve E. Beebe, and Matthew W. Blair. 2006. “Common Bean Breeding for Resistance against Biotic and Abiotic Stresses: From Classical to MAS Breeding.” *Euphytica* 147:105–31.
- Mundt, C. C. 2002. “Use of Multiline Cultivars and Mixtures for Disease Management.” *Annu. Rev. Phytopathol* 40:381–410.
- Murphy, Kevin, Doug Lammer, Steve Lyon, Brady Carter, and Stephen S. Jones. 2005. “Breeding for Organic and Low-Input Farming Systems: An Evolutionary–Participatory Breeding Method for Inbred Cereal Grains.” *Renewable Agriculture and Food Systems* 20(1):48–55.
- Murphy, Kevin M., Kimberly G. Campbell, Steven R. Lyon, and Stephen S. Jones. 2007. “Evidence of Varietal Adaptation to Organic Farming Systems.” *Field Crops Research* 102(3):172–77.
- Myers, Robert. 2002. *DRY EDIBLE BEANS*. Jefferson Institute Alternative Crops Guide. ([www.extension.iastate.edu/alternativeag/cropproduction/pdf/drybean\\_crop\\_guide.pdf](http://www.extension.iastate.edu/alternativeag/cropproduction/pdf/drybean_crop_guide.pdf))
- Parker, Travis A., Antonia Palkovic, and Paul Gepts. 2020. “Determining the Genetic Control of Common Bean Early-Growth Rate Using Unmanned Aerial Vehicles.” *remote sensing* 12(11): 1748.
- Ponisio, L. C., L. K. M’Gonigle, K. C. Mace, J. Palomino, P. de Valpine, and C. Kremen. 2014. “Diversification Practices Reduce Organic to Conventional Yield Gap.” *Proceedings of the Royal Society: Biological Sciences* 282(1799): 1396–1396.
- Posa-Macalincag, Maria-Carmela T., George L. Hosfield, Kenneth F. Grafton, Mark A. Uebersax, and James D. Kelly. 2002. *Quantitative Trait Loci (QTL) Analysis of Canning Quality Traits in Kidney Bean (Phaseolus Vulgaris L.)*. Vol. 127.
- Renato Corte, Hercules, Magno Antonio Patto Ramalhol, Flávia Maria Avelar Gonçalves, and Ângela de Fátima Barbosa Abreu. 2002. *Natural Selection for Grain Yield in Dry Bean Populations Bred by the Bulk Method*. *Euphytica* 123: 387-393.
- Rivera, Ana, Marçal Plans, Josep Sabaté, Francesc Casañas, Joan Casals, Aurora Rull, and Joan

- Simó. 2018. "The Spanish Core Collection of Common Beans (*Phaseolus Vulgaris* L.): An Important Source of Variability for Breeding Chemical Composition." *Frontiers in Plant Science* 9:1642.
- SARE. n.d. "About Northeast SARE." *Northeast Sustainable Agriculture and Education*. Retrieved December 7, 2020 (<https://northeast.sare.org/about/>).
- Santalla, M., M. Lema, A. P. Rodiño, A. M. González, A. B. Monteagudo, and A. M. De Ron. 2005. "Improvement of Large-Seeded Common Bean Cultivars under Sustainable Cropping Systems in Spain." *Euphytica* 142(1–2):85–95.
- Santalla, Marta, Miguel Angelo Fueyo, A. Paula Rodino, Inès Montero, and Antonio M. De Ron. 1999. "Breeding for Culinary and Nutritional Quality of Common Bean (*Phaseolus Vulgaris* L.) in Intercropping Systems with Maize (*Zea Mays* L.)." *Biotechnol. Agron. Soc. Environ* 3(4):225–29.
- Schwartz, Howard F., Shree P. Singh, and Howard F. Schwartz. 2010. "Review : Breeding Common Bean for Resistance to Insect Pests and Nematodes Review : Breeding Common Bean for Resistance to Insect Pests and Nematodes." *Canadian Journal of Plant Science* 91(2).
- Seyfang, Gill. 2006. "Ecological Citizenship and Sustainable Consumption: Examining Local Organic Food Networks." *Journal of Rural Studies* 22(4):383–95.
- Shelton, Adrienne C. and William F. Tracy. 2017. "Cultivar Development in the U.S. Public Sector." *Crop Science* 57(4):1823–35.
- Silbernagel, M. J. and R. M. Hannan. 1988. "Utilization of Genetic Resources in the Development of Commercial Bean Cultivars in the U.S.A." Pp. 561-596 in *Genetic Resources of Phaseolus Beans*, edited by P. Gepts. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Singh, Shree. 1999. *Common Bean Improvement in the Twenty-First Century*. Dordrecht, The Netherlands: Springer.
- Singh, Shree. 2001. "Broadening the Genetic Base of Common Bean Cultivars." *Crop Science* 41(6):1659.
- Smith, Margaret E., Fernando Castillo G., and Francisco Gómez. 2001. "Participatory Plant Breeding with Maize in Mexico and Honduras." *Euphytica* 122: 551–63.
- Sorrells, Mark, Heather Darby, Julie C. Dawson, Phil Atkins, Michael Davis, Elizabeth Dyck, and Ellen Mallory. 2015. *Farm-Based Selection and Seed Production of Varieties of Bread Wheat, Spelt, Emmer, and Einkorn Adapted to Organic Systems in the Northeast*. Northeast Sustainable Agriculture Research and Education.

- Sperling, Louise, Michael E. Loevinsohn, and Beatrice Ntabomvura. 1993. "Rethinking the Farmer's Role in Plant Breeding: Local Bean Experts and on-Station Selection in Rwanda." *Experimental Agriculture* 29(4):509–19.
- Swegarden, Hannah R., Craig C. Scheaffer, and Thomas. Michaels. 2016. "Yield Stability of Heirloom Dry Bean (*Phaseolus Vulgaris* L.) Cultivars in Midwest Organic Production." *HortScience* 51(1):8–14.
- Swegarden, Hannah R, Craig C. Sheaffer, and Thomas E. Michaels. 2016. "Yield Stability of Heirloom Dry Bean (*Phaseolus Vulgaris* L.) Cultivars in Midwest Organic Production". *HortScience* 51(1):8-14.
- Swegarden, Hannah Rae. 2015. "Selection of Commercial and Heirloom Common Bean (*Phaseolus Vulgaris* L.) for Organic Production in Minnesota." University of Minnesota.
- Thomas, Mathieu, Elise Demeulenaere, Julie C. Dawson, Abdul Rehman Khan, Nathalie Galic, Sophie Jouanne-Pin, Carine Remoue, Christophe Bonneuil, and Isabelle Goldringer. 2012. "On-Farm Dynamic Management of Genetic Diversity: The Impact of Seed Diffusions and Seed Saving Practices on a Population-Variety of Bread Wheat." *Evolutionary Applications* 5(8):779–95.
- USDA-AMS. 2019. "Bean Market News 2019 Summary USDA Agricultural Marketing Service."
- USDA-NASS. 2017. *2017 Census of Agriculture-NY*.
- USDA-NASS. n.d. "Dry Beans: Planted Acreage by Year, US and Major States." Retrieved March 6, 2019 ([https://www.nass.usda.gov/Charts\\_and\\_Maps/Dry\\_Beans,\\_Dry\\_Peas,\\_and\\_Lentils/dbinten.php](https://www.nass.usda.gov/Charts_and_Maps/Dry_Beans,_Dry_Peas,_and_Lentils/dbinten.php)).
- USDA NASS. 2019. "2019 Organic Survey."
- Villa, Tania Carolina Camacho, Nigel Maxted, Maria Scholten, and Brian Ford-Lloyd. 2005. "Defining and Identifying Crop Landraces." *Plant Genetic Resources* 3(3):373–84.
- Virk, D. S. and B. K. Mangat. 1991. "Detection of Cross over Genotype × Environment Interactions in Pearl Millet." *Euphytica* 52(3):193–99.
- Vollmann, Johann and Michelle Menken. 2011. "Soybean: Breeding for Organic Farming Systems." Pp. 203–14 in *Organic Crop Breeding*. Oxford, UK: Wiley-Blackwell.
- Watson, Amy, Sreya Ghosh, Matthew J. Williams, William S. Cuddy, James Simmonds, María Dolores Rey, M. Asyraf Md Hatta, Alison Hinchliffe, Andrew Steed, Daniel Reynolds, Nikolai M. Adamski, Andy Breakspear, Andrey Korolev, Tracey Rayner, Laura E. Dixon, Adnan Riaz, William Martin, Merrill Ryan, David Edwards, Jacqueline Batley, Harsh

Raman, Jeremy Carter, Christian Rogers, Claire Domoney, Graham Moore, Wendy Harwood, Paul Nicholson, Mark J. Dieters, Ian H. Delacy, Ji Zhou, Cristobal Uauy, Scott A. Boden, Robert F. Park, Brande B. H. Wulff, and Lee T. Hickey. 2018. "Speed Breeding Is a Powerful Tool to Accelerate Crop Research and Breeding." *Nature Plants* 4(1):23–29.

Wilker, Jennifer, Alireza Navabi, Istvan Rajcan, Frédéric Marsolais, Brett Hill, Davoud Torkamaneh, and K. Peter Pauls. 2019. "Agronomic Performance and Nitrogen Fixation of Heirloom and Conventional Dry Bean Varieties Under Low-Nitrogen Field Conditions." *Frontiers in Plant Science* 10.

Witcombe, J. R., A. Joshi, K. D. Joshi, and B. R. Sthapit. 1996. "Farmer Participatory Crop Improvement. I. Varietal Selection and Breeding Methods and Their Impact on Biodiversity." *Experimental Agriculture* 32(4):445–60.

## CHAPTER II

### IMPROVEMENT OF DISEASE RESISTANCE IN SPECIALTY DRY BEAN MARKET CLASSES FOR REGIONAL MARKETS

#### ***Significance of Seedborne Pathogens in Common Bean***

Bean growers across the world have traditionally saved and replanted their seed year after year, and in many countries the majority of growers still do. This practice reduces input costs and facilitates specific adaptation of the crop, especially if the farmer practices selection or culling from season to season (Döring et al. 2011; Thomas et al. 2012). However, this also allows for the accumulation of seed borne pathogens in the population if susceptible cultivars are grown. With roughly 50% of bean pathogens being seed borne, this is a persistent problem for bean growers, especially in more humid regions such as the Northeastern United States (Ellis and Gálvez E. 1980).

The solution to this problem in the United States has been to use resistant varieties and buy certified seed from the arid west, where disease pressures are lower and inspection-based certification programs exist. However, even in arid areas disease resistance is important to production of pathogen-free seed, and recent increased use of overhead irrigation has led to a resurgence of bacterial pathogens in seed crops (Wohleb and du Toit 2011). As certified seed of heirloom or other specialty varieties is not available, growers in the Northeast must save their own seed or buy from uncertified commercial sources. As a result of these practices, many have observed endemic disease issues such *Bean common mosaic virus*, halo blight and bacterial blight (Darby 2016; Baer, personal communication). Due to its large seed size, hot water treatments that are effective for smaller seeds are ineffective for common bean. Steam treatment

has promise as a viable alternative, but results have been inconclusive and more research is needed (Darby et al. 2019).

Improved disease resistance to the major seedborne pathogens in varieties and populations outside of commodity market classes would contribute to regional staple crop production systems by improving grower profitability, reducing pesticide use, and improving access to high-quality, pathogen-free seed. There are other seed borne pathogens of significance in the Northeast (and the west, where much of our bean seed is grown) in addition to the three pathogens discussed in this chapter; *Colletotrichum lindemuthianum*, *Xanthomonas campestris* and *Bean common mosaic virus*. These include *Pseudomonas syringae pv. syringae*, causal organism of bacterial brown spot, and *Pseudomonas syringae pv. phaseolicola*, causal organism of halo blight. *Curtobacterium flaccumfaciens pv. flaccumfaciens*, causal organism of bean bacterial wilt, is also a significant pathogen in western states, though reports in the Northeast are less common. The latter pathogens will not be discussed further as resistance breeding for them was outside the scope of current work.

### ***Plant-pathogen interactions in the incidence of seed borne disease***

Interaction between seeds and pathogens can manifest by way of seed infection, in which the pathogen is present within seed tissue, or infestation, in which the pathogen is present on the surface of the seed only. Seed infection requires certain seed tissues to be infected with the pathogen; which tissues can be infected varies among different pathogens. Transmission of viruses requires infection of embryo tissue, while fungi and bacteria do not require embryo infection to be transmitted to progeny, and often are transmitted through infestation of the pathogen on the seed surface (Agarwal and Sinclair 1997).

*Bean common mosaic virus* and *X. campestris pv. phaseoli* can both infect all seed tissues including the embryo. Despite this, Agarwal and Sinclair (1997) clarify that “successful seed infection is the exception and not the rule”, and plants can commonly be infected without seed being infected. However, inoculum is frequently transmitted on the surface of the seed, especially in grain and pulse crops where threshing of dry plant material is the mode of harvest (C. Smart, personal communication). Virus to seed transmission happens via systemic infection of the plant during early stages of development, resulting in infected ovular tissue, or via infected pollen into the developing ovule. The bacterium *X. campestris* can be transmitted via foliar infection or infection of pods. In general, severity of mother plant infection is highly correlated to seed transmission, a condition which is more likely when infected seed is planted (Agarwal and Sinclair 1997).

### ***Selected Seedborne Pathogens of Common Bean in the Northeast***

#### **Common Bacterial Blight**

Common bacterial blight in bean is caused by the pathogenic organisms *X. campestris pv. phaseoli* (*Xcp*) and *X. axonopodis pv. fuscans* (O’Leary 2018). The two pathovars differ in their lab culture phenotype but in the field cause indistinguishable symptoms and often co-occur in the same field, on the same plants (S. Singh and Miklas 2015). Multiple strains of *X. campestris pv. phaseoli* exist, with *Xcp25* being the most common strain in the United States, but another strain *Px100* was also found to be widespread. Differential pathogenicity was found between species and strains, with *Xcp25* and *Px100* being among the most virulent. (Duncan et al. 2011; Viteri et al. 2014).

Strategies for management of common bacterial blight include crop rotation, cultural practices, sowing pathogen-free seed and, for non-organic production, use of bactericides (Viteri et al. 2014). All aerial parts of the plant can be infected though symptoms primarily appear on leaves. Infected seed is not necessarily identifiable visually but pod symptoms may indicate seed infection. Characteristic symptoms include small water-soaked spots on the underside of leaves, becoming irregularly shaped necrotic spots with a lemon-yellow border, often starting at the leaf margin where bacteria can more easily enter plant tissue via hydathodes. Plants grown from infected seed often have lesions on cotyledons, nodes and primary leaves (S. Singh and Miklas 2015). Bacteria proliferate asymptotically on leaves in conditions of warm temperatures and high humidity, and attack plant tissues when a threshold population is reached (O'Leary 2009). The disease is favored by temperatures between 20°C and 30°C, and does not proliferate in temperatures below 16°C (S. Singh and Miklas 2015).

Seed yield losses of up to 45% due to infection have been reported (S. Singh and Miklas 2015). Intermediate to strongly resistant cultivars tend to slow the increase of bacterial populations on leaves, also delaying or preventing eventual vascular spread of the disease, thereby reducing chances of seed infection. Planting of infected seed does not necessarily result in systemic transmission of bacteria from vascular tissue to growing seeds. *X. campestris* can survive epiphytically on many non-host species including soybean and corn (S. Singh and Miklas 2015).

Seed infection or infestation is the most common means of bacterial blight epidemics, serving as the primary inoculum, and secondary inoculum is transported by rain, water and agricultural implements (O'Boyle et al. 2007). Bacteria do not survive in soil, and this is not known to be a means of pathogen transmission, though *X. campestris* can survive on plant

residue in the field (O’Leary 2009). Some copper-based pesticides are organically approved and can be used as a preventative measure to stop or reduce occurrence of disease, but genetic resistance and cultural practices are the primary means of control. However, the primary cultural strategy to prevent disease has been to produce new seed each year in regions of the country where arid conditions reduce pathogen proliferation is less favored (Wohleb and du Toit 2011).

Genetic resistance to common blight is characterized by quantitative inheritance of mostly additive effects, and heritability of bacterial blight resistance is considered low to moderate. However, some gene-for-gene inheritance of bacterial blight resistance has been identified in *P. acutifolius* (S. Singh and Miklas 2015). Differential response to resistance genes occurs across varying environments, genetic backgrounds and different plant organs (O’Boyle, Kelly, and Kirk 2007). An interaction between resistance locus and pathogen strain has also been observed (S. Singh and Miklas 2015). These factors make pyramiding of multiple resistance loci the most effective strategy for development of highly resistant cultivars. Currently there are no breeding lines or cultivars available with high levels of resistance to aggressive bacterial strains within plant pods, indicating a need for future research (S. Singh and Miklas 2015).

Low levels of resistance have been identified in the Mesoamerican common bean gene pool, including the *SAP6* QTL which derives from a Great Northern landrace. As genetic resistance to bacterial blight is low to moderate in *P. vulgaris*, resistance genes from *P. coccineus* and *P. acutifolius* have been introgressed (R. W. Duncan et al. 2011). The *SU91* locus is one of these genes, originating in *P. acutifolius* (O’Leary 2009).

More than 20 major and minor effect QTL have been reported to affect common blight resistance. Different QTL confer resistance in different plant parts, and against different bacterial strains. For example, the *P. acutifolius*-derived QTL such as *SU91* do not confer resistance in

bean pods and the *SAP6* QTL confers resistance in leaves and pods to less aggressive strains only (Viteri et al. 2014). There are six identified linked markers available for common bacterial blight resistance. The *SAP6* QTL derives from a Great Northern landrace, a Mesoamerican type. The *SU91* and *BC420* QTL derive from *P. acutifolius*, or tepary bean. Two large-seeded Andean breeding lines with moderate resistance to *X. campestris* are ‘USDK-CBB-15’ and ‘USCR-CBB-20,’ which derive their resistance from the *SU91* and *SAP6* loci described above (Miklas et al. 2011; Miklas et al. 2006). There is evidence for an epistatic relationship between the *BC420* and *SU91* loci, with no resistance evident in genotypes possessing the *BC420* locus alone. Other potential germplasm for improvement of resistance in the Andean gene pool includes two cranberry lines bred by Dr. Tim Porch of the USDA by backcrossing USCR-CBB-20 to the Mesoamerican resistant line ‘VAX 6’, which has shown higher resistance to *X. campestris* than USDK-CBB-15 and USCR-CBB-20. The additional locus that ‘VAX 6’ provides is likely *Xa11.4* (Viteri et al. 2014).

In a comparison of direct disease selection (DDS) and marker-assisted selection (MAS) methods in breeding for resistance by Duncan et al. (2012), more resistant lines were obtained through the DDS pipeline compared to MAS (12 vs 6 lines). The *SU91*, *SAP6* and *BC420* SCAR markers were used. Mean disease severity index was 3.3 for DDS and 4.2 for MAS in a severe infection environment in the greenhouse. Under moderate field pressure, MAS line performance was not significantly different from DDS line performance. *SU91* contributed more significant resistance, though most resistant lines had all three markers. Some lines with all three markers had a susceptible phenotype. A study by Yu et al. (2000), in contrast, found MAS to be 33% less expensive than DDS and 94.2% accurate using the *BC420* locus. These results indicate the value of a combined approach to breeding, using MAS in the early stages to screen large numbers of

progeny before doing DDS (Duncan et al. 2012; Singh and Miklas 2015). This is also necessary to maintain minor resistance QTL (P. Miklas et al. 2006).

Due to the quantitative nature of pathogen resistance, a combination of gamete selection and recurrent selection followed by self-pollination could be useful to pyramid multiple genes from diverse sources. The use of four or more parents in three-way or modified double crossing schemes may be necessary to pyramid loci. Generation of large numbers of F<sub>1</sub> seed from multiple-parent crosses is likely to be more effective than a large number of single crosses and backcrosses, and will ultimately produce highly resistant lines in a shorter period of time (S. Singh 2001). Breeding lines possessing pyramided resistance loci can subsequently be crossed with elite parent cultivars to recover seed characteristics and traits essential for commercial production without losing their common bacterial blight resistance (Viteri et al. 2014). Concurrent selection for desirable characteristics such as seed coat, seed size and growth habit with disease resistance is also recommended to ensure that later generations have the desired phenotypes present. Especially in multiparent crosses, use of the elite parent with superior disease resistance as the last crossing parent was found to produce superior ultimate breeding lines compared to using a susceptible landrace parent. Alternatively, a larger F<sub>1</sub> population could overcome this limitation (Carmen et al. 2005).

There is some evidence for linkage between disease resistance QTL and undesirable agronomic traits. The *BC420* locus is linked to the *V* locus for dark seed coat color and purple flowers, limiting its usefulness in some classes of bean. In addition, there is evidence for linkage between small seed size and high levels of resistance, originating from the *P. acutifolius* gene pool (S. Singh and Miklas 2015). There is a potential negative correlation between presence of

the *SU91* allele and seed yield, though this was only observed in one out of two trial sites (O'Boyle, Kelly, and Kirk 2007).

### **Anthracnose**

Anthracnose, caused by the fungus *Colletotrichum lindemuthianum*, is a major fungal pathogen affecting dry bean production in tropical and temperate regions, and is endemic to New York, Michigan and Ontario (Miklas et al. 2003). Anthracnose can cause severe yield losses or complete crop loss due to early leaf senescence or plant death, as well as blemishes on seed that affect market value. Most severe loss occurs when contaminated seed is planted. Other symptoms are long, angular brick-red to purple lesions that follow the leaf vein, usually on the underside of the leaf. Fungal conidia are spread by wind or rain splashing. Cooler and humid conditions are optimal for proliferation of conidia, so yield losses tend to be more severe in wetter seasons. Stems, petioles and pods develop lesions, which appear circular, dark and sunken with a red halo (Mohammed 2013).

High pathogenic variability characterizes *C. lindemuthianum*, with 247 races identified worldwide, and 117 races in the United States alone (de Lima Castro et al. 2017; Padder et al. 2017). Race 73, a particularly virulent strain, is predominant in North America's major bean producing regions such as Michigan and North Dakota, and in an anthracnose diversity study of North and South America, represented over 25% of all isolates collected. Races 7 and 65 are also prevalent in the United States (Zuiderveen et al. 2016).

*Colletotrichum lindemuthianum* is a hemi-biotrophic fungus that can reproduce both asexually and sexually. The pathogen can be seedborne only when the pods of a plant are infected, and the more pods are infected on the plant, the higher rate of disease in the seeds. The fungus will infect the plant stems before appearing on pods. Presence of the fungus is visible on

the seed coat, and sometimes in the cotyledons as well. The pathogen can survive in viable, air-dried bean seed for five years, and for more than two years in plant residue under field conditions (Mohammed 2013). However, repeated drying and rewetting of residue reduces the lifespan and populations of the fungus, so seed transmission is generally the major form of disease transmission rather than reinfection from field debris. Plowing under debris and using a two-year crop rotation can further reduce reinfection from field residue. Other species hosts include fava bean, mung bean and pea (Mohammed 2013).

Twenty-nine major independent genes for anthracnose have been identified, originating in both primary gene pools. The majority show complete dominance for resistance to a particular pathogen race. All loci segregate independently except for Co-3 and Co-9, which are two allelic forms of the same gene (Vazin et al. 2015). Significant sources of resistance to Middle American pathogen races can be found in Andean germplasm and transferred to Middle American cultivars (de Lima Castro et al. 2017). All identified resistance genes display a qualitative mode of resistance, following the gene-for-gene interaction model (Flor 1955).

Co-4, Co-5 and Co-6 are the most effective combination of resistance genes for North America, all of which are Mesoamerican in origin (Kelly et al. 2003; de Lima Castro et al. 2017; Miklas et al. 2006). Marker-assisted pyramiding of loci has successfully been implemented in pinto and other market classes, but some unsuccessful efforts show concurrent phenotyping should be done as well. The disease is highly variable across environments with many different races of the pathogen, making selection for resistance across regions very difficult. Resistance breeding is also complicated by the fact that pathogenic races tend to be gene pool specific in their virulence (Zuiderveen et al. 2016). The Co-4 gene is most broadly effective across

geographic areas, but different complements of resistance loci are best suited for different regions (Miklas et al. 2006).

### **Bean Common Mosaic Virus**

*Bean common mosaic virus* (BCMV) is a potyvirus commonly transmitted by infected seed or by aphids. Seed transmission of the virus has been shown to occur through infection of the embryo or cotyledon tissue, so rates of seed transmission are generally lower than rates of seed infection. Contamination of seed stocks is a significant problem, and even certified seed stocks have been shown to contain up to 1% of infected seed. This can occur from segregation of resistance genes presumed to be homozygous, or from strains that overcome varietal resistance. Especially if low levels of virus-infected seed are present in a field, aphid-mediated transmission can quickly increase the spread of the virus. However, different species of aphids have drastically different levels of pathogen transmission, so this relationship should not be assumed (Worrall et al. 2015). BCMV is closely related to *Bean common mosaic necrosis virus* (BCMNV), and it has been demonstrated that recombination between the two viruses can occur when there is a simultaneous infection. Common symptoms of both viruses include dwarfing, mosaic leaf pattern, and chlorosis, though severe yield reduction can occur without any of these visible symptoms (Worrall et al. 2015).

Differential pathogenicity of the diverse strains of BCMV have been identified, as well as differing symptoms that interact to varying degrees with the commonly deployed *I* gene, which is characterized by single gene-for-gene interaction. Some virus strains cause a hypersensitive resistance response at moderate temperatures but cause systemic necrosis of the plant, also known as black root syndrome, at temperatures above 30 degrees celsius (Worrall et al. 2015). Other strains cause such necrosis symptoms at lower temperatures in the presence of the *I* gene,

and are known as temperature-independent necrosis inducing strains (Kelly et al. 1995; Kelly et al. 1983).

All strains of BCMNV cause systemic necrosis in plants possessing the *I* gene, regardless of temperature. In regions of Africa where BCMNV is endemic, this necessitates using other resistance genetics than the *I* gene alone. BCMV is more widespread in North America than BCMNV, but both have been identified in the region (Kelly, Saettler, and Morales 1983; Worrall et al. 2015).

Resistance to BCMV is characterized by 4 loci: the monogenic dominant *I* gene and three recessive, multi-allelic loci *bc-3*, *bc-1<sup>2</sup>*, *bc-u*. The *bc-1* and *bc-u* loci are loosely linked (Naderpour et al. 2010). The *I* gene has not been cloned, but is presumed to be part of the TIR-NB-LRR (Toll/interleukin-1-nucleotide binding site-leucine rich repeat) class of resistance gene based on inter-species homology (Worrall et al. 2015). The *I* gene confers a hypersensitive resistance response to BCMV, as well as five other related potyviruses. The recessive loci *bc-u*, *bc-i*, *bc-2*, and *bc-3* confer constitutive resistance that prevents spread of virus proteins (Kelly et al. 2003). The combination of the *I* gene and recessive *bc-3* resistance gene is effective in providing complete resistance to BCMV due to their different resistance mechanisms (Miklas et al. 2006). The *bc-3* gene has an epistatic interaction with the *I* gene, necessitating marker-assisted selection to confirm the presence of both genes rather than phenotypic confirmation (Kelly, Afanador, and Haley 1995). However, as aphid transmission of BCMNV and BCMV is non-persistent, field resistance can be obtained with the *I* gene locus only, as an aphid carrying BCMNV will only transmit necrosis-inducing BCMNV to a single plant in a given field, as the single plant will die and the aphid will clear itself of the BCMNV virus in the act of feeding on

that single plant (T. Parker, personal communication, 2019). Co-dominant genetic markers are available for both the *I* gene and *bc-3* resistance alleles (Bello et al. 2014; Naderpour et al. 2010).

### ***Improvement of specialty dry bean germplasm for Northeastern regional food systems***

As detailed above, high demand exists for dry beans produced in the Northeastern United States outside of commodity market classes. Currently available germplasm includes heirloom and landrace population-varieties valued for their culinary quality and aesthetic value, in particular large-seeded beans from the Andean center of origin. Commodity market classes are characterized by various seed sizes, seed coat colors, and some variation in seed coat pattern. However, a partial white pattern is not represented by any commodity market class and as it is a recessive trait, alleles for this partly-colored phenotype, characterized by white markings set on a solid or patterned colored background, are not present in modern bean cultivars (Bassett and McClean 2000). The goal of this breeding work was to develop lines with partial seed coat pattern as well as favorable agronomic traits that would represent a unique product for regional staple crop systems. Traditional population-varieties possessing the partly-colored seed coat phenotype are susceptible to destructive diseases including BCMV, common blight and anthracnose that are easily transmitted by seed. Common blight is particularly destructive in regions of high precipitation and humidity such as the Northeastern United States. The development of cultivars that combine non-commodity seed traits and superior disease resistance and yield of modern germplasm would support profitable production of dry beans for regional food systems and facilitate regional seed production through enhanced disease resistance.

While resistance to common bacterial blight is complex and identification of new resistance loci is still needed, available breeding lines with multiple resistance loci have been

underutilized in modern cultivar development (S. Singh and Miklas 2015). In addition, all deployment of resistant breeding lines has occurred within major market classes such as navy, kidney and pinto. To the best of our knowledge, the following experiment marks the first instance of introgression of significant common bacterial blight resistance QTL outside of commodity market classes. Selection for resistance to BCMV and anthracnose was also conducted.

### **Materials and Methods**

Parents were selected based on a literature review and complementary traits (Darby et al. 2016; Darby and Cummings 2017, 2018; S. Singh and Miklas 2015) Two traditional dry bean varieties that have been grown extensively in the Northeast, ‘Jacob’s Cattle’ and ‘Calypso,’ were selected as parents with the partly colored seed coat phenotype. ‘Jacob’s Cattle’ likely is  $t^{cf}/t^{cf}$  at the *T* locus while ‘Calypso’ is likely to be *t/t* (Bassett and McClean 2000). Both of these traditional varieties were originally cultivated by indigenous communities in the Americas or the Caribbean, as were many heirloom bean varieties (Dwivedi et al. 2019). They both exhibit type I determinate growth habit and are likely to be race Nueva Granada in the Andean gene pool (Singh et al. 1991). The breeding line ‘USDK-CBB-15,’ a dark red kidney developed by USDA-ARS plant breeder Dr. Phillip Miklas, was selected as the source of resistance to common blight, anthracnose and BCMV as well as high yield [Table 2.1]. ‘USDK-CBB-15’ also possesses a type I determinate growth habit. ‘USDK-CBB-15’ possesses the *I* gene for hypersensitive resistance to BCMV, the *SAP6* and *SU91* loci for partial resistance to common bacterial blight, and the *Co-1* and *Co-2* loci for resistance to some races of *C. lindemuthianum* (Miklas et al. 2006; P. Miklas, personal communication, 2018).

Seed of ‘Jacob’s Cattle’ was obtained from Hudson Valley Seed Company (Accord, NY), seed of ‘Calypso’ was obtained from Seed Saver’s Exchange (Decorah, IA), and seed of “USDK-CBB-15” was obtained from Dr. Phillip Miklas of USDA-ARS (Pullman, WA). Five plants of each parent were grown in Cornell mix media in container culture in the fall of 2018 in the Guterman Research Center on the campus of Cornell University. Initial crosses (‘Calypso’ x ‘USDK-CBB-15’ and ‘Jacob’s Cattle’ x ‘USDK-CBB-15’) were made using hand-pollination, and seed was harvested from individual pods of each parent plant with a successful pollination event. Three “mother” plants of ‘Jacob’s Cattle’ and four of ‘Calypso’ had successful pollinations. All 19 F<sub>1</sub> seeds obtained from crosses were planted in individual pots in winter 2018-2019 in the Guterman Research Center to advance to the F<sub>2</sub> generation.

Approximately 175 F<sub>2</sub> seeds for each biparental cross were planted in peat pots in the certified organic greenhouse at the Guterman Research Center in spring of 2019. All seeds were inoculated with Guard-N N<sub>2</sub>-fixing bacteria immediately before planting. Individual F<sub>2</sub> seedlings were screened for the *SU91* and *SAP6* loci for common blight resistance and the *I* gene for BCMV resistance using a combination of dominant SCAR markers in the case of *SU91* and *SAP6* loci (P. N. Miklas et al. 2000; Pedraza Garcia et al. 1997) and a co-dominant CAPS marker for the *I* locus (Bello et al. 2014). A multiplex polymerase chain reaction (PCR) allowed simultaneous genotyping of the *SU91* and *SAP6* loci following the protocol of O’Boyle et al. (2007). Screening for the co-dominant CAPS marker for the *I* locus was conducted following the protocol of Bello et al. (2014). Plants that had at least one resistance allele for two or more resistance loci (approximately 56% selection proportion) were space-planted at 0.3 meter spacing within rows into the field on certified organic land at Homer C. Thompson Research Farm in June of 2019 and grown to maturity. Plants were irrigated approximately weekly. Individual

plants were phenotyped for seed coat phenotype and plants possessing the partly-colored phenotype ( $t/t$  or  $t^{cf}/t^{cf}$ ) were harvested and selected for advancement (approximately 25% selection proportion). Eight  $F_2$  plants were selected for advancement in a winter greenhouse generation.

Eighteen seeds of each selected  $F_2$  plant were planted in the greenhouse in winter 2020, for a total of 144  $F_{2:3}$  plants. Seedlings were screened with co-dominant high-resolution melt (HRM) markers for *I*, *SAP6* and *SU91* loci (Table 2.2, Lobaton et al. 2018, Soler-Garzon et al. 2021). This marker technology allowed for the detection and selection of plants that are heterozygous or homozygous for the resistance loci as well as a faster genotyping method without the use of a gel-based assay, and protocols followed Wang et al. (2005) and Lobaton et al. (2018). Primer sequences can be found in Table 2.2. The PCR reaction included 10  $\mu L$  genomic DNA, 1x Taq buffer, 0.2 mM of dNTPs mix, 0.2 mM of each primer (two allele-specific forward primers and a common reverse primer), .6x EvaGreen dye, and 0.1  $\mu L$  of Taq polymerase for a total reaction volume of 20  $\mu L$ . The amplification protocol consisted of 2 min initial denaturation at 95°C, and 39 cycles of 15 sec at 94 °C, 20 sec at 58 °C (annealing temperature), and 20 sec at 72 °C. This was followed by 5 min at 75 °C. Profile fluorescence-detection was conducted using a QuantStudio 7 Real-time PCR thermocycler (Applied Biosystems) with initial denaturation for 1 min at 95 °C and a ramped melting curve step from 70 °C to 95 °C with an increment of 0.5 °C per 20 seconds under continuous fluorescence detection. Plants possessing at least one resistance allele at two or more loci were grown to maturity and seed coat phenotype was assessed. At maturity, individual plants were again selected that possessed the partly-colored seed phenotype ( $t/t$  or  $t^{cf}/t^{cf}$ ), resulting in 34 plants selected out of 146 initially planted.

In summer of 2020, 34 F<sub>3:4</sub> families were planted in an augmented design with two replications and four blocks at the East Ithaca Research Farm on Cornell University campus. The resistant parent “USDK-CBB-15” was used as a check variety and was planted in each of the four blocks. Individual plots consisted of two 2.2-m rows planted with a “Precision Garden Seeder” push seeder (EarthWay, Bristol IN USA) at 0.76 m spacing between rows and approximately 50 plants per plot. All seeds were inoculated with “Guard-N®” N<sub>2</sub>-fixing bacteria (Verdesian, Cary NC USA) immediately before planting. Bulked tissue was taken from each family at approximately 21 days after planting and co-dominant HRM markers were run for the *I*, *SAP6* and *SU91* loci as previously described, as well as the *Co-1* locus for resistance to some races of *C. lindemuthianum*. Screening for the *Co-1* locus marker followed the same HRM protocol as outlined above for the *I*, *SAP6* and *SU91* loci, though with an annealing temperature of 68 °C (Zuiderveen et al. 2016). Each F<sub>3:4</sub> family was rated on a whole-plot basis for maturity time, vigor, lodging and disease incidence under natural disease pressure. Lodging was rated on a percent-incidence basis; for example, a rating of 20 indicating that approximately 20% of plants in the plot had lodged. Foliar disease was also rated on a percent-incidence scale; for example, a rating of 40 was given to plots in which 40% of plants exhibited foliar disease symptoms. Plant vigor was rated on a zero to five scale with a score of zero indicating very low vigor and a score of five indicating high vigor. Maturity time was rated on a one to three scale, with one indicating early maturity time and three indicating late maturity time. The predominant disease symptoms observed were identified as common bacterial blight using visual diagnostic methods (Wohleb and du Toit 2011). Where families were segregating for seed coat pattern, only individual plants with the partly colored phenotype were selected. Individual plants were also visually selected based on plant architecture and vigor.

For each phenotype measured in the field, the following mixed linear model was fitted to the data using the lme4 package in R version 3.5.2 (2018-12-20) -- "Eggshell Igloo" (Bates et al. 2015; R Core Team 2017).

$$y_{ijk} = \mu + \alpha_i + b_j + \alpha b_{ij} + \beta_k + \varepsilon_{ijk}$$

$y_{ijk}$  is the observed response of the  $k^{\text{th}}$  breeding line in the  $i^{\text{th}}$  row-pair in the  $j^{\text{th}}$  column-pair and the  $k^{\text{th}}$  column-pair.  $\alpha_i$  is the effect of the  $i^{\text{th}}$  row-pair,  $b_j$  is the effect of the  $j^{\text{th}}$  column-pair,  $\alpha b_{ij}$  is the interaction effect of the  $i^{\text{th}}$  row-pair and the  $j^{\text{th}}$  column-pair,  $\beta_k$  is the effect of the  $k^{\text{th}}$  breeding line and  $\varepsilon_{ijk}$  is the random error term. Field partition coordinates row-pair and column-pair were used as model terms due to high observed within-field variation, rather than replicate and block terms. A variogram was analyzed to examine the data for spatial correlation, but as no correlation was found an autocorrelation was not implemented (Zuur et al. 2009). Least-squares means for breeding line foliar disease ratings were calculated using the ‘emmeans’ function in the emmeans package (Lenth 2020). The same analysis was also conducted for vigor and lodging ratings. Families were then pooled based on bulked common bacterial blight resistance genotype, into one of two pools; 1) ‘sap6SU91’ which denotes two susceptible alleles at the *SAP6* locus and one or more resistant alleles at the *SU91* locus, or 2) ‘SAP6SU91’ which denotes one or more resistance alleles at both *SAP6* and *SU91* loci. No F<sub>4</sub> families homozygous for susceptible alleles at both loci were present, but these two pools were compared to susceptible and resistant parent genotypes instead. Least-squares means for merged genotype values were calculated using the ‘emmeans’ function in the emmeans package (Lenth 2020); Tukey’s honestly significant difference (HSD) was performed using the ‘cld’ function in the multcomp package (Hothorn and Westfall 2008) for pairwise comparisons between least squares means at the  $\alpha=0.05$  significance level.

Table 2.1 Seedborne diseases and causal pathogens included in resistance breeding program, including type of genetic resistance resources available and known resistance loci for which genetic markers have been identified. Loci deployed in the current study are shown in bold.

Disease	Causal organism	Characterization of Resistance and Deployment	Known resistance loci <sup>a</sup>	Citation
Common bacterial blight	<i>Xanthomonas campestris</i> pv. <i>phaseoli</i> (Synonym: <i>Xanthomonas axonopodis</i> pv. <i>phaseoli</i> )	Polygenic resistance; Moderately well characterized; intermediate deployment	<b>SAP6</b> <b>SU91</b> <b>BC420</b> Xa11.4 <sup>OV1,OV3</sup>	(Viteri et al. 2014)
BCMV	<i>Bean common mosaic virus</i>	Major gene resistance; well characterized; widespread deployment	<b>I</b> <i>bc-1</i> <i>bc-2</i>	(Singh and Miklas 2015; Viteri et al. 2014)
Anthracnose	<i>Colletotrichum lindemuthianum</i>	Polygenic, race-specific resistance; deployment challenged by wide virulence diversity	<b>Co-1</b> <i>Many other loci</i>	(de Lima Castro et al. 2017)

<sup>a</sup> Target disease resistance loci are in bold.

Table 2.2 Primers and annealing temperatures used in a High-Resolution Melt (HRM) assay to determine genotype at four resistance loci *SAP6*, *SU91*, *I* and *Co-1*, which are markers for resistance genes for common bacterial blight, *Bean common mosaic virus* and anthracnose (Lobaton et al. 2018, Zuiderveen et al. 2016, Soler-Garzon et al. 2021).

Locus	Forward Primer A	Forward Primer B	Reverse Primer	Annealing Temp (°C)
<i>SAP6</i>	gcgggcCTTTACCAC ATT TCCACTATTTTTTC	gcgggcagggcggcCTTTA CCACATTTCCACTAT TTTTG	AAACTTTCGCACTT TCTCATCA	56
<i>SU91</i>	gcgggcCAGTGATG TTAATGCGCAAT	gcgggcagggcggcCCAGT GATGTTAATGCGCAA A	TTTTGATGGAGGGA AGCAAC	61
<i>I</i>	gcgggcCAAAGTGC TAGAGGCATGAT CA	gcgggcagggcggcCAAAG TGCTAGAGGCATGAT CT	TGGTTATCATTTCAT TGTGAAGTCAATG	58
<i>Co-1</i>	gcgggcTGTTGGGTT CTGCGAAGAA	gcgggcagggcggcTGTTG GGTTCTGCGAAGAG	GCGCCCCATTCTTC ATGTACT	68

## **Results**

In the F<sub>4</sub> generation, approximately 81% of loci in the genome are fixed (Kelly and Cichy 2012), indicating that we were likely to see some segregation for visible traits in our family plots. Substantial segregation for seedcoat phenotypes was observed, both in partial pattern phenotype as well as background pattern and color. Seventeen F<sub>3:4</sub> families were observed to be fixed at the recessive *T* locus for partly colored seed coat. A subset of examples demonstrating the range of seed phenotypes observed is shown in Figure 2.1 .

Using genetic markers, all F<sub>3:4</sub> families were genotyped for the four disease resistance loci of interest. Out of these 17 families, six families were either homozygous or heterozygous for all four disease resistance loci *I*, *SAP6*, *SU91* and *Co-1*. One family was homozygous for resistance alleles at all four targeted loci (Table 2.2).

In field experiments phenotyping the F<sub>3:4</sub> families for foliar disease incidence under natural disease pressure, some plots appeared more similar to the ‘USDK-CBB-15’ parent in larger plant stature, longer maturity time and/or low incidence of foliar pathogens, while other plots more resembled the ‘Jacob’s Cattle’ or ‘Calypso’ parent with smaller plant stature, shorter maturity time and/or high incidence of foliar pathogens. Significant differences in foliar disease incidence at the  $\alpha=0.05$  significance level were observed between USDK-CBB-15, the parent with both *SAP6* and *SU91* common blight resistance alleles, and the two heirloom parents, ‘Jacob’s Cattle’ and ‘Calypso,’ which are recessive for both alleles (Figure 2.2). Possibly due to high within-field variability of growing conditions, comparisons between differing common blight resistance genotypes of pooled F<sub>4</sub> families by genotype were not found to be significant at the  $\alpha=0.05$  significance level. An attempt was made to control for within-field variability using coefficients for plot row and column interaction in the mixed linear model, but substantial

variance remained even after introducing these coefficients into the model (Table 2.3).

Resistance at the *SAP6* locus did not appear to confer additional resistance when compared to lines with resistance at the *SU91* locus only (Figure 2.2). This could indicate the presence of more virulent strains of *X. campestris*, as *SAP6* only confers resistance in plant leaves and pods in response to less virulent strains.

Figure 2.1 Seed phenotypes of a subset of F<sub>4</sub> families from two biparental crosses, displaying diverse seed coat colors and patterns present across all F<sub>4</sub> families. Some families continue to show segregation for one or more seed coat traits.



Table 2.3 Seedcoat phenotype and genotype at four resistance loci, *I*, *SAP6*, *SU91* and *Co1*, of 17 partial-patterned F<sub>3:4</sub> bean lines and three parent lines<sup>1</sup> from which they were derived. Loci represent resistance genes for *Bean common mosaic virus*, common bacterial blight and anthracnose. Genotype score ‘a’ represents homozygous resistant, ‘b’ represents homozygous susceptible and ‘ab’ represents heterozygous genotypes. Genotypes were analyzed based on a bulk tissue sample of four plants per plot.

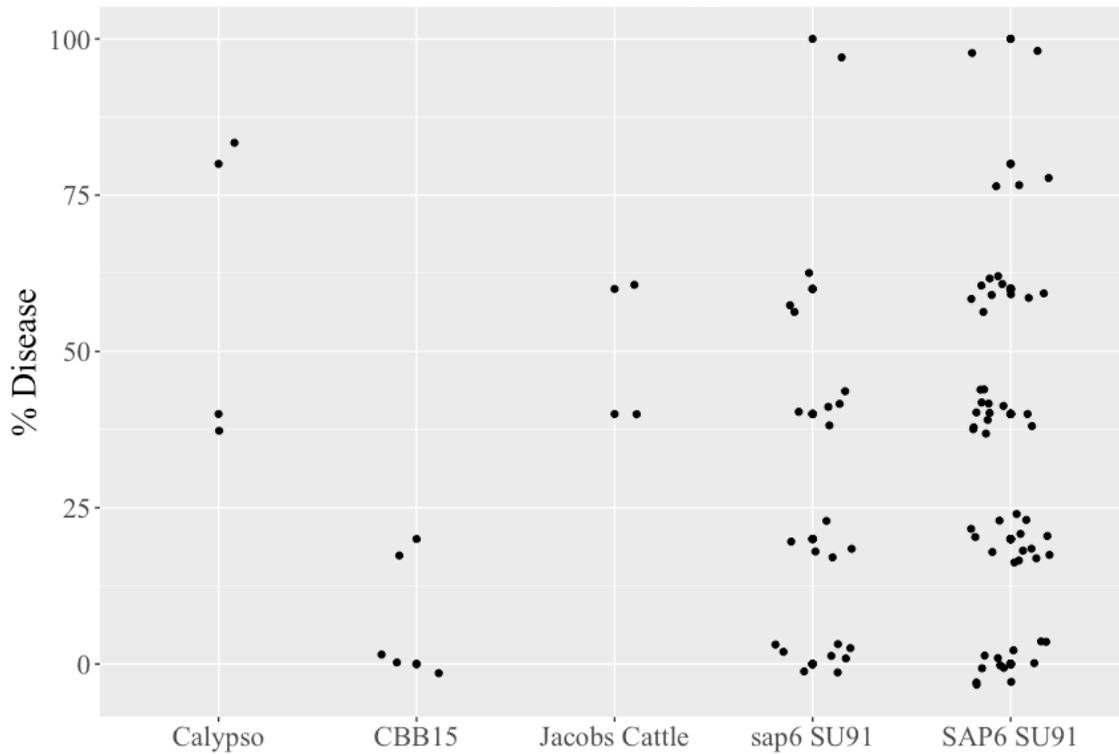
F <sub>4</sub> Family	<i>I</i>	<i>SAP6</i>	<i>SU91</i>	<i>Co1</i>
20-3050	a	b	a	ab
20-3023	a	b	a	a
20-3042	a	ab	ab	ab
20-3043	ab	b	ab	b
20-3045	b	a	ab	ab
20-3025	b	ab	ab	b
20-3036	a	ab	a	ab
20-3048	a	a	a	b
20-3053	a	a	a	ab
20-3029	a	b	a	a
20-3030	a	b	a	a
20-3031	a	b	a	a
20-3046	ab	b	a	ab
20-3051	a	a	ab	ab
20-3022	a	ab	a	a
20-3033	a	b	a	ab
20-3040	a	a	a	a
20-3056	a	ab	ab	ab
20-3057	ab	ab	a	ab
20-3058	ab	a	a	ab

<sup>1</sup> Parents were ‘USDK-CBB-15’, ‘Calypso’, and ‘Jacob’s Cattle’

Table 2.4 Random effects of linear mixed model used to evaluate percent disease ratings for parent and bulked F<sub>4</sub> families in a replicated field experiment. Column Pair and Row Pair represent gridded location within the experimental plot. The ‘lme4’ package was used in R to create the model (Bates et al. 2015).

Group	Name	Variance	Standard Deviation
ColumnPair:RowPair	(Intercept)	35.90	5.991
RowPair	(Intercept)	0.00	0.00
ColumnPair	(Intercept)	87.17	9.337
Residual		363.23	19.059

Figure 2.2 Percent disease ratings for common bacterial blight of parent and bulked F<sub>4</sub> family genotypes pooled by resistance genotype at two loci, SU91 and SAP6. Ratings were evaluated on a percent scale in 20% increments under natural disease pressure during the 2020 field season.



## *Discussion*

Due to limited seed in the F<sub>3:4</sub> generation, field-scale evaluations were limited in their power to distinguish statistically significant differences between lines, and assessment of performance across environments and years was not possible. Important next steps in the variety development process for these lines would be replicated yield testing and more accurate field selection for disease resistance. As bacterial blight and anthracnose resistance loci present in breeding lines are likely to only confer partial resistance or resistance to a limited number of races, phenotypic evaluation of resistance would be recommended. In particular, results obtained in this study resemble previous studies showing loose correlations between resistance genotypes and phenotypic disease scores, with some lines possessing both resistance alleles still demonstrating high levels of susceptibility in the field. This underscores the importance of combining MAS with DDS methodologies (Duncan et al. 2012).

Another reason for compromised phenotyping accuracy in the field may have been differential levels of inoculum present on the seeds of different lines, as the F<sub>3:4</sub> lines planted from greenhouse-grown seed were planted adjacent to other lines (not included in the experiment) that were planted from F<sub>2:3</sub> seed that originated in 2019 field trials with heavy disease incidence of bacterial blight. This could have resulted in “hotspots” within the 2020 experimental plots that introduced additional variance to foliar disease incidence. Some significant differences were found for plant vigor and lodging at the  $\alpha = .1$  level, however results may not be replicable in terms of use for selection of superior families, due to high standard residual error in the analysis (Table 2.3).

Bacterial diseases such as common blight tend to be prevalent on common bean under field conditions in our region, however, a more accurate disease phenotyping strategy for common

blight may be an inoculated disease incidence trial, which would likely improve uniformity of pathogen pressure across experimental plots and ensure that only the targeted bacterial pathogen is present. Seeds could be surface sterilized before planting to decrease external inoculum on seed surface and increase the chances of even bacterial presence across the field following inoculation. As anthracnose was not observed in experimental plots in any of the three growing seasons, an inoculated trial or collaboration with a grower with consistent anthracnose pressure may be needed. As several races of *C. lindemuthianum* are likely present in the Northeast and information on which races are present is not readily available, a field trial may be more useful than inoculation with lab collections of the pathogen. Alternatively, fungal spores could be collected from natural outbreaks in the region, including infected seed, and an inoculum disease trial could be conducted using the collected pathogen.

Dry bean cultivars are typically selected from one or a few intermediate-generation plants, producing an extremely uniform final cultivar. This strategy could be employed in this instance by selecting single plants in the F<sub>5</sub> generation, or continued bulk selections within or between families could continue, producing a cultivar or population that is adequately uniform in traits such as maturity time or seed size, which is important for successful commercial production, while maintaining genetic diversity that may improve yield stability and field-scale stress tolerance as well as disease resistance over time. Approaches such as evolutionary breeding and the use of multilines offer a precedent for this population approach, especially when selecting for heterogeneous or marginal environments (Mundt 2002; K. Murphy et al. 2005). Other considerations for future genetic improvement include the engagement of supply chain stakeholders such as farmers, consumers and chefs in the selection process. This could include multi-location trials of F<sub>5</sub> or later populations by farmers, who could evaluate performance in

their environment and make selections at the individual plant or family level. This participatory evaluation strategy could also leverage larger population sizes to engage chefs or consumers in culinary quality analyses of selected families. This strategy, in which diverse stakeholders are engaged in the evaluation and selection process at an intermediate generation when sufficient seed numbers are achievable and populations have already been selected for traits such as disease resistance, has been successfully been deployed in other self-pollinated crops in the Northeast such as wheat (Kissing-Kucek 2017b).

### ***Conclusion***

In summary, this experiment was successful in generating multiple breeding lines that combine a partial seedcoat phenotype from heirloom parents ‘Calypso’ and ‘Jacob’s Cattle’ with major loci for resistance to significant bacterial, fungal and viral pathogens in the Northeast United States. These lines also incorporate superior agronomic performance for commercial production from the modern ‘USDK-CBB-15’ parent line. The partial seedcoat phenotype is not present in any major market classes and is likely to be appealing for consumers interested in specialty and heirloom bean varieties, while providing important agronomic traits for commercial growers. As these lines are currently in the F<sub>4</sub> generation, further screening is needed to screen and select for disease resistance phenotypes and superior agronomic performance, as well as evaluate consumer-facing traits such as culinary quality. Especially for common blight, selection of resistant lines could be improved by inoculated disease trials that better control intra-field variability in inoculum load. As maturity time between the heirloom and modern parents were also substantially different, there will likely be potential to select lines more suited to areas with longer growing seasons of the region, such as Central and Western New York and Pennsylvania, as well as quicker maturing lines that would be more suited to northern regions such as Maine.

These lines could be carried forward as pure line cultivars, or maintained as diverse populations that would be ideally suited to on-farm selection and adaptation.

## REFERENCES

- Agarwal, V. K. and J. B. Sinclair. 1997. *Principles of Seed Pathology (2nd edition)*. Boca Raton: Lewis Publishers.
- Bassett, Mark J. and Phil McClean. 2000. *A Brief Review of the Genetics of Partly Colored Seed Coats in Common Bean*. Bean Improvement Collaborative Reports.
- Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using Lme4." *Journal of Statistical Software* 67(1):1–48.
- Bello, Marco H., Samira M. Moghaddam, Mark Massoudi, Phillip E. McClean, Perry B. Cregan, and Phillip N. Miklas. 2014. "Application of in Silico Bulk Segregant Analysis for Rapid Development of Markers Linked to Bean Common Mosaic Virus Resistance in Common Bean." *BMC Genomics* 15(1):903.
- Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". *Journal of the Royal Statistical Society* 57(1): 289-300.
- Carmen, M., Asensio-S. Manzanera, Carmen Asensio, and Shree Singh. 2005. "Introgressing Resistance to Bacterial and Viral Diseases from the Middle American to Andean Common Bean." *Euphytica* 143:223–28.
- Darby, Heather. 2016. *Northeast Dry Bean Production Guide*. UVM Northwest Crops and Soils Program.
- Darby, Heather and Erica Cummings. 2017. *2016 Heirloom Dry Bean Variety Trial*. UVM Northwest Crops and Soils Program.
- Darby, Heather and Erica Cummings. 2018. *2017 Heirloom Dry Bean Variety Trial*. UVM Northwest Crops and Soils Program.
- Darby, Heather, Erica Cummings, Lily Calderwood, Abha Gupta, Julian Post, and Sara Ziegler. 2016. *2015 Heirloom Dry Bean Variety Trial*. UVM Northwest Crops and Soils Program.
- Darby, Heather, Rory Malone, Erica Cummings, and Hillary Emick. 2019. *The Effects of Seed Steam Treatment on Dry Bean Yield and Presence of Pests & Disease*. UVM Northwest Crops and Soils Program.
- Döring, Thomas F., Samuel Knapp, Geza Kovacs, Kevin Murphy, and Martin S. Wolfe. 2011. "Evolutionary Plant Breeding in Cereals-Into a New Era." 3.
- Duncan, Robert, Robert Gilbertson, and Shree Singh. 2012. "Direct and Marker-Assisted Selection for Resistance to Common Bacterial Blight in Common Bean." *Crop Science* 52.

- Duncan, Robert W., Shree Singh, Robert L. Gilbertson, and W. Duncan. 2011. "Interaction of Common Bacterial Blight Bacteria with Disease Resistance Quantitative Trait Loci in Common Bean." *Phytopathology* 101(4):425.
- Dwivedi, Sangam, Irwin Goldman, and Rodomiro Ortiz. 2019. "Pursuing the Potential of Heirloom Cultivars to Improve Adaptation, Nutritional, and Culinary Features of Food Crops." *Agronomy* 9(8):1–26.
- Ellis, M. A. and Guillermo E. Gálvez E. 1980. "Seed Pathology." in *Bean production problems: Disease, insect soil and climatic constraints of Phaseolus vulgaris*, edited by G. E. Schwartz, Howard F.; Gálvez E. Centro Internacional de Agricultura Tropical (CIAT).
- Flor, H. H. 1955. "Host-Parasite Interaction in Flax Rust - Its Genetics and Other Implications." *Phytopathology* 45:680–85.
- Kelly, James D., L. Afanador, and S. D. Haley. 1995. "Pyramiding Genes for Resistance to Bean Common Mosaic Virus." *Euphytica* 82:207.
- Kelly, James D. and Karen A. Cichy. 2012. "Dry Bean Breeding and Production Technologies." Pp. 23–54 in *Dry Beans and Pulses Production, Processing and Nutrition*. Oxford, UK: Blackwell Publishing Ltd.
- Kelly, James D., P. Gepts, P. N. Miklas, and D. P. Coyne. 2003. "Tagging and Mapping of Genes and QTL and Molecular Marker-Assisted Selection for Traits of Economic Importance in Bean and Cowpea." *Field Crops Research* 82(2–3):135–54.
- Kelly, James D., A. W. Saettler, and M. Morales. 1983. *New Necrotic Strain of Bean Common Mosaic Virus in Michigan*.
- Kissing-Kucek, Lisa. 2017. "Participatory Wheat Breeding for Organic Production". Cornell University.
- Russell Lenth (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.2-1. <https://CRAN.R-project.org/package=emmeans>
- de Lima Castro, Sandra Aparecida, Maria Celeste Gonçalves-Vidigal, Thiago Alexandre Santana Gilio, Giselly Figueiredo Lacanallo, Giseli Valentini, Vanusa da Silva Ramos Martins, Qijian Song, Marta Zulema Galván, Oscar P. Hurtado-Gonzales, and Marcial Antonio Pastor-Corrales. 2017. "Genetics and Mapping of a New Anthracnose Resistance Locus in Andean Common Bean Paloma." *BMC Genomics* 18(1):306.
- Lobaton, Juan David, Tamara Miller, Juanita Gil, Daniel Ariza, Juan Fernando de la Hoz, Alvaro Soler, Steve Beebe, Jorge Duitama, Paul Gepts, and Bodo Raatz. 2018. "Resequencing of Common Bean Identifies Regions of Inter-Gene Pool Introgression and Provides Comprehensive Resources for Molecular Breeding." *The Plant Genome* 11(2):0.

- Miklas, P. N., J. D. Kelly, and S. P. Singh. 2003. "Registration of Anthracnose-Resistant Pinto Bean Germplasm Line USPT-ANT-1." *Crop Science* 43(5):1889.
- Miklas, P. N., S. P. Singh, H. Terán, J. D. Kelly, and J. R. Smith. 2011. "Registration of Common Bacterial Blight Resistant Cranberry Dry Bean Germplasm Line USCR-CBB-20." *Journal of Plant Registrations* 5(1):98–102.
- Miklas, P. N., J. R. Smith, and S. P. Singh. 2006. "Registration of Common Bacterial Blight Resistant Dark Red Kidney Bean Germplasm Line USDK-CBB-15." *Crop Science* 46(2):1005–7.
- Miklas, P. N., JR Smith, R. Riley, KF Grafton, SP Singh, G. Jung, and DP Coyne. 2000. "Marker-Assisted Breeding for Pyramided Resistance to Common Bacterial Blight in Common Bean." *Annual Report of the Bean Improvement Cooperative* 43:39–40.
- Miklas, Phillip, James D. Kelly, Steve E. Beebe, and Matthew W. Blair. 2006. "Common Bean Breeding for Resistance against Biotic and Abiotic Stresses: From Classical to MAS Breeding." *Euphytica* 147:105–31.
- Mohammed, Amin. 2013. "An Overview of Distribution, Biology and the Management of Common Bean Anthracnose." *Journal of Plant Pathology & Microbiology* 04(08).
- Mundt, C. C. 2002. "Use of Multiline Cultivars and Mixtures for Crop Disease Management." *Annu. Rev. Phytopathol* 40:381–410.
- Murphy, Kevin, Doug Lammer, Steve Lyon, Brady Carter, and Stephen S. Jones. 2005. "Breeding for Organic and Low-Input Farming Systems: An Evolutionary–Participatory Breeding Method for Inbred Cereal Grains." *Renewable Agriculture and Food Systems* 20(1):48–55.
- Naderpour, Masoud, Ole Sogaard Lund, Richard Larsen, and Elisabeth Johansen. 2010. "Potyviral Resistance Derived from Cultivars of Phaseolus Vulgaris Carrying Bc-3 Is Associated with the Homozygotic Presence of a Mutated EIF4E Allele." *Molecular Plant Pathology* 11(2):255–63.
- O'Boyle, Patrick, James Kelly, and William Kirk. 2007. "Use of Marker-Assisted Selection to Breed for Resistance to Common Bacterial Blight in Common Bean." *Journal of the American Society for Horticultural Science* 132(3):381–86.
- O'Leary, Michael. 2009. "Bacterial Blights of Common Bean: Application of Traditional, Molecular and Genomic Sequences for Improved ID and Understanding of Mechanisms of Pathogenesis." University of California, Riverside.
- Padder, B. A., P. N. Sharma, H. E. Awale, and J. D. Kelly. 2017. "Colletotrichum Lindemuthianum, the Causal Agent of Bean Anthracnose." *Journal of Plant Pathology* 99(2):317–30.

- Pedraza Garcia, Fabio, Gerardo Gallego, Stephen Beebe, and Joseph Tohme M. 1997. “[SCAR and RADP Markers for the Common Bacterioses (CBB) Resistance of Beans].” *Agris*.
- R Core Team. 2017. “R: A Language and Environment for Statistical Computing.”
- Singh, Shree. 2001. “Broadening the Genetic Base of Common Bean Cultivars.” *Crop Science* 41(6):1659.
- Singh, Shree and Phillip Miklas. 2015. “Breeding Common Bean for Resistance to Common Blight: A Review.” *Crop Science* 55(3):971.
- Singh, Shree P., Paul Gepts, and Daniel G. Debouck. 1991. “Races of Common Bean (*Phaseolus Vulgaris*, Fabaceae).” *Economic Botany* 45(3):379–96.
- Thomas, Mathieu, Elise Demeulenaere, Julie C. Dawson, Abdul Rehman Khan, Nathalie Galic, Sophie Jouanne-Pin, Carine Remoue, Christophe Bonneuil, and Isabelle Goldringer. 2012. “On-Farm Dynamic Management of Genetic Diversity: The Impact of Seed Diffusions and Seed Saving Practices on a Population-Variety of Bread Wheat.” *Evolutionary Applications* 5(8):779–95.
- Vazin, Maryam, Alireza Navabi, Peter K. Pauls, Mary Ruth McDonald, and Chris Gillard. 2015. “Characterization of Anthracnose Resistance in Common Bean.” University of Guelph.
- Viteri, Diego M., Perry B. Cregan, Jennifer J. Trapp, Phillip. N. Miklas, and Shree P. Singh. 2014. “A New Common Bacterial Blight Resistance QTL in VAX 1 Common Bean and Interaction of the New QTL, SAP6, and SU91 with Bacterial Strains.” *Crop Science* 54(4):1598–1608.
- Viteri, Diego, Henry Terán, Mari Carmen Asensio-S.-Manzanera, Carmen Asensio, Timothy Porch, Phillip Miklas, and Shree Singh. 2014. “Progress in Breeding Andean Common Bean for Resistance to Common Bacterial Blight.” *Crop Science* 54(5):2084.
- Wang, Jun, Karen Chuang, Mandeep Ahluwalia, Sarika Patel, Nanette Umblas, Daniel Mirel, Russell Higuchi, and Soren Germer. 2005. “High-Throughput SNP Genotyping by Single-Tube PCR with T<sub>m</sub>-Shift Primers.” 39(6).
- Wohleb, Carrie and Lindsey du Toit. 2011. *Common Bacterial Blight and Halo Blight Two Bacterial Diseases of Phytosanitary Significance for Bean Crops in Washington State*. WSU Extension Fact Sheet FS038E.
- .
- Worrall, Elizabeth A., Francis O. Wamonje, Gerardine Mukeshimana, Jagger J. W. Harvey, John P. Carr, and Neena Mitter. 2015. “Bean Common Mosaic Virus and Bean Common Mosaic Necrosis Virus.” *Advances in Virus Research* 93:1–46.

- Yu, K., S. J. Park, and V. Poysa. 2000. "Marker-Assisted Selection of Common Beans for Resistance to Common Bacterial Blight: Efficacy and Economics." *Plant Breeding* 119(5):411–15.
- Zuiderveen, Grady H., Bilal A. Padder, Kelvin Kamfwa, Qijian Song, and James D. Kelly. 2016a. "Genome-Wide Association Study of Anthracnose Resistance in Andean Beans (*Phaseolus Vulgaris*)." *PLoS ONE* 11(6).
- Zuur, Alain F., Elena N. Ieno, Neil J. Walker, Anatoly A. Saveliev, and Graham M. Smith. 2009. "Violation of Independence – Part II." Pp. 161–91 in *Mixed Effects Models and Extensions in Ecology with R*.

## CHAPTER III

### EXPLORATION OF INTRA-VARIETAL DIVERSITY OF ‘JACOB’S CATTLE’ DRY BEAN WITHIN DECENTRALIZED SEED SAVER NETWORKS

#### *Introduction*

In the 20<sup>th</sup> century, modern breeding efforts for commercial production have led to the gradual replacement of region-specific varietal strains that are reproduced on farm with a relatively narrow genetic pool of elite cultivars developed to be broadly adapted to high input, high productivity growing environments (Ceccarelli, 1989; Bellucci et al., 2013; Gioia et al., 2019). Consolidation of the plant breeding and seed industries has further centralized development and maintenance of modern crop varieties (Howard, 2015). Industry consolidation has further restricted germplasm access for public sector breeders, where public sector germplasm for crop improvement is typically limited to within the breeding program, from other public breeding programs or from the USDA National Plant Germplasm Collection (NPGS) (Kelly, 2010). The result of this for farmers is that seed of a genetically uniform cultivar, grown in one of a small number of seed producing regions, is typically purchased each year rather than being replanted from the prior season’s crop. The advantages of this system include seed quality assurance and predictable variety performance (Kelly, 2010). However, this system also drastically reduces intra-variety genetic diversity, minimizing the potential buffering effects of genetic variation at the field scale in response to environmental stress (Dawson and Goldringer, 2012). This system also prevents continued natural and human selection within many diverse environments across the landscape, a process which has formed the basis for modern crop domestication and diversity over millennia (Thomas et al., 2012; Stone et al., 2019). This age-old

stewardship of crop varieties on a farm or in a garden is defined as *In-situ* management, in contrast to *ex situ* germplasm conservation strategies employed by gene banks such as the NPGS collection, in which collected accessions are propagated only as frequently as is necessary to maintain seed viability (Thomas et al., 2011). The phrase “genetic resource management” will be used in this study, rather than the commonly used terms “conservation” or “preservation”, which inaccurately imply an unchanging or static composition within a plant population across space and time. (Klaedtke et al., 2017).

The importance of *In-situ* farmer management to crop evolution and adaptation is widely recognized in the context of the Global South, particularly in centers of domestication (Hufford et al., 2019). However, *In-situ* germplasm management receives little attention in industrialized agricultural systems of the Global North. Despite this lack of attention, traditional crop varieties present today within farmer and seed saver networks continue to be cultivated across wide agro-ecological gradients for dozens to hundreds of years, diverging through dynamics of natural and artificial selection, de novo mutation, outcrossing, environmentally triggered genetic changes, and genetic drift (Stone et al., 2019). Past work indicates that autogamous crops managed *In-situ* show evidence of significant changes in allele frequencies, even over relatively short time scales (Döring et al., 2011; Klaedtke et al., 2017; Bhandari and Gauchan, 2018). Examining variation within *In-situ* stewarded varieties has specific implications for dry bean (*Phaseolus vulgaris*). Traditional dry bean varieties are increasingly valued in industrialized countries for their culinary and aesthetic quality, historical significance, and potential adaptation to specific environmental conditions (Swegarden et al., 2016; Brouwer et al., 2016; Wilker et al., 2019). Recent work has begun to evaluate North American traditional bean varieties for yield and other useful traits as well as profitability for commercial production (Darby and Cummings 2018; Swegarden et al.,

2016), but single commercial seed sources are typically utilized, and an evaluation of intra-varietal diversity of traditional varieties resulting from *In-situ* management within decentralized seed saving networks in North America is lacking.

In the Global North, grassroots movements of seed saving and regional seed company networks has arisen in response to trends of decreasing cropping system diversity at both species and variety levels (Veteto, 2008; Aguilar et al., 2015). A central example in these efforts in the United States has been Seed Saver's Exchange, a not-for-profit organization started in 1975 and located in Decorah, Iowa (USA). Seed Savers Exchange maintains a germplasm collection of heirloom crop varieties and manages a seed exchange network through which seed savers can list and exchange seeds (Seed Savers Exchange Heirloom Seeds, 2021). The Exchange, through which thousands of varieties are offered by hundreds of listers across the United States, most not commercially available, is an invaluable resource for freelance plant breeders (Deppe, 2021). This network also offers a relatively rare example in industrialized countries of sustained *In-situ* management of crop varieties over the course of nearly fifty years.

The purpose of this study was to evaluate the extent of genetic and phenotypic diversity of a traditional North American dry bean variety, 'Jacob's Cattle', maintained within decentralized seed saving networks. 'Jacob's Cattle', also known as 'Trout' or 'Cattle', was selected for the study due to its widespread cultivation across North America both in commercial production as well as gardener and seed saver networks. 'Jacob's Cattle' or 'Trout', whose origins in North America spans centuries, is also described as a market class rather than a variety (P. McClean, personal communication, 2020). The origins of 'Jacob's Cattle' were not able to be confirmed by the authors, but this variety was certainly cultivated by native communities in the Northeast since at least European colonial contact, if not prior. 'Jacob's Cattle' is of race Nueva

Grenada within the Andean gene pool, with center of origin in the Andean region of South America (Cichy et al., 2015). We obtained 18 unique sources of ‘Jacob’s Cattle’ bean from Seed Savers Exchange networks, farmer networks and the USDA-ARS germplasm collection, as well as five unique sources of a commercial light red kidney variety for comparison. All seed sources were genotyped using Genotyping-by-Sequencing (GBS) and all ‘Jacob’s Cattle’ sources were grown in replicated field plots to explore possible correlation of genetic diversity to observable phenotypic diversity. These data were additionally compared to sequence data of a set of North American traditional dry bean varieties obtained from a University of Minnesota study. Overall, this study advances our knowledge of the effects of *In-situ* seed stewardship on crop diversity by measuring genetic and phenotypic divergence within ‘Jacob’s Cattle’ bean seed sources stewarded by seed savers and farmers.

### ***Materials and Methods***

#### ***Germplasm***

In September of 2019, accessions to be included in the analysis were identified through several channels: 1) the National Plant Germplasm System’s Germplasm Information Resources Network (NPGS-GRIN); 2) the Seed Exchange portal hosted by Seed Savers Exchange, a not-for-profit organization based in Decorah, Iowa; 3) an internet search of farm-based seed company inventories; and 4) snowball sampling of farmers in Maine, believed to be the historical region of origin of ‘Jacob’s Cattle’ in North America. Accessions were identified by variety name, which included either the name ‘Jacob’s Cattle’ or by morphological characteristics that showed close similarity to ‘Jacob’s Cattle’, despite a different name. Seed samples were requested and received from 18 distinct accessions were received with the number of seeds provided varying (**Table 3.1**).

Seed sources varied from home gardens to commercial farms, and consequently population sizes varied widely but not all source population sizes were able to be confirmed, so these data were not included in the study. Some sources had only been cultivating the variety for a few years, while one seed saver source had been growing and saving seed of ‘Jacob’s Cattle’ for 30 years. Source 706, which came from the NPGS-GRIN collection, was visually identified as being an admixture of ‘Jacob’s Cattle’ and ‘Anasazi’, the latter of which is a Southwestern heirloom with similar color and pattern to ‘Jacob’s Cattle’ but of Middle American origin and thus highly genetically distinct from ‘Jacob’s Cattle’. Two individuals of each type were included in subsequent genotyping, which confirmed visual identification of admixed seed as being of Middle American origin, likely race Durango. These two individuals were included in downstream analyses as ‘Anasazi’, rather than ‘Jacob’s Cattle’. Finally, as a check comparison, an attempt was made to acquire 18 sources of ‘Cal Early’ light red kidney, a standard commercial variety. Only five sources were identified and seed was acquired from each of these, which is consistent with our experience of highly centralized seed production for modern cultivars.

Sequence data for additional traditional dry bean genotypes was obtained from University of Minnesota to provide further context for the study (NCBI – SRA Bioproject ID: PRJNA667092). These data included five heirloom varieties and two market class checks. For four of the heirloom varieties, ‘Jacob’s Cattle Gold’, ‘Lina Sisco’s Bird Egg’, ‘Tiger’s Eye’, and ‘Peregion’, 20 individual plants, originating from pure lines that had been selected from single commercial seed sources, were genotyped. For the remaining varieties, ‘Lariat’ pinto, ‘Eclipse’ black and ‘Painted Pony’ heirloom, single individuals were genotyped (Swegarden and

Michaels, personal communication). This sequence data was pooled with sequence data obtained from the present study, realigned and SNPs called using methods described below. Passport data for genotypes included in the study can be found in Table 3.1.

#### *Genotyping and genetic diversity analysis*

In Winter of 2020, five seeds from each source were planted in Cornell mix media and grown in container culture in the Guterman Research Center on the campus of Cornell University (Ithaca, NY USA). Leaf tissue of the first trifoliolate was collected 21 days after planting from four individuals per source for genotyping. For the purposes of subsequent field evaluations in Summer of 2020, all five plants from sources with fewer than 150 seeds were grown to maturity and harvested for the purpose of increasing seed for field trial plots. DNA was extracted using the Qiagen DNEasy 96 Plant Kit (Qiagen Inc., Valencia, CA, USA) according to the manufacturer's instructions.

A 96-plex Genotyping-By-Sequencing (GBS) library (Elshire *et al.* 2011) was prepared and sequenced on a NovaSeq 6000 (Illumina, San Diego, CA, USA) with shared-lane paired-end 150 bp reads at the University of Wisconsin-Madison Biotechnology Center (Madison, WI, USA). Reads were aligned to the *Phaseolus vulgaris* genome (v2.1) (DOE-JGI and USDA-NIFA) with the “bwa” aligner in the GBSv2 pipeline in TASSEL 5 (Glaubitz *et al.*, 2014). In VCFtools (Danecek *et al.*, 2011), SNPs that were not biallelic, had extreme mean read depths (<3, or > 30), low minor allele frequency (<0.05), or were missing in >40% of samples were removed. SNPs were lastly filtered to exclude loci with >10% heterozygosity to remove any other likely misaligned reads. As common bean is highly inbreeding typically (90-99%), high proportions of heterozygosity at the same loci across sources was deemed unlikely. Raw reads from a University of Minnesota GBS analysis of five heirloom dry bean varieties were also

aligned and SNPs called along with ‘Jacob’s Cattle’ using the same pipeline as above (NCBI – SRA Bioproject ID: PRJNA667092). In these data, 20 individuals of a single source of each of five varieties were genotyped, along with single plant individuals of two market class check varieties representing two races of the Middle American gene pool (Singh et al., 1991).

Total SNP counts within populations and sub-populations for ‘Jacob’s Cattle’, ‘Cal Early’ and each of the five other traditional bean varieties were tallied using TASSEL 5 (Bradbury et al., 2007). A Principal Components Analysis (PCA) of two datasets were conducted in TASSEL5: 1) All varieties and seed sources 2) ‘Jacob’s Cattle’ individuals only. PCA results were plotted in ggplot2 (Wickham, 2016). A Discriminant Analysis of Principle Components (DAPC) was conducted using the ‘find.clusters’ function in the ‘adegenet’ package (Jombart, 2008) to obtain k-means group assignments for sub-populations. In the analysis, Bayesian Information Criterion (BIC) values were minimized as a statistical measure of goodness of fit.

Minor alleles of individual genotypes grouped by assigned cluster, including both homozygous and heterozygous calls, were then plotted by physical location across all eleven chromosomes to visualize SNP distribution and identify haplotype blocks. In this instance, “minor alleles” represent a genotype at a given locus that differs from that of the predominant genetic cluster identified within seed source genotypes. Based on these data, a separate plot was subsequently drawn that included a subset of seed sources as well as ‘Lariat’ with known Middle American gene pool assignment to evaluate for evidence of inter-gene pool recombination. Only SNP positions with polymorphism in the ‘Jacob’s Cattle’ dataset were included, meaning that positions at which a divergent genotype would only be assigned to pinto or black bean genotypes were not included. This allowed greater resolution to analyze similarity between ‘Jacob’s Cattle’ sources and Middle American genotypes, but does not allow accurate analysis of genetic

divergence between ‘Jacob’s Cattle’ and Middle American genotypes. 1356 SNPs were included in this analysis. Data was plotted in ggplot2 (Wickham, 2016).

*Phenotypic diversity measurements and analysis*

In summer of 2020, seeds from all 18 ‘Jacob’s Cattle’ sources were planted in an augmented design with 2 replications and 4 blocks at the East Ithaca Research Farm on Cornell University campus (Ithaca, NY USA). Phenotypic data was not collected on ‘Cal Early’ accessions or any lines for which sequence data was obtained from the University of Minnesota. Four sub-populations were selected as checks were planted in each of the four blocks, while the remaining sub-populations were each planted in two out of four blocks. Individual plots consisted of two 2.2-m rows planted with a “Precision Garden Seeder” push seeder (EarthWay, Bristol IN USA) at 0.76 m spacing between rows and approximately 50 plants per plot. All seeds were inoculated with “Guard-N®” N<sub>2</sub>-fixing bacteria (Verdesian, Cary NC USA) immediately before planting.

Phenotypic data measured in each plot consisted of main stem length, number of main stem nodes, number of pods per plant, total seed yield per plant in grams, and 100-seed weight in grams. Each measurement was conducted on four randomly selected plants per plot. Using results from the DAPC model described above, individual seed sources were assigned to one of five genetic clusters and phenotypic data was analyzed by cluster rather than individual seed source. For each phenotype measured in the field, a mixed linear model was fitted to the data using the lme4 package in R version 4.0.3. with genetic cluster included as a fixed effect and row-pair and column-pair as random effects (Bates et al., 2015; R Core Team, 2017). Field partition coordinates row-pair and column-pair were used as model terms due to high observed within-field variation, rather than replicate and block terms. A variogram was analyzed to

examine the data for spatial correlation, but as no correlation was found an autocorrelation was not implemented (Zuur et al., 2009). Least-squares means for mean cluster trait values were calculated using the 'emmeans' function in the emmeans package (Lenth, 2020); Tukey's Honestly Significant Difference (HSD) was performed using the 'cld' function in the multcomp package (Hothorn and Westfall, 2008) for pairwise comparisons between least squares means at the  $\alpha=0.05$  significance level. Phenotypic data was plotted in ggplot2 (Wickham, 2016).

Table 3.1 Passport data for all seed sources of ‘Jacob’s Cattle’, ‘Cal Early’ light red kidney, ‘Lina Sisco’s Bird Egg’, ‘Tiger’s Eye’, ‘Peregrion’, ‘Eclipse’, and ‘Lariat’ included in the study. ‘MA’ indicates Middle American origin.

Variety	Source Type	Source Name	Location	Center of Origin/Race
<b>Jacob's Cattle</b>				Andean/Nueva Granada
20-701	Commercial	Jacob's Cattle	Copake, NY	
20-702	Commercial	Jacob's Cattle	Lynden, WA	
20-703	Commercial	Wink's Jacob's Cattle	Nictaux, Nova Scotia	
20-704	Commercial	Jacob's Cattle	Quincy, WA	
20-706	NPGS-GRIN	Jacob's Cattle	Pullman, WA	
20-707	NPGS-GRIN	Jacob's or Dutch Cattle	Pullman, WA	
20-708	Seed Saver	Jacob's Cattle Gasless	Clinton, ME	
20-709	Seed Saver	Jacob's Cattle	Viroqua, WI	
20-710	Seed Saver	Jacob's Cattle Gasless	Arkansaw, WI	
20-711	Seed Saver	Jacob's Cattle	Eugene, OR	
20-712	Seed Saver	Jacob's Cattle Amish	Illinois	
20-713	Seed Saver	Deep Red Trout	Illinois	
20-714	Seed Saver	Mammoth Trout	Illinois	
20-715	Seed Saver	Coach Dog	Illinois	
20-717	Commercial	Jacob's Cattle	Exeter, ME	
20-718	Farm	Jacob's Cattle	Berwick, ME	
20-719	Commercial	Jacob's Cattle	Unknown	
<b>Other Traditional</b>				
Tiger's Eye	Commercial	Tiger's Eye	Unknown	Andean/Unknown
Jacob's Cattle Gold	Commercial	Jacob's Cattle Gold	Unknown	Andean/Unknown
Peregrion	Commercial	Peregrion	Unknown	Andean/Unknown
Lina Sisco's Bird Egg	Commercial	Lina Sisco's Bird Egg	Unknown	Andean/Unknown
Painted Pony	Commercial	Painted Pony	Unknown	Andean/Unknown
<b>Cal Early</b>				Andean/Nueva Grenada
20-721	Breeder	Cal Early	California	
20-722	Commercial	Cal Early	Idaho	
20-723	Commercial	Cal Early	California	
20-724	Commercial	Cal Early	Unknown	
20-751	Commercial	Cal Early	Idaho	
<b>Eclipse</b>	Commercial	Eclipse	Unknown	MA/Mesoamerican
<b>Lariat</b>	Commercial	Lariat	Unknown	MA/Durango

## ***Results***

After filtering, a subset of 1225 SNPs across 18 sources and 69 total individuals were utilized for genetic analysis within the ‘Jacob’s Cattle’ populations only. For analyses that included ‘Cal early’ check and University of Minnesota accessions, 9093 SNPs were used.

The number of SNPs within individual ‘Jacob’s Cattle’ sources ranged from 1 to 93. In general, the number of SNPs was substantially less within sources compared to the pooled ‘Jacob’s Cattle’ data (Table 3.2). Across all five ‘Cal Early’ sources, 624 SNPs remained after filtering, and within-source SNP counts ranged from 1 to 111. SNP counts within the four single-source heirloom varieties were overall lower than ‘Jacob’s Cattle’ or ‘Cal Early’, with the exception of ‘Peregion’. 72 SNPs were identified within ‘Lina Sisco’s Bird Egg’, 163 within ‘Jacob’s Cattle Gold’, 62 within ‘Tiger’s Eye’ and 2119 within ‘Peregion’. Substantially higher SNP counts within ‘Peregion’ are likely due to its identification as a seed mixture with two distinct seed phenotypes (Swegarden, 2015).

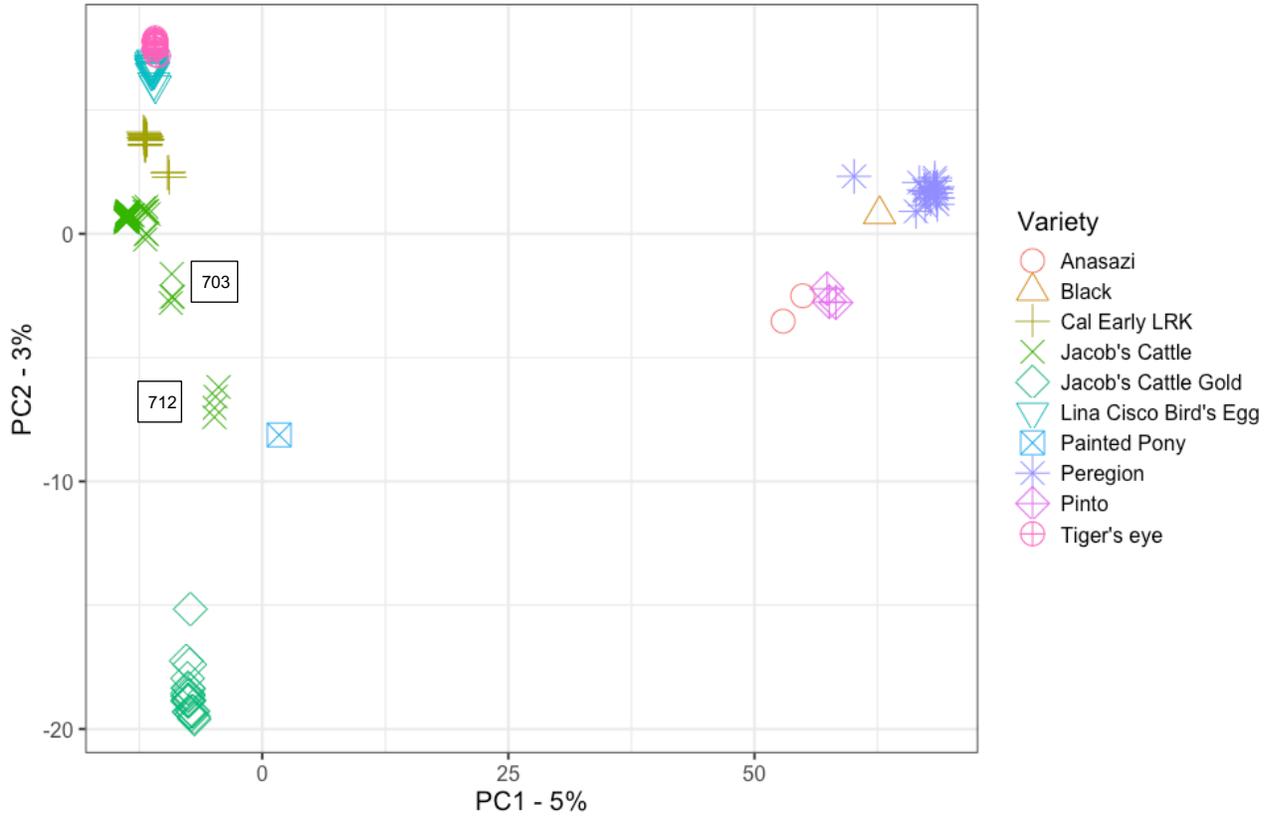
Table 3.2 Total SNP counts within varieties and seed sources. 18 distinct seed sources of ‘Jacob’s Cattle’, 5 distinct sources of ‘Cal Early’ light red kidney, and single sources of ‘Lina Sisco’s Bird Egg’, ‘Tiger’s Eye’, and ‘Peregrion’ were included in the analysis. SNPs were identified via Genotyping-by-Sequencing (GBS). Only individuals with multiple individuals genotyped are included in this table.

Variety	Source ID	Number of Individuals	Number of Seed Sources	Number of SNPs
Jacob's Cattle	all	70	17	1403
	20-701	4	1	9
	20-702	4	1	8
	20-703	4	1	54
	20-704	4	1	3
	20-707	4	1	10
	20-708	4	1	63
	20-709	4	1	1
	20-710	4	1	14
	20-711	4	1	8
	20-712	4	1	93
	20-713	4	1	15
	20-714	4	1	35
	20-715	4	1	8
	20-717	4	1	19
	20-718	4	1	13
20-719	4	1	8	
‘Cal Early’	all	19	5	624
	20-721	4	1	82
	20-722	4	1	111
	20-723	4	1	70
	20-724	3	1	1
	20-751	4	1	71
‘Lina Sisco’s Bird Egg’	single source	19	1	72
‘Jacob's Cattle Gold’	single source	20	1	163
‘Tiger's Eye’	single source	20	1	62
‘Peregrion’	single source	20	1	2119

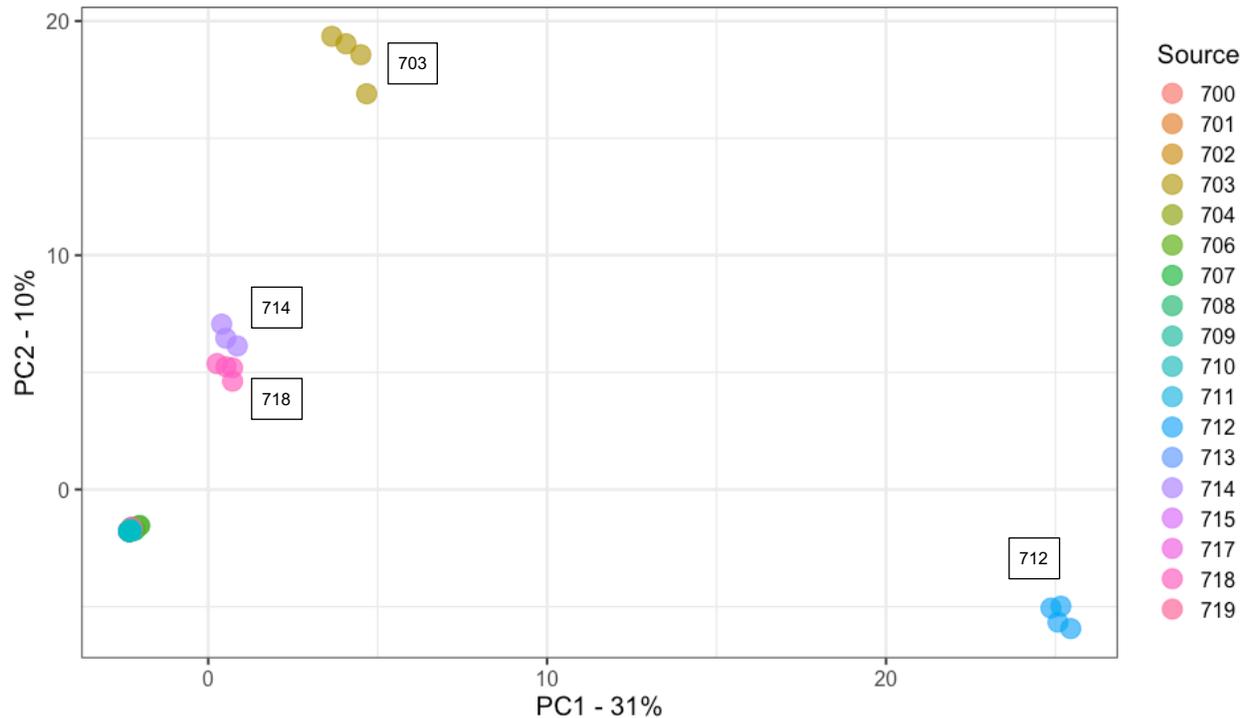
A Principal Components Analysis (PCA) of ‘Jacob’s Cattle’ with ‘Cal Early’ and other heirloom beans showed differentiation along Principal Component 1 (PC1), with two identifiable clusters. From the relative positions of known check varieties, genotypes associated with higher mean PC1 values can be assigned to the Middle American gene pool, while Andean genotypes are associated with lower PC1 values. PC1 accounted for 5% of total variance (Figure 3.1). PC2 primarily differentiated between sub-groups within Andean and Middle American gene pools. Within the Middle American cluster, ‘Peregion’ clusters more closely with ‘Eclipse’ than ‘Lariat’, indicating that ‘Peregion’ likely belongs to the Mesoamerican race. ‘Jacob’s Cattle’ shows greater overall differentiation between seed sources when compared to the five ‘Cal Early’ sources. Three out of four single-source heirloom varieties, ‘Tiger’s Eye’, ‘Lina Sisco’s Bird Egg’, and ‘Jacob’s Cattle Gold’, show little differentiation between individuals, indicating low diversity within single seed sources. The exception is ‘Peregion’, which is a Middle American variety that has been previously characterized as a mixture, and would therefore likely demonstrate higher diversity within source (Swegarden, 2015). The contextualization of ‘Jacob’s Cattle’ seed sources within a PCA of other traditional bean varieties and market class checks demonstrates the scale of divergence between ‘Jacob’s Cattle’ sources, with evidence of greater intra-varietal differentiation between some ‘Jacob’s Cattle’ sources compared to differentiation between distinct varieties.

Figure 3.1 a-b. Scatter plot of PC1 and PC2 values for Principal Component Analysis (PCA) of a) ‘Jacob’s Cattle’ (18 sources), ‘Cal Early’ (5 sources), 4 single-source heirloom varieties, and two market class checks, black and pinto, and b) ‘Jacob’s Cattle’ seed sources only. Color of data point indicates seed source origin. PCA of 18 JC seed sources were obtained using GBS data analyzed in TASSEL5 (Glaubitz et al. 2014)

a)



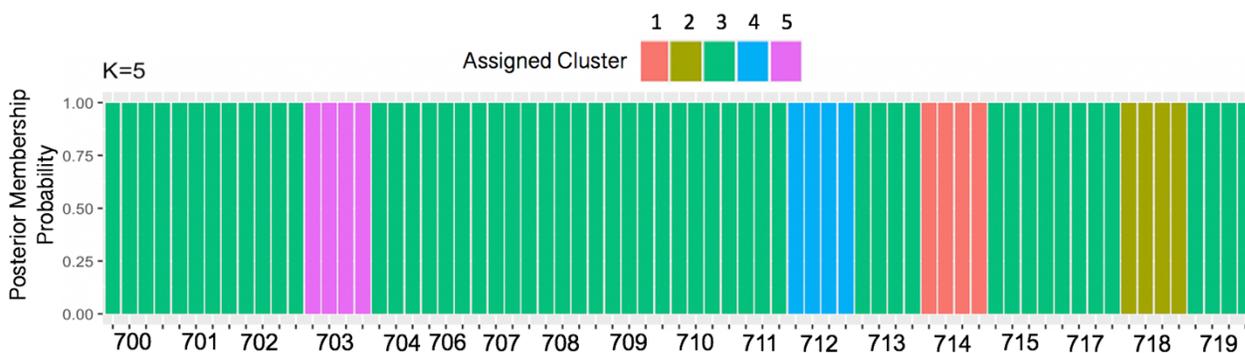
b)



A second PCA of 'Jacob's Cattle' seed sources further demonstrates differentiation within 'Jacob's Cattle' germplasm (Figure 3.1b). Fourteen seed sources tightly cluster in the lower left-hand corner of the plot, indicating little genetic differentiation. Four other sub-populations, however, show significant divergence between sources and a low level of divergence within source. A single seed source, 712, shows significant divergence across the PC1 axis which comprises 31% of total variance in the analysis.

Discriminant Analysis of Principal Components (DAPC) analysis resulted k=5 genetic clusters. Groups 1, 2, 4 and 5 each consisted of a single seed source comprised of four individual genotypes. Group 3 consisted of 14 seed sources comprised of 53 individual genotypes collectively. Genetic cluster assignments were subsequently used to explore significant differences between field phenotypes and visualize SNP distributions.

Figure 3.3. Results from Discriminant Analysis of Principal Components (DAPC) indicate  $k=5$  clusters across 18 seed sources using Bayesian Information Criterion (BIC) to assess goodness of fit. Results are displayed in a bar chart of posterior membership probability for each seed source across  $k=5$  genetic clusters. 4 individual genotypes for each seed source were included, except for source 706, for which only one individual remained after filtering.



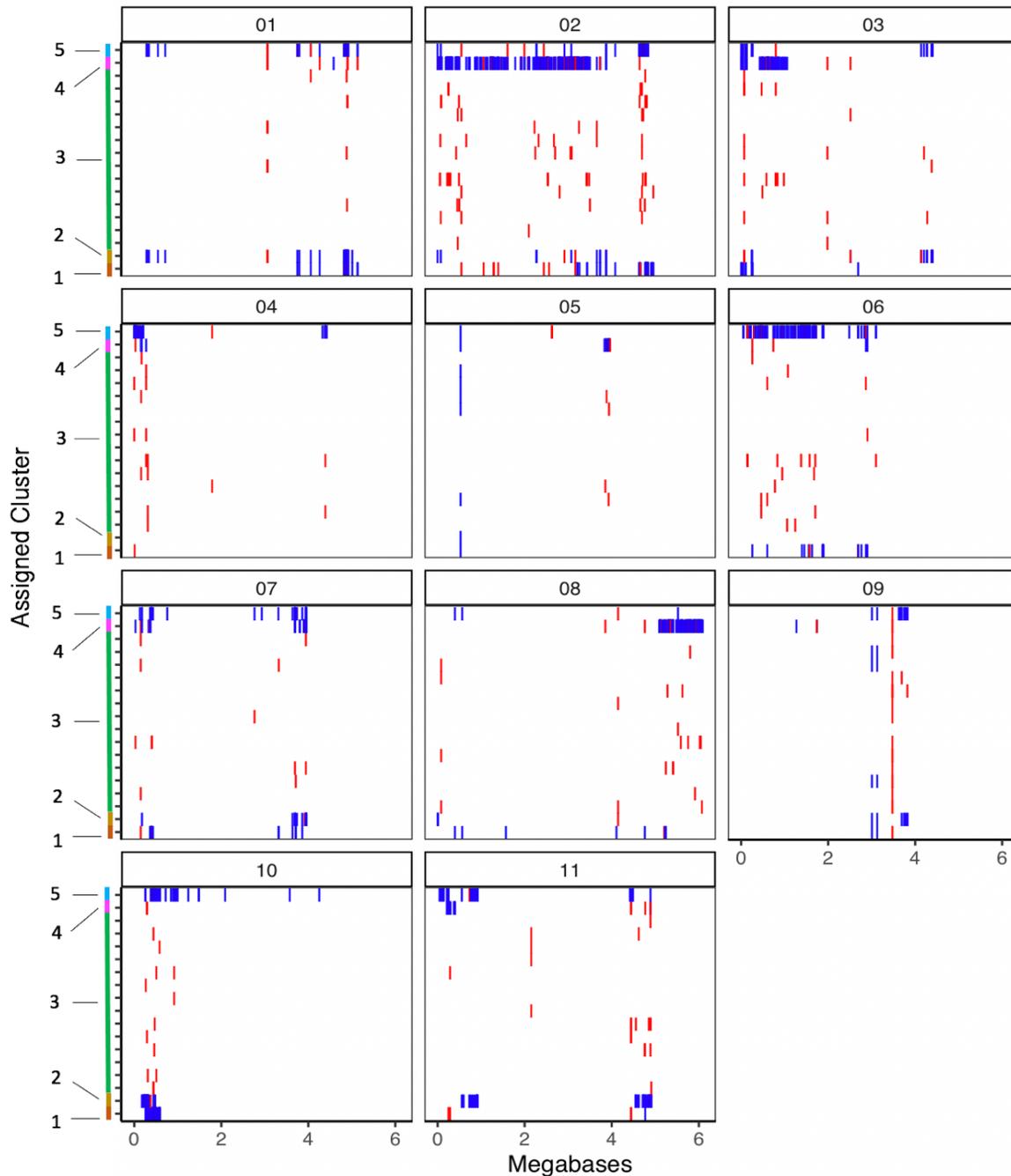
Distribution of minor alleles of individual seed sources, grouped by genetic cluster and plotted by physical location, show the presence of large minor allele haplotype blocks present in some genetic clusters, but not others (Figure 3.3a). In particular, clusters 4 and 5, which consist of seed sources 712 and 703 respectively, show particular haplotype divergence from the majority of seed sources represented by cluster 3. Clusters 1 and 2 also demonstrate some areas of divergent haplotypes, though blocks are less frequent and smaller in size. In the plot comparing minor allele frequency of five ‘Jacob’s Cattle’ clusters to heirloom and market class checks, minor allele haplotypes in clusters 4 and 5, which consist of sources 712 and 703 respectively, show several regions of similarity to ‘Anasazi’, black and pinto genotypes, which are of Middle American origin (Figure 3.3b). Minor allele distributions suggest that two ‘Jacob’s Cattle’ seed sources, cluster 4 and 5, likely outcrossed with a Middle American genotype. Cluster 4 displayed several large haplotype blocks shared with Middle American genotypes, particularly on chromosomes 2, 3 and 8. Cluster 5 displayed a large haplotype block on chromosome 6, with

smaller regions homologous to Middle American checks on chromosomes 1, 2, 7, 10 and 11.

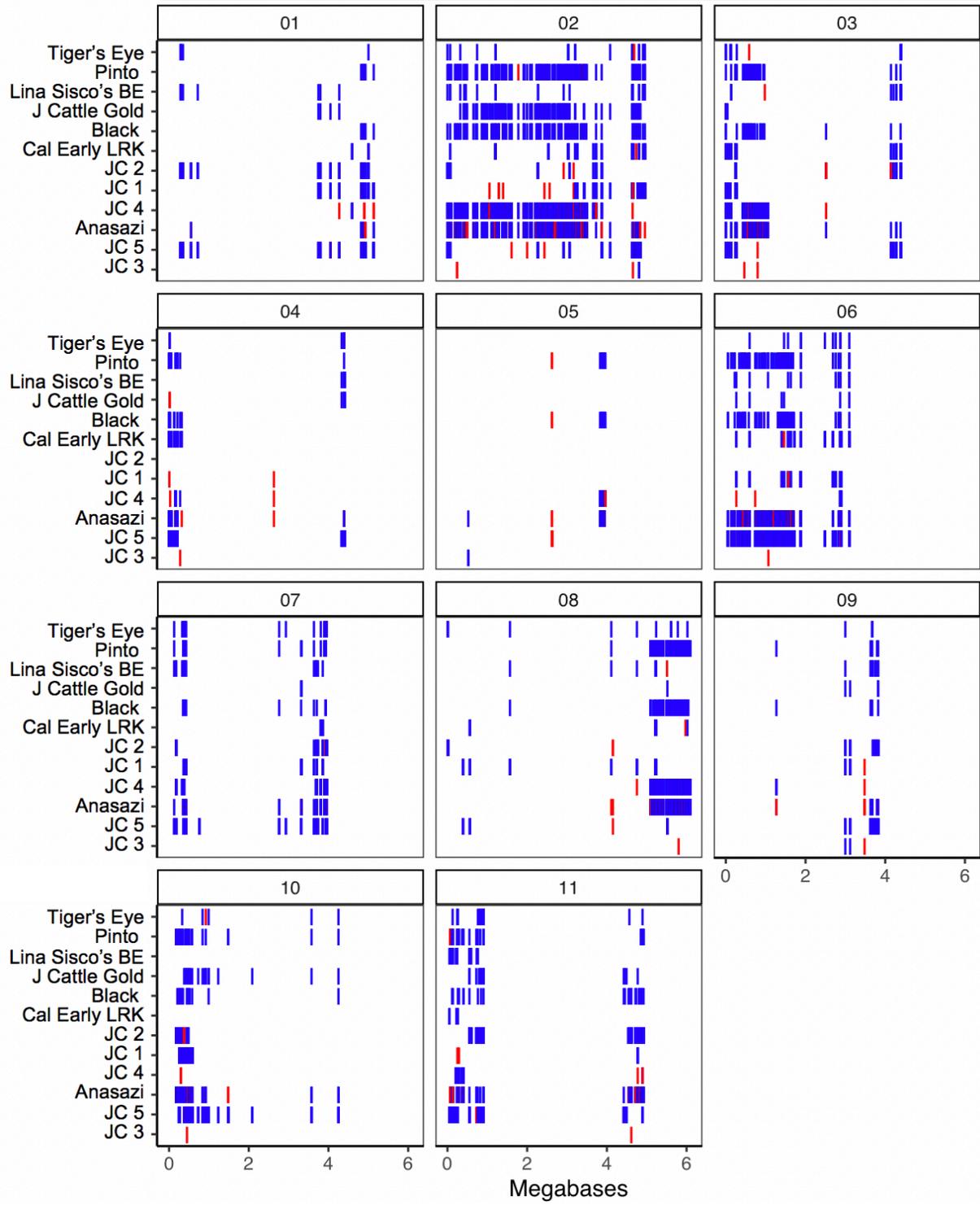
These regions of shared haplotypes likely explain divergence of sources 712 and 703 seen in the PCA of 'Jacob's Cattle' seed sources (Figure 3.1b).

Figure 3.4 a-b. Distribution of minor allele genotype at SNPs identified across chromosomes. Homozygous calls are displayed in blue and heterozygous calls are displayed in red, for: a) all Jacob's Cattle seed sources grouped by genetic cluster along y-axis; b) representative genotypes from each of five Jacob's Cattle clusters, heirloom and market class checks across all chromosomes. 'Lina Sisco's BE' represents 'Lina Sisco's Bird Egg', 'Cal Early LRK' represents 'Cal Early' light red kidney. Exact chromosome lengths vary, for individual chromosome lengths see *Phaseolus vulgaris* genome v2.1 (DOE-JGI and USDA-NIFA).

a)



b)



Field phenotypes demonstrated significant differences between k-means clustered group assignments for three out of five measured traits as determined using Tukey's HSD ( $\alpha=0.05$ ). In particular, clusters 4 and 5 had significantly smaller mean 100-seed weight than the three other clusters. Cluster 4 also had significantly longer mean stem length and higher mean pod number per plant (Figure 3.4). No significant differences were observed for mean plant yield. Diversity in seed shape, color and pattern were also noted, though these were not measured quantitatively (Figure 3.5).

Figure 3.5. Boxplots of phenotypic trait values of 18 ‘Jacob’s Cattle’ seed sources for 100-seed weight in grams (top left), plant yield in grams (top right), pods per plant (bottom left) and stem length in centimeters (bottom right), grouped by five assigned genetic clusters. Clusters denoted with the same letter are not significantly different as determined using Tukey’s HSD ( $\alpha=0.05$ ).

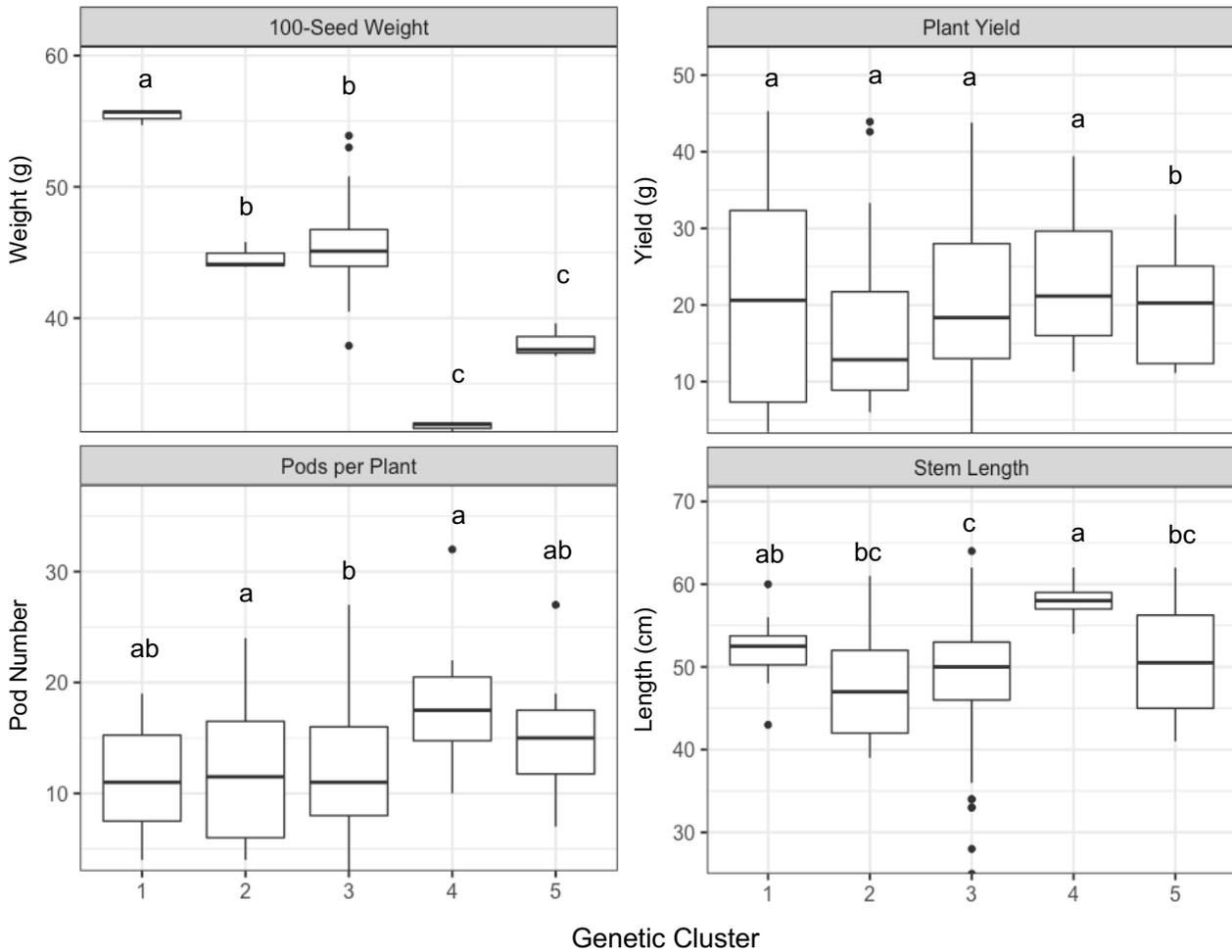
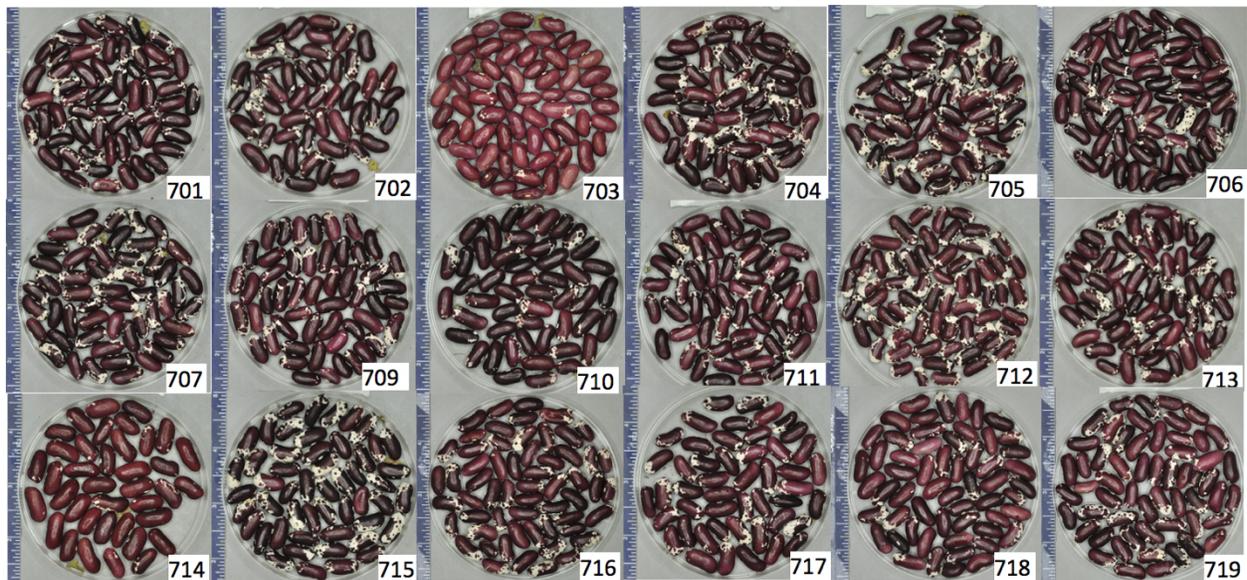
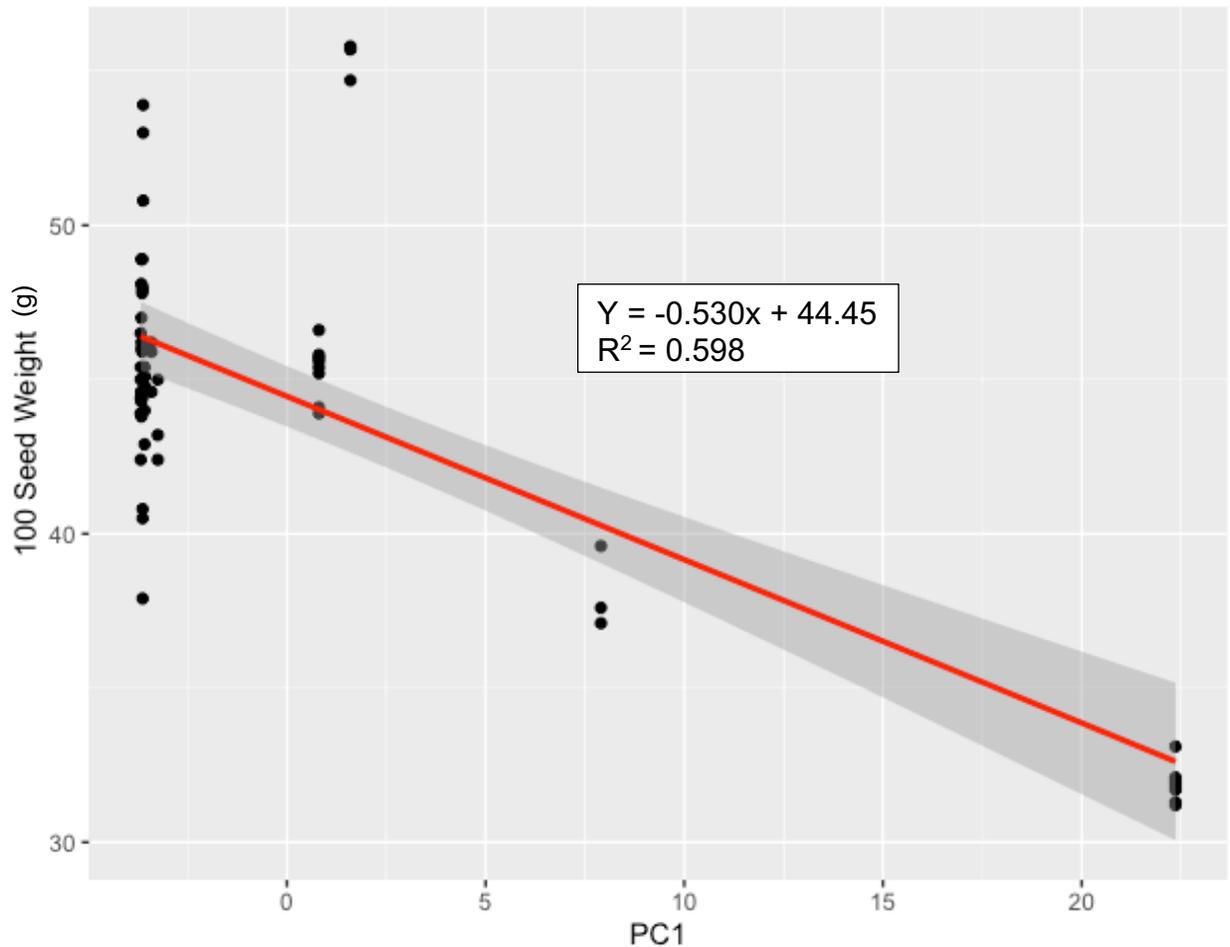


Figure 3.6 Seed phenotypes of each of the 18 seed sources of ‘Jacob’s Cattle’ included in the experiment, labelled by source number in bottom right corner of each panel. Differences in seed phenotypes can be observed.



A significant relationship between mean trait value for 100-seed weight and PC1 values for all genotypes was identified through regression, resulting in an  $R^2$  value of .60 and a p-value of  $1.73 \times 10^{-13}$ . This further supports the hypothesis that genetic divergence is a driver in observed diversity of measured phenotypic traits across ‘Jacob’s Cattle’ seed sources, and suggests that genetic divergence occurred due to historical introduction of Middle American genetic material into one or more ‘Jacob’s Cattle’ seed sources. This admixture could be the causal factor for phenotypic indicators of Middle American origin, including smaller seed size, higher pod number and greater plant height for cluster 4, which is represented by data points with highest PC1 values (Figure 3.6) (Singh et al., 1991).

Figure 3.7 PC1 values for all 18 ‘Jacob’s Cattle’ seed sources plotted against measurements for 100-seed weight phenotypic trait.



Overall, results indicate that observed genetic diversity is correlated to phenotypic diversity. Most notable were differences in 100-seed weight, pod number and stem length, and a high correlation between mean 100-seed weight and PC1 values. Along with the presence of large haplotype blocks distributed across the eleven chromosomes that appears to account for the majority of SNPs present, these findings indicate that observed genetic and phenotypic diversity within some seed sources is likely to be the result of outcrossing. Furthermore, it appears that some instances of outcrossing occurred with genotypes of Middle American races highly unrelated to ‘Jacob’s Cattle’, which is of Andean origin.

## *Discussion*

Polymorphism within ‘Jacob’s Cattle’ was largely observed between seed sources rather than within seed sources, indicating that genetic bottlenecks within networks of seed distribution (i.e. the exchange of very small amounts of seed), are likely commonplace. The 14 seed sources assigned to cluster 3 appeared to be highly related, which may suggest a relatively recent common seed source, or simply an absence of genetic divergence after seed sources were distributed. This cluster may offer limited potential for selection or adaptation when faced with contrasting environments. However, seed sources 712 (‘Jacob’s Cattle Amish’), 714 (‘Mammoth Trout’), 703 (‘Wink’s Jacob’s Cattle’) and 718 (‘Jacob’s Cattle’) demonstrated significant differentiation from each other and all other seed sources, forming distinct genetic clusters in the DAPC model. Sources 703 and 712 in particular show evidence of one or more historical outcross events with a Middle American genotype. While these sources still largely cluster with Andean genotypes in a PCA (Figure 3.1), several minor allele haplotype blocks that are shared between these sources and Middle American genotypes (Figure 3.3b) indicates that such a cross likely occurred. This outcrossing event may have been followed by seed saver selection back towards a ‘Jacob’s Cattle’-like phenotype that could explain the relatively low frequency of Middle American haplotypes in the current genotypes, as segregating material with Middle American traits were removed from the gene pool.

It should be noted that three out of four of the divergent seed sources were named distinctly, indicating an understanding that they represent a unique strain of the variety. However, other uniquely named strains ‘Coach Dog’, ‘Dutch Cattle’, ‘Deep Red Trout’, and ‘Jacob’s Cattle Gasless’ clustered with the majority of other seed sources. This indicates that naming of a variety or strain may or not capture meaningful strain diversity at the genetic level.

Two of the four divergent sources were obtained from the Seed Saver's Exchange network, the third from a regional farm-based seed company, and the fourth from a commercial farm in Maine. The two accessions procured from the NPGS-GRIN collection clustered with the majority of other sources in cluster 3. This finding that ex-situ germplasm conservation resources do not capture the extent of intra-varietal diversity maintained within *In-situ* seed saver networks is supported by similar comparisons in European seed systems (Negri and Tiranti, 2010; Enjalbert et al., 2011). 'Jacob's Cattle' sources also demonstrated significantly higher diversity between sources than did the commercial kidney variety, indicating that more centralized seed production systems typical for elite cultivars may not generate as much diversity with varieties, perhaps due to stricter controls on varietal purity and identity preservation (AOSCA, 2021).

More broadly, this case study indicates that, for traditional varieties being cultivated within decentralized seed saving networks, meaningful intra-varietal diversity is likely to be found at both genetic and phenotypic levels, even in autogamous crops such as common bean. However, local adaptation via allele and haplotype frequency change is likely contingent upon either an initial population with sufficient genetic variance, or the occurrence of an outcrossing event to generate such variance. Interestingly, past findings indicate that outcrossing in autogamous crops such as common bean increases when presented with a stressful environment, perhaps one evolutionary mechanism to initiate short-term crop adaptation (Klaedtke et al., 2017). Though beyond the scope of this study, it is important to note that phenotypic divergence within crop varieties in response to environmental conditions can also occur at the heritable epigenetic level, changes that in some cases contribute to local adaptation (Galloway, 2005).

Within the context of institutional plant breeding efforts in the Global North, germplasm maintained within decentralized stewardship networks comprised of home gardeners, farmers and freelance plant breeders is rarely recognized as a significant source of genetic diversity or adaptive traits, as for instance, compared to formal germplasm repositories (Almekinders 2000; Galluzzi et al. 2010). In part, strict commodity market class standards such as canning quality may play a role in disincentivizing use of novel germplasm in formal breeding programs (Kelly, 2010). In contrast, freelance plant breeders operate primarily in organic or low-input farming systems and emphasize distinct breeding goals such as adaptation to local environment and novel flavor and culinary quality (Deppe, 2021). Freelance plant breeders have made extensive use of seed saver networks such as Seed Saver's Exchange as well as regional farm-based seed companies in sourcing germplasm (Deppe, 2021).

### ***Conclusion***

Significant diversity across 'Jacob's Cattle' seed sources seems to have been generated by the introduction of new genetic material into stewarded populations via outcrossing, followed by subsequent selection by one or more seed savers over the course of population stewardship. Four out of eighteen seed sources exhibited significant genetic divergence, underscoring the important role that decentralized seed saving networks play in stewarding meaningful genetic diversity within traditional crop varieties.

These findings are meaningful for seed savers, farmers and plant breeders interested in accessing diverse genetic resources of traditional crop varieties. Aggregation of multiple varietal seed sources or strains from across seed saver networks may be a useful tool to facilitate regional selection and adaptation, especially given that individual seed sources show evidence of significant genetic bottlenecks and low within-source diversity. This strategy could constitute

“crowdsourcing” of intra-varietal diversity, alternately conceived of as a single varietal core collection. This strategy could be especially meaningful in cultural or geographical centers of origin, or when seeking to adapt a variety to a new environment. These results are also significant for formal breeding programs interested in using germplasm that has been managed *In-situ* for genetic improvement. Plant breeders may wish to take into account potential seed source divergence when evaluating germplasm in variety trials or for parental selection. Sources could either be evaluated by trialing side-by-side, or simply pooled in the hopes of increasing a baseline level of genetic diversity for future selection.

*In-situ* management of crop diversity within seed saver networks represents an understudied venue, and further study is needed to improve our understanding of how these complex and diffuse seed systems may influence crop adaptation to diverse environmental conditions. However, current evidence suggests that decentralized farmer and gardener seed saver networks may play an important role in protecting diversity and security of crop genetic resources, a key component of global food security.

## REFERENCES

- Aguilar, J., Gramig, G. G., Hendrickson, J. R., Archer, D. W., Forcella, F., and Liebig, M. A. 2015. Crop Species Diversity Changes in the United States: 1978–2012. *PLOS ONE* 10(8):e0136580.
- Almekinders, C. 2000. The Importance Of Informal Seed Sector And Its Relation With The Legislative Framework. GTZ-Eschborn.
- AOSCA. 2021. Seed Certification - Association of Official Seed Certifying Agencies. <https://www.aosca.org/programs-and-services/seed-certification/>.
- Bates, D., Maechler, M., Bolker, B., and Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1):1–48.
- Bellucci, E., Bitocchi, E., Rau, D., Nanni, L., Ferradini, N., Giardini, A., Rodriguez, M., Attene, G., and Papa, R. 2013. Population structure of barley landrace populations and gene-flow with modern varieties. *PLoS ONE* 8(12).
- Bhandari, B. and Gauchan, D. 2018. Intra-varietal diversity in landrace and modern variety of rice and buckwheat. *Journal of Agriculture and Environment*.
- Bradbury, P., Zhang, Z., Kroon, D., TM, C., Y, R., and ES, B. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23(19):2633–2635.
- Brouwer, B., Winkler, L., Atterberry, K., Jones, S., and Miles, C. 2016. Exploring the role of local heirloom germplasm in expanding western Washington dry bean production. *Agroecology and Sustainable Food Systems* 40(4):319–332.
- Ceccarelli, S. 1989. Wide adaptation: How wide? *Euphytica* 40(3):197–205.
- Cichy, K. A., Porch, T. G., Beaver, J. S., Cregan, P., Fourie, D., Glahn, R. P., Grusak, M. A., Kamfwa, K., Katuramu, D. N., McClean, P., Mndolwa, E., Nchimbi-Msolla, S., Pastor-Corrales, M. A., and Miklas, P. N. 2015. A *Phaseolus vulgaris* diversity panel for andean bean improvement. *Crop Science* 55(5):2149–2160.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R., Lunter, G., Marth, G., Sherry, S. T., McVean, G., and Durbin, R. 2011. The Variant Call Format and VCFtools. *Bioinformatics* 27(15): 2156–2158.
- Darby, H. and Cummings, E. 2018. 2017 Heirloom Dry Bean Variety Trial.
- Deppe, C. S. 2021. Freelance Plant Breeding. *Plant Breeding Reviews* 44(1): 113–186.

Döring, T. F., Knapp, S., Kovacs, G., Murphy, K., Wolfe, M. S., Döring, T. F., Knapp, S., Kovacs, G., Murphy, K., and Wolfe, M. S. 2011. Evolutionary Plant Breeding in Cereals—Into a New Era. *Sustainability* 3(10):1944–1971.

Enjalbert, J., Dawson, J. C., Paillard, S., Rhoné, B., Rousselle, Y., Thomas, M., and Goldringer, I. 2011. Dynamic management of crop diversity: From an experimental approach to on-farm conservation. *Comptes Rendus - Biologies* 334(5–6):458–468.

Galloway, L. F. 2005, April 1. Maternal effects provide phenotypic adaptation to local environmental conditions. *New Phytologist* 166(1): 93-100.

Galluzzi, G., Eyzaguirre, P., Negri, V., Galluzzi, G., and Eyzaguirre, P. 2010. Home gardens: neglected hotspots of agro-biodiversity and cultural diversity. *Biodiversity and Conservation* 19:3635–3654.

Gioia, T., Logozzo, G., Marzario, S., Zeuli, P. S., and Gepts, P. 2019. Evolution of SSR diversity from wild types to U.S. Advanced cultivars in the Andean and Mesoamerican domestications of common bean (*Phaseolus vulgaris*). *PLoS ONE* 14(1).

Glaubitz, J., Casstevens, T., Lu, F., Harriman, J., Elshire, R., Sun, Q., and Buckler, E. 2014. TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLoS ONE* 9(2): e90346

Howard, P. 2015. Intellectual Property and Consolidation in the Seed Industry. *Crop Science* 55:1–7.

Hufford, M. B., Berny Mier Y Teran, J. C., and Gepts, P. 2019, April 1. Crop Biodiversity: An Unfinished Magnum Opus of Nature. *Annual Review of Plant Biology* (70): 727:751.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24 (11):1403–1405.

Kelly, J. D. 2010. The Story of Bean Breeding: dry bean production and breeding research in the U.S. Michigan State University.

<http://www.hrt.msu.edu/pbgp/Links/plantbreedingintro.html>

Klaedtke, S., Caproni, L., Klauck, J., de la Grandville, P., Dutartre, M., Stassart, P., Chable, V., Negri, V., Raggi, L., Klaedtke, S. M., Caproni, L., Klauck, J., De la Grandville, P., Dutartre, M., Stassart, P. M., Chable, V., Negri, V., and Raggi, L. 2017. Short-Term Local Adaptation of Historical Common Bean (*Phaseolus vulgaris* L.) Varieties and Implications for In Situ Management of Bean Diversity. *International Journal of Molecular Sciences* 18(3):493.

Negri, V. and Tiranti, B. 2010. Effectiveness of in situ and ex situ conservation of crop diversity. What a *Phaseolus vulgaris* L. landrace case study can tell us. *Genetica* 138(9):985–998.

*Phaseolus vulgaris* v2.1, DOE-JGI and USDA-NIFA, <http://phytozome.jgi.doe.gov/>

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Seed Savers Exchange Heirloom Seeds. 2021. <https://www.seedsavers.org/>.

Singh, S. P., Gepts, P., and Debouck, D. G. 1991. Races of common bean (*Phaseolus vulgaris*, Fabaceae). *Economic Botany* 45(3):379–396.

Stone, S., Boyhan, G., and McGregor, C. 2019. Inter- and Intracultivar Variation of Heirloom and Open-pollinated Watermelon Cultivars. *HortScience* 54(2):212–220.

Swegarden, H.R. 2020. GBS of Heirloom Common Bean Pure Lines. NCBI – SRA Bioproject ID: PRJNA667092. <https://www.ncbi.nlm.nih.gov/bioproject/>

Swegarden, H. R. 2015. Selection of commercial and heirloom common bean (*Phaseolus vulgaris* L.) for organic production in Minnesota. University of Minnesota.

Swegarden, H. R., Sheaffer, C. C., and Michaels, T. E. 2016. Yield Stability of Heirloom Dry Bean (*Phaseolus vulgaris* L.) Cultivars in Midwest Organic Production. *HORTSCIENCE* 51(1): 8-14.

Thomas, M., Dawson, J. C., Goldringer, I., and Bonneuil, C. 2011, March 1. Seed exchanges, a key to analyze crop diversity dynamics in farmer-led on-farm conservation. *Genetic Resources and Crop Evolution* 58(3): 321-338.

Thomas, M., Demeulenaere, E., Dawson, J. C., Khan, A. R., Galic, N., Jouanne-Pin, S., Remoue, C., Bonneuil, C., and Goldringer, I. 2012. On-farm dynamic management of genetic diversity: the impact of seed diffusions and seed saving practices on a population-variety of bread wheat. *Evolutionary applications* 5(8):779–95.

Veteto, J. R. 2008. The history and survival of traditional heirloom vegetable varieties in the southern Appalachian Mountains of western North Carolina. *Agriculture and Human Values* 25(1):121–134.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

Wilker, J., Navabi, A., Rajcan, I., Marsolais, F., Hill, B., Torkamaneh, D., and Pauls, K. P. 2019. Agronomic Performance and Nitrogen Fixation of Heirloom and Conventional Dry Bean Varieties Under Low-Nitrogen Field Conditions. *Frontiers in Plant Science* 10.

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., and Smith, G. M. 2009. Violation of Independence – Part II. *Mixed Effects Models and Extensions in Ecology with R* p. 161–191.

## APPENDIX I

### NORTHERN ORGANIC VEGETABLE IMPROVEMENT COLLABORATIVE 2020 SQUASH VARIETY TRIAL REPORT

#### Trial Goal

This trial evaluated commercially available kabocha-type *Cucurbita maxima* winter squash. The goal of the trial was to evaluate currently available varieties as well as two newly released varieties for productivity, disease resistance, eating quality and storage ability. In our region, *C. maxima* squash are often severely afflicted by bacterial wilt brought on by striped cucumber beetle infestation.

Variety	Color	OP/Hybrid	Company
Marine Grey	Blue	OP	Uprising Seeds
Winter Sweet	Blue	F1	Johnny's Selected Seeds
JWS 14-4069 Pink Kabocha	Pink	F1	Johnny's Selected Seeds
JWS 17-4547 Green Kabocha	Green	F1	Johnny's Selected Seeds
Stella Blue	Blue	OP	Seed Revolution Now
Vanity	Blue	F1	Levantia
Silver Bell	Blue	OP	Restoration Seeds
Sunshine	Orange	F1	Johnny's Selected Seeds
Blue Kuri	Blue	OP	Adaptive Seeds



*Blue Kuri at harvest*

#### Trial Specifications

The trial was seeded on June 1st and transplanted on June 14<sup>th</sup> into raised beds with black plastic mulch. Plants were 24” apart in-row, and beds were 9’ apart to allow for harrowing between beds. Plants were covered with row cover from transplant until first flowers appeared to protect from cucumber beetles. Plants were watered with drip irrigation approx. once weekly. Trial was planted in an augmented design with 2 replicates and 4 blocks, with the check variety “Stella Blue” planted in all four blocks. All other entries were planted in two blocks. Squash were harvested on September 18<sup>th</sup> in order to avoid a killing frost, and some immature fruit were present at harvest. In particular, the two Johnny’s breeding lines had a higher proportion of

unripe fruit. Relatively low cucumber beetle pressure was observed in the trial, in contrast to many on-farm trials this season. Bacterial wilt and powdery mildew incidence was also low so was not recorded.

Fruit were cured at approximately 75 degrees and ambient humidity for two weeks, and stored at 50 degrees with ambient humidity. Fruit were sorted, weighed and counted 30 days after harvest and 60 days after harvest. Fruit quality data was taken on two individual fruits from each plot. Brix was calculated by freezing and thawing fruit samples and measuring with a handheld refractometer and dry matter was calculated by measuring fruit samples before and after dehydration. Brix and dry matter are commonly used to assess fruit quality in squash, with higher brix and dry matter values correlating to superior eating quality. In taste tests, fruit were also rated on a 1-5 scale for appearance, texture, sweetness, acidity, bitterness, and intensity of flavor.

## Results

All entries in the trial were selected to conform to the kabocha-type market class, though Stella Blue and Silver Bell included at least some individuals with kuri-type fruit. An attempt was also made to choose varieties with small to medium sized fruit. Due to differences in fruit size, yield based on fruit number vs. weight will be very different, so depending on market strategy that may influence how yield is measured (i.e. price per pound vs price per fruit), retail vs. wholesale, etc. All weights are in kilograms (=2.2 lbs). All figures are color-coded, with green highlight indicated the top performing entries for a given measured variable.

### Marketable Yield

Two entries, 'Marine Gray' and to a lesser extent 'JWS 17-4547 Green Kabocha,' had a higher proportion of fruit with scarring/scaling at harvest. Otherwise fruit quality was high with little rot, likely due to the dry growing season conditions in 2020. 'Vanity' and 'Stella Blue' had the highest average individual fruit size at roughly 2 kg, while 'JWS 17-4547 Green Kabocha' had the smallest fruit size at roughly .5 kg. Fruit photos are included at the end of the report. In general, Silver Bell, Sunshine F<sub>1</sub> and Vanity F<sub>1</sub> were highest yielding by weight. JWS 17-4547 Green Kabocha yielded significantly higher by number due to its much smaller fruit size. 'Blue Kuri' and 'JWS- 14-4069 Pink Kabocha' also yielded well by both count and weight. 'TotalNum' is the total number of fruit per plot, 'MktNum' is the marketable number of fruit per plot, 'MktWeight' is the marketable weight of fruit per plot in kg.

Table 1. Yield Performance and Fruit Size: Table of Means and Groups. Note that letters indicate entries which are not statistically different for a particular trait. For example, all entries with an 'a' are not statistically different for a given trait.

EntryName	TotalNum	MktNum	MktWeight	AvgFruitWght				
Blue Kuri	23.12	b	22.89	b	36.3	ab	1.61	a
JWS 14-4069 Pink Kabocha	22.62	b	22.39	b	35.3	ab	1.62	a
JWS 17-4547 Green Kabocha	70.38	a	68.11	a	35.1	ab	0.5	b
Marine Grey	16.16	b	15.48	b	17.75	b	1.21	ab
Silver Bell	29.16	b	26.98	b	41.05	a	1.52	a
Stella Blue	17.86	b	16.54	b	33.43	ab	2.02	a
Sunshine	21.6	b	21.25	b	37.84	ab	1.71	a
Vanity	17.05	b	17.16	b	37.79	ab	2.21	a
Winter Sweet	19.16	b	16.48	b	24.15	ab	1.43	ab

## Storage

'MktWht30DAH' and 'Num30DAH' denote percent by weight and number respectively of stored fruit that were marketable after 30 days of storage. The same measurements were taken at 60 days after harvest, denoted as 60DA

Most entries stored fairly well at the 30 day check, except for Sunshine which had significantly more loss to storage rot. 'Silver Bell' and 'Stella Blue' were the best storsers.

Table 2. Storage performance: Table of Means and Groups. Note that letters indicate entries which are not statistically different for that trait. For example, all entries with an 'a' are not statistically different for that trait.

EntryName	MktWht30DAH		Num30DAH		MktWht60DAH		Num60DAH	
Blue Kuri	87.15	ab	91.89	a	75.4	a	81.2	ab
JWS 14-4069 Pink Kabocha	85.33	ab	87.35	a	79.39	a	86.76	ab
JWS 17-4547 Green Kabocha	81.18	ab	87.15	a	52.49	ab	57.18	ab
Marine Grey	90.97	ab	92.49	a	61	ab	61.17	ab
Silver Bell	96.65	a	98.05	a	95.58	a	97.03	a
Stella Blue	94.43	a	93.3	a	87.46	a	86.61	a
Sunshine	60.83	b	60.07	b	29.73	b	37.77	b
Vanity	90.19	ab	93.62	a	90.47	a	92.28	a
Winter Sweet	88.7	ab	89.72	a	71.46	ab	76.19	ab

## Fruit Quality

In taste tests, ‘JWS 14-4069 Pink Kabocha,’ ‘Stella Blue’ and ‘Sunshine’ were the most preferred while ‘Marine Grey,’ ‘Silver Bell’ and ‘JWS 17-4547’ were the least preferred. ‘Marine Grey’ and ‘Silver Bell’ also had the lowest Brix and dry matter values, while ‘JWS 17-4547’ had a low Brix but fairly high dry matter. ‘Vanity’ had very high Brix and dry matter values, although it did not rank as highly in flavor ratings. Some varieties such as ‘Winter Sweet’ are advertised as tasting better after several months of storage, so possibly did not perform as well for that reason.

Table 3. Brix and Dry Matter of Harvested Fruit: Table of Means and Groups. Note that letters indicate entries which are not statistically different for that trait. For example, all entries with an 'a' are not statistically different for that trait.

EntryName	Brix		DryMatter	
Blue Kuri	10.03	ab	18.5	ab
JWS 14-4069 Pink Kabocha	11.85	a	20.54	ab
JWS 17-4547 Green Kabocha	9.25	ab	20.29	ab
Marine Grey	8.57	ab	14.34	ab
Silver Bell	6.43	b	11.91	b
Stella Blue	9.08	ab	16.89	ab
Sunshine	10.8	a	16.67	ab
Vanity	12.45	a	25.11	a
Winter Sweet	9.2	ab	16.51	ab

Table 4. Flavor Assessment of Harvested Fruit: Table of Means and Groups. Note that letters indicate entries which are not statistically different for that trait. For example, all entries with an 'a' are not statistically different for that trait.

EntryName	Appearance		Texture		Sweetness		Acidity		Bitterness		Intensity	
Blue Kuri	2.5	a	4.06	a	3	a	3.96	a	1	b	4.48	a
JWS 14-4069 Pink Kabocha	5	a	5.06	a	3.5	a	3.96	a	1	b	3.48	a
JWS 17-4547 Green Kabocha	2.5	a	4.44	a	2.5	a	1.04	a	2.5	ab	3.02	a
Marine Grey	3	a	5.08	a	2	a	2.54	a	3	ab	3.56	a
Silver Bell	4	a	3.08	a	1	a	1.54	a	3.5	a	5.06	a
Stella Blue	4	a	4.75	a	2.25	a	3	a	1	b	2.75	a
Sunshine	5	a	3.32	a	4.5	a	1.55	a	1	b	2.98	a
Vanity	2.5	a	3.8	a	3.5	a	1.97	a	1	ab	2.9	a
Winter Sweet	3	a	3.58	a	2	a	2.04	a	2.5	ab	2.06	a

**Representative fruit photos for each trial entry**

Vanity



Stella Blue



Silver Bell



Winter Sweet



JWS 17-4547 Green Kabocha



JWS 14-4069 Pink Kabocha



Blue Kuri



Stella Blue



## Sunshine



*This project is funded by USDA NIFA Award # 2018-51300-28430. More trial results @ [varietytrials.eorganic.info](http://varietytrials.eorganic.info)*

## APPENDIX II

### CORNELL SMALL FARMS PROGRAM *SMALL FARMS QUARTERLY* ARTICLES

#### What Can Variety Trialing Do for Your Vegetable Farm?

April 1, 2019

By Kristen Loria

**Put those glowing seed catalog descriptions to the test and find varieties that are the best fit for your own farming environment.**



*Evaluating basil quality and Downy Mildew resistance. Courtesy of Mazourek Lab*

A variety trial entails growing different varieties of a crop alongside each other in order to directly compare their performance across any number of characteristics. It can be highly controlled and scientific or very informal. Conducting a variety trial on your farm is a simple idea that can produce long-lasting benefits to your farm enterprise.

#### **Why do a variety trial?**

Each year new varieties crop up in your favorite seed catalogs — maybe your old standby variety

is no longer offered, or you have been dissatisfied with it and are looking for alternatives.

Perhaps your farmer's market customers or chefs are asking for a crop or crop type you haven't grown before. Catalog descriptions can give relevant information — for example, disease resistance attributes, days to maturity and beautiful photos, but they don't tell the whole story.

Much of these catalog descriptions might not reflect the conditions on your own farm, and here lies the main reason that it might be worth your while to conduct a variety trial on your farm: identifying what varieties perform best in your specific farming system and environment.

In addition, you might be looking for specific quality traits that are hard to glean from the catalog. For example, rainbow carrots might be really popular at your market stand, but you haven't been satisfied with the variety you have been growing. Perhaps you want to find out what varieties have the best flavor, or which store the best for your winter sales. Or perhaps you want a variety with strong tops that won't snap off when you are bunching them. Any of these traits could be the basis for a great on-farm trial, and you know best what traits are most important for your own farm.

Third, as our weather becomes more extreme and unpredictable, you might want to evaluate what varieties you will be able to count on when the weather gets weird. Although you can't control the weather during your variety trial, testing out which varieties perform best in stress from drought, flooding, cold snaps — whatever the season throws at you — could help your cropping systems be more resilient in future years.

#### **How to do a variety trial?**

First, you need to select what crop(s) you want to trial. Crop planning can be a time of lofty goals for the season ahead, but it's important to be realistic about your capacity and prioritize what to focus your efforts on. So, a good approach is to think about the biggest crop challenges

or frustrations you have had in past seasons and focus your trial efforts where they will have the biggest pay off for you.

Once you have picked the crops you will trial, you also want to figure out exactly what your goals are for the trial, and what you want to evaluate. Do you want to find a slicing tomato that won't crack when grown in the field, or a cilantro variety that won't bolt as fast in hot weather? Think about what you will need to do and when, all the way from seed to harvest in order to get the information from the trial that you need.

### **Picking your varieties**

#### 1. "Check" varieties

These are varieties that you already know are likely to perform a certain way — including the common "workhorse" varieties that you or other growers in your region count on. If your trial is focused on resistance to disease, that means including varieties that are likely susceptible as well as those that are resistant. This allows you to better draw contrasts between your varieties.

#### 2. Open pollinated vs. Hybrid

In many vegetable crops, older traditional varieties tend to be open pollinated (OP), while new varieties are either hybrids or open pollinated. If you are considering turning this into a breeding project down the line, selecting open-pollinated versus hybrid varieties will affect your choice of potential breeding material. *More on that in future issues!*

### **Designing your trial**



*Sweet pepper variety trial at Cornell University. Courtesy of Mazourek Lab.*

In general, when conducting a variety trial, you want each variety to get the same exact "treatment" so that all differences you observe are due to the variety itself. A few common strategies can help you do this, and get the most out of your trial that you can.

#### 1. Replication and Randomization

It's a good idea to do at least one replication to get a sense for how many differences

between the reps you observe. Randomizing the order of variety plots within a rep is another way to minimize environmental effects. Also try to orient plots to minimize the effect of variation in your field (wet pockets, rocky spots).

#### 2. Border rows

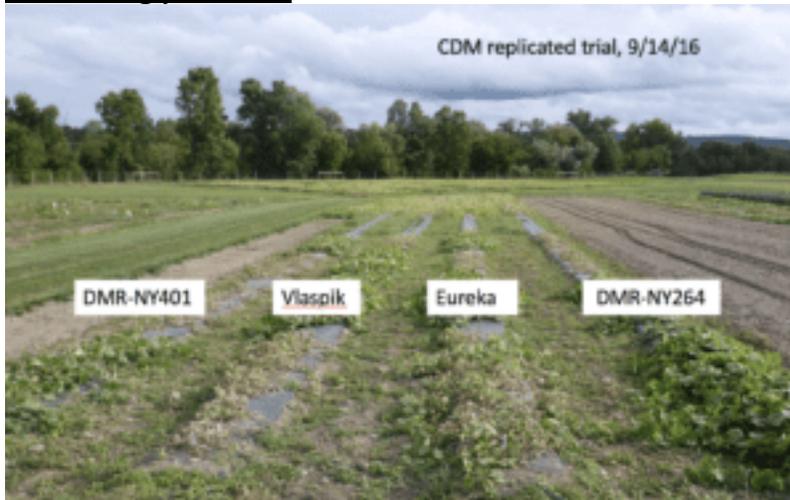
A well-known source of variation in a crop field is "edge effect," when environmental conditions are substantially different for the outermost rows, which can fudge your trial results. A good way to avoid this is to plant border rows surrounding your trial that you don't plan to evaluate.

#### 3. Make a map

To make sure you don't lose track of what's what, it's a good idea to overdo it when it comes to documenting your trial layout. You should label your plots in the field itself (most trials are

“blind” to minimize bias, which means assigning numbers to each variety rather than writing names on the stakes).

### **Evaluating your trial**



*Example of a downy mildew resistance trial in cucumbers. Sometimes the results are obvious! Courtesy of Mazourek Lab.*

The most important thing in evaluating an on-farm trial is to make sure you have a reliable method to get the information you want, while being realistic about your ability to spend time on the trial.

1. Keep good records

Keeping good records of the trial such as seed source, seeding or planting date, days to emergence, maturity and/or harvest etc. will be useful no matter what trait you are evaluating. If you just want to evaluate yield, you can record yields whether they are weekly fruit harvests or a bulk root harvest. If you are looking for disease resistance, you probably want to use a scale for rating disease severity, and also collect yield data.

2. You're the boss

There's no need to go overboard and take data you don't really need. Write down some notes, flag plots that you like at different stages of growth, whatever you think will be useful to your future self.

### **References and More Resources:**

The Grower's Guide to Conducting On-Farm Variety Trials. Organic Seed Alliance. 2018.

[Trials and Selection Webinar](#). Organic Seed Alliance. 2017.

[Variety Trial Planning Worksheet](#). Organic Seed Alliance.

[Trial Evaluation Worksheet](#). Organic Seed Alliance.

## **On-Farm Plant Breeding Pt. I: Getting Started with Diversity**

By Kristen Loria  
July 1, 2019

### **Learn how to get started with your own on-farm plant breeding project.**

Often as growers we rely on our favorite seed companies to provide us with the varieties we grow. This strategy usually works well, and there are a lot of great varieties out there both new and old. However, by their nature most modern commercial varieties (we will focus on vegetables, but this applies to other crops too) were developed to perform well across as broad a geographic area as possible, with a focus on the major commercial production regions for that crop. These varieties may not perform so well in non-target climates.

Breeding investment also tends to focus on the most commonly grown crops, so if there is an obscure crop that's important on your farm (i.e. ground cherries or watermelon radish), odds are the commercially available options will be more limited — and the opportunity will be ripe for an on-farm plant breeding project. Major seed corporations also emphasize hybrid varieties in some crops, which are much more expensive and must be re-purchased every season.

Also, modern varieties are almost exclusively bred in conventional systems where synthetic fertilizer, abundant water and at least some chemical controls were used. Because environmental conditions on organic farms tend to be very different in terms of nutrient availability, weed competition and other stressors, organic growers especially find that conventionally bred varieties don't perform as well or as reliably on their farms.

All of these factors are common reasons that farmers and gardeners decide to try their hand at plant breeding. Plant breeding on the farm allows us to develop varieties that fit our specific farm environments, with the traits that we most value. However as useful as that can be to our business, it's also important to realize that plant breeding is a long-term endeavor and takes additional time and attention in the growing season, so it's probably going to be hard to sustain unless it's an activity that you feel curious and passionate about.

### **Getting Started**

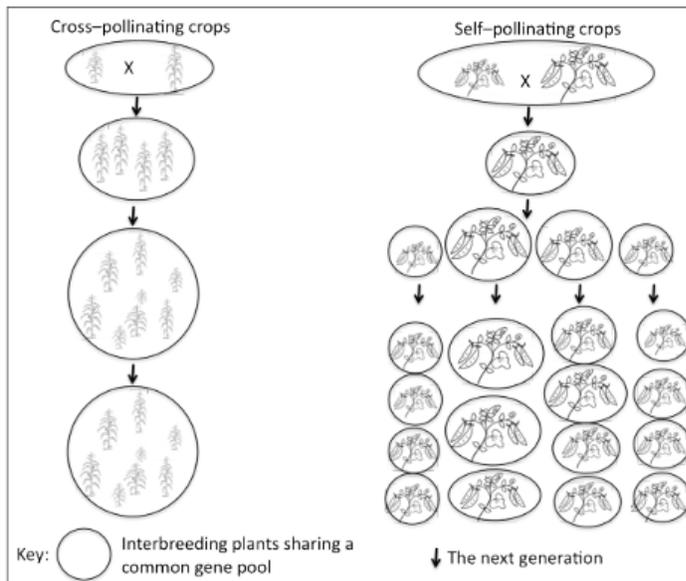
The best plant breeding projects arise when you are already familiar with a crop, the traits that are important (at least to you) for that crop, and different varieties that are available. Unless you already have that familiarity, it's worthwhile to spend some time trialing varieties and observing them closely. Through that process you often will notice interesting differences or traits that might spark an idea: "I love the fruit quality of this winter squash, but I'd prefer a bush growth habit," etc.

It might already be clear to you that an important part of plant breeding is seed-saving! Saving seeds from crops to replant the following year used to be a part of every farm, but it's much less common now. Some crops are very easy to grow seed for — if you are a bean or corn grower, you already are a seed grower. Other annual fruiting crops, like tomatoes, squash and peppers, are quite straightforward. Some crops, such as kale, beets and carrots, are biennial and need more than one season to flower and produce seeds. Make sure you know what will be required in order to grow seed for the crop you want to work on and think about whether this process will realistically be able to fit into your farm systems.

## Crop Pollination Strategies

The pollination strategy of your crop makes a big difference both in terms of what type of population you will be starting with, and the strategies you will use to improve it. A quick internet search will tell you whether your crop is an out-crosser (cross-pollinating) or an inbreeder (self-pollinating), if you don't already know. Pollination strategy is a spectrum and some crops like peppers and squash often do both.

If you start with a cross-pollinated crop (i.e. corn), you can assume that plants planted near each other will freely intermate. This is helpful if you want to select within an existing variety or create a diverse population by intermating several different varieties, but it limits your ability to maintain "pure" varieties in a field without large amounts of space to isolate them. Another useful thing to look up is isolation distances for your crop — how far apart you need to plant two outcrossing plants if you don't want them to cross. Often pollen from out-crossers is carried by wind, so distances can be large. If you have limited space, consider picking a crop that you would be comfortable with growing just one variety or population, or perhaps pick a self-pollinating crop to work on instead. Some backyard plant breeders ask to grow plots at neighbor's properties to achieve isolation distances!



*Difference in population structure of self-pollinated vs. cross-pollinated crop. Courtesy of the Organic Seed Alliance.*

Highly outcrossing species also suffer from inbreeding depression, which means that if forced to self-pollinate or intermate repeatedly with very genetically similar plants, vigor and overall quality of the variety will decrease due to loss of genetic diversity (genes become more homozygous). To avoid inbreeding depression, it's important to grow a sufficiently large number of plants in your breeding plot to maintain diversity within the crossing

population. Some outcrossing species exhibit self-incompatibility, which means they can't create viable offspring by self-pollination. This is an evolutionary strategy to ensure genetic diversity and can affect breeding approaches.

Self-pollinating crops (i.e. beans, peas) are unlikely to outcross even if grown right next to each other. That means that they will naturally form what breeders call "pure lines" — each plant self-pollinates each generation to create a very genetically uniform and true-breeding line. Plants will be homozygous for most or all genes. Modern varieties of inbreeding crops are usually pure lines, which means they were often descended from just one high-performing plant. Older varieties such as heirlooms or landraces are less likely to have descended from a pure line, so you might see more diversity between different plants in your variety. With inbreeding crops, it is more likely that you will have to make an initial cross in order to introduce enough variation into the population to select from.

## Picking Your Parents

It's important to remember that in order to achieve gain from selection (i.e. improve the performance of your variety over time), you need to start out with a genetically diverse population. In particular, you want a population that has genetic variation for the trait or traits that you are most interested in improving. This means that either you find a variety or population that already has variation within it, or you make a cross between two varieties to introduce new variation. That genetic variation is the buffet of traits that you select from throughout the project, and if most or all of your plants in a field are already genetically very similar, you won't be able to select plants that are any better than the rest.



*Perennial Kale Grex, an example of an initial population with high variation for numerous traits. Courtesy of the Experimental Farm Network.*

One complication is that some of the variation we see between plants in the field is due to environmental effects rather than genetic differences (phenotype = genotype x environment), so in order to make sure you are looking at a genetically diverse population of plants, you will want to know something about where that population came from. You will be much more effective at selecting for traits that are highly heritable, rather than those that are more influenced by environmental conditions. In the initial vetting of parent varieties, the easiest way to

determine if the variation in phenotype you see is heritable is to plant multiple replications of your variety plots.

There are several sensible starting populations for a plant breeding project on the farm.

- Use an existing open-pollinated variety, especially an heirloom or landrace, which has obvious variability in the trait/s you are interested in.
- Make a cross by taking two varieties that show variability in the trait/s you are interested in, and either cross them by hand (self-pollinators) or interplant them in a plot (cross-pollinators).
- Use a commercial hybrid, as these are the first generation offspring ( $F_1$ ) of two distinct parents — you can think of the seed company as already having made your initial cross for you. If there is a hybrid variety that you like (and there are no intellectual property restrictions on it, see below) you will grow out the offspring and see segregation for whatever traits those two parents differed by. This works for both cross-pollinated and self-pollinated crops.
- Some smaller seed companies are starting to sell “breeder’s grex” populations, which just mean a population created by intermating several different parent varieties. These are

sold with the express intention of being a plant breeding “starter kit” — there is sure to be plenty of diversity present, and you get to select what you like best for your own farm ecosystem.

You can use any seed to start a plant breeding project, unless there are plant patents or “bag tags” that restrict this use. Odds are you will get your parents from seed catalogs that you are already using. Other sources can be seed saver organizations, the Open Source Seed Initiative, small seed companies that often focus on unusual or regionally adapted varieties, or the USDA-ARS Germplasm Resources Information Network.

## On-Farm Variety Improvement Part II: Selection and Plant Breeding

By Kristen Loria

October 7<sup>th</sup>, 2019

*This article is a follow up to “On-Farm Plant Breeding Pt. I: Getting Started with Diversity”, published in the Summer 2019 SFQ. Starting there will provide more context to understand the following information.*

### What is Selection?



*An example of breeding of cucumbers with selection for fruit shape and quality. Mazourek Lab / CALS.*

Selection is the basis of planting breeding. When we talk about selection as plant breeders, we are referring to artificial selection — we walk into our field and advance a subset of plant individuals (by saving their seeds/tubers/cuttings and replanting them) that meet our objectives. Looking back through time, this is how crops were domesticated and improved — farmers picked out the plants that

yielded the best, had large edible parts, weren't bitter, hairy or poisonous, etc. This is in contrast to natural selection or “survival of the fittest,” which of course happens in natural ecosystems, but also happens in agricultural fields too. The ability for you to practice selection in your field depends on genetic diversity that is expressed in the plant's phenotype. Ways to generate that diversity was the focus of Part I in this series.

In summary, you generated enough diversity to select upon by making a cross between two different varieties, planting a hybrid variety without any intellectual property restrictions, or identifying an open-pollinated variety or population that already displays variation for the traits you want to improve.

### Helpful terms:

**Allele:** one of several versions of a gene; different versions have different functions or are non-functional. Usually caused by a genetic mutation at some point in the crop's history.

**Heritability:** the proportion of observed (phenotypic) variation that is genetically determined, and therefore heritable by offspring.

**Heterozygous:** Two different alleles for a gene present in plant; one copy of each.

**Homozygous:** Two copies of the same allele are present.

**Segregation:** In a population, segregation refers to phenotypic variation resulting from the random assortment of alleles into pollen/egg cells, creating offspring with different phenotypes.

For example, some offspring of a heterozygous purple-podded pea plant produce purple pods and some produce green pods.

### What will you select for and how?

A very important step in the process is deciding what your goal is: what trait or traits will you select for? And if there is more than one, which will you prioritize? Another very important step is ensuring that the trait you want to improve is **heritable**. That means that the variation you observe is mostly due to genetic variation (which you can select for) rather than environmental variation (which you can't). Heritability is a relative concept — it's value changes across time and space. Examples of traits that generally have low heritability includes yield, seedling vigor and susceptibility to disease or stressors such as drought. Traits that tend to be more highly heritable include fruit or root color and shape, days to flowering or protein content of grain, or plant architecture. This doesn't mean that you can't make any improvement selecting for low heritability traits on your farm, just that it will take longer and environmental variation (and strategies to control for it) will play a more important role.

A related concept is simple vs. complex traits, which refers to how many genes are responsible for a trait or phenotype. For example, yield is generally based on many, many genes while flower color is often based on one gene. Disease resistance can be simple or complex, depending on the disease and resistance mechanism.



*There is clear diversity within an open-source carrot population. Courtesy of Open Source Seed Initiative.*

The next question to ask is how you will evaluate your trait. It is a trait you can easily walk through the field and identify? Or will you need to collect data – perhaps by cooking and evaluating flavor, measuring yield or doing a storage trial? Make sure you know how you will evaluate and determine superior plants. Another issue is when evaluation of a trait can only occur after a crop has been harvested — as

with fruit quality evaluation in winter squash. That means you will have to harvest the whole crop (and possibly store it) before you can evaluate and identify superior plants. All of these factors will influence the logistics and time commitment of your project.

### Selection Strategies

In general, the goal is to improve the traits that you care about, whatever they may be. You will do this by selecting the plants you like best and discarding (or selling, or eating) those you don't. In commercial cultivars, stability of a variety is very important — breeders strive for nearly complete uniformity from one plant to another, from one season to the next. Commercial growers often value this uniformity because they know exactly what they will be harvesting, when it will be mature and how much it will yield. To get a stable variety, you generally have to practice more intense selection repeatedly over many seasons to “weed” out any unwanted variation.

An important thing to remember is that selection inherently means reducing the genetic diversity of your starting population. This is fine, but it's also worth considering that the more diverse a population or variety, the better the potential for future adaptation or improvement. In cross-pollinating crops, maintaining a minimum population size is also necessary to prevent inbreeding

depression. Therefore, it may be worth preserving genetic variation that doesn't have any negative impact on what you are trying to accomplish. For example, variation in baby kale leaf shape makes for a more interesting salad mix!

Formally trained plant breeders have many strategies for practicing selection, from old-school to genomic. Here are some fundamental strategies that are practical for on-farm plant breeding without an excessive amount of labor, time or infrastructure. Over several seasons, you may use multiple strategies at different stages of the process.

### **Mass Selection**

The most common method of selection on the farm is mass selection, which simply means selecting the best plants in a plot or field and saving their seed separately to plant out in the next generation. This can be done either by “roguing” out poor performers (negative selection), or by picking the best performers (positive selection). This method works especially well with outcrossing species as it's easy to maintain the larger population sizes you need, and the genetics of the pool will slowly improve as your selected plants continue to intermate over time. The disadvantage of this method is that you are selecting based on single plant characteristics, which are inevitably affected by environmental conditions. In mass selection of outcrossing crops, you also don't know who your “pollen parent” is, so you are only selecting based on traits of the mother plant.

You will gain more from mass selection in cross-pollinated crops if you remove inferior plants from the field before pollination occurs. Otherwise those inferior plants will contribute pollen to future generations, decreasing gain from selection.

Mass selection is also used to maintain and improve varieties after they have been developed. For example, if you have an heirloom corn variety that you grow every year, you might save ears separately from the best plants each year to use as your seed for the next year, rather than taking your seed from the entire field's harvest.

### **Family Selection**

Family selection entails first selecting single plants (as in mass selection) but seeds from individual plants are then saved in separate bags and the following year, each bag is sown in a single row. Each row constitutes a “family.” For a self-pollinated crop, each plant in a row will have the same mother and father plant (“S1 family” for self-pollination), while for cross-pollinated crops a row can either be a “full-sib” family if it was produced by a controlled cross between two different plants, or a “half-sib family” if mother plant was allowed to freely cross-pollinate, resulting in plants with the same mother but different father.

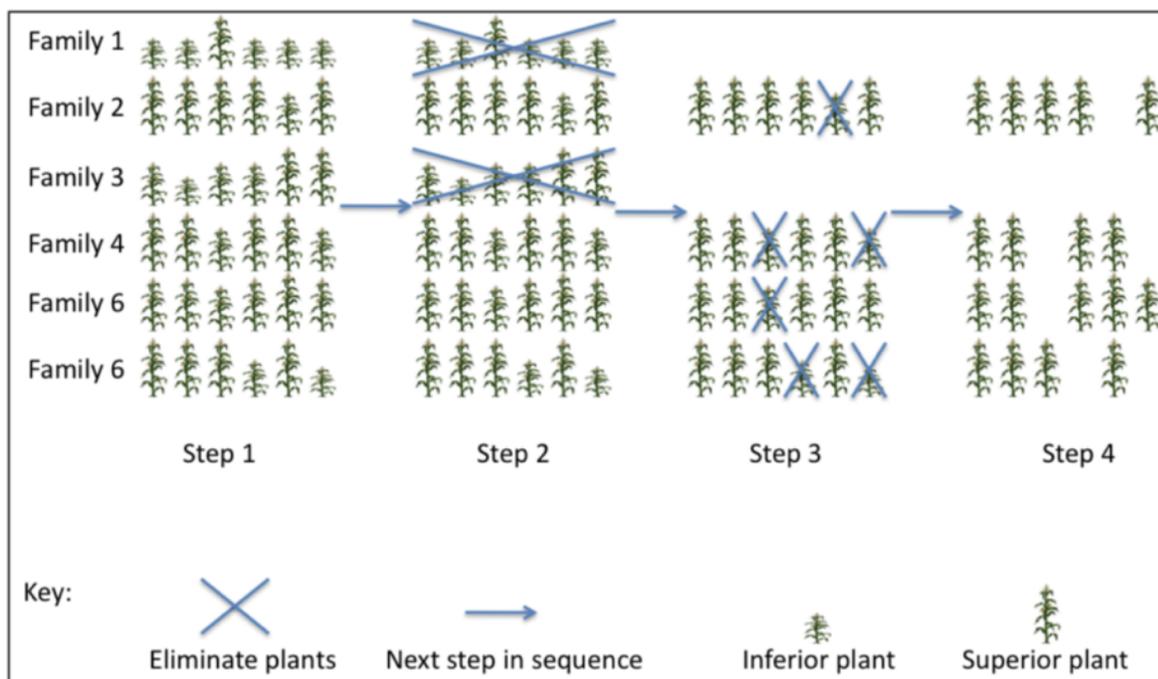
You can then advance or eliminate entire families together, as well as eliminating plants within families. Then aggregate the seed from all selected plants in selected families, plant it out in a row again next year and repeat the selection process. Because there is more relatedness and therefore genetic similarity within rows than between them, you can more easily sort through the environmental noise and identify plants that are genetically superior. While this method takes more work and space, it is useful in minimizing environmental effects that can greatly affect the phenotype of a plant.

Again, with cross-pollinated crops, this scheme works best if you physically remove inferior families or plants before pollination occurs. At the end of several generations of family selection, the winning families are ultimately combined to form a final population.

Family selection is also useful in self-pollinating crops for eliminating recessive traits in a **segregating** population that you don't want. For example, I have a bean population with purple pods (a dominant trait), some **homozygous** and some **heterozygous**. If I grow out my bean offspring in family rows, I will see which rows are 100% purple-podded, which tells me that there are no **recessive alleles** hidden in that family. I then select those rows and get a true-breeding purple podded bean.

### Thinking about Environmental Effects

As discussed earlier, some traits are more affected by environmental conditions than others. Selecting for a trait that is known to be highly heritable will likely go far in allowing you to successfully improve your crop without worrying too much about environmental effects. For this reason, it's a good idea to take into account field conditions when you are selecting — either use the gridded selection method to select the best plants from different sectors of the field, or use your knowledge of the field and plot layout to minimize environmental effects. Another option if you are doing family selection is to plant multiple replications of your family plots in different places (divide up each bag of seed amongst three different smaller plots). That will help illuminate whether a family is really standout, or if it just happened to be planted in an advantageous location. The first article in this series discusses replication in trialing.



*Selecting on a grid is a strategy for managing environmental variation across a field that affects phenotype. Courtesy of Organic Seed Alliance*

### What Next?

This series is intended to be a quick crash course in on-farm plant breeding, which has ideally gotten you hooked and left you wanting more. There are lots of wonderful and comprehensive resources available that can guide you more deeply into the intricacies of genetics, plant breeding methods, and more. Below is a list of where you might start looking.

Another important topic not covered here is seed saving. As methods for this vary from crop to crop, we recommend taking advantage of the many resources available on this subject.

**Sources for more Info:**

Breeding Organic Vegetables: A Step-by-Step Guide for Growers. Rowen White and Bryan Connoly. Published by NOFA-NY.

<https://eorganic.info/sites/eorganic.info/files/u2/BreedingOrganicVegetables-2011.pdf>

Breed Your Own Vegetable Varieties. Carole Deppe. Chelsea Green Publishing: 2000.

The Organic Seed Grower: A Farmer's Guide to Vegetable Seed Production. John Navazio. Chelsea Green Publishing: 2012.

**Organic Seed Alliance Resources:**

Introduction to On-Farm Plant Breeding. 2017.

[https://seedalliance.org/wp-content/uploads/2017/04/introduction\\_to\\_on-farm\\_organic\\_plant\\_breeding.pdf](https://seedalliance.org/wp-content/uploads/2017/04/introduction_to_on-farm_organic_plant_breeding.pdf)

The Grower's Guide to Conducting On-Farm Variety Trials. 2018.

<https://seedalliance.org/publications/growers-guide-conducting-farm-variety-trials/>

**Extension.org:**

Seeds and Seed Production in Organic Farming Resource Guide

<https://articles.extension.org/pages/59459/seed-and-seed-production-in-organic-farming-systems>

## APPENDIX III

### DRY BEAN GROWER NEEDS ASSESSMENT

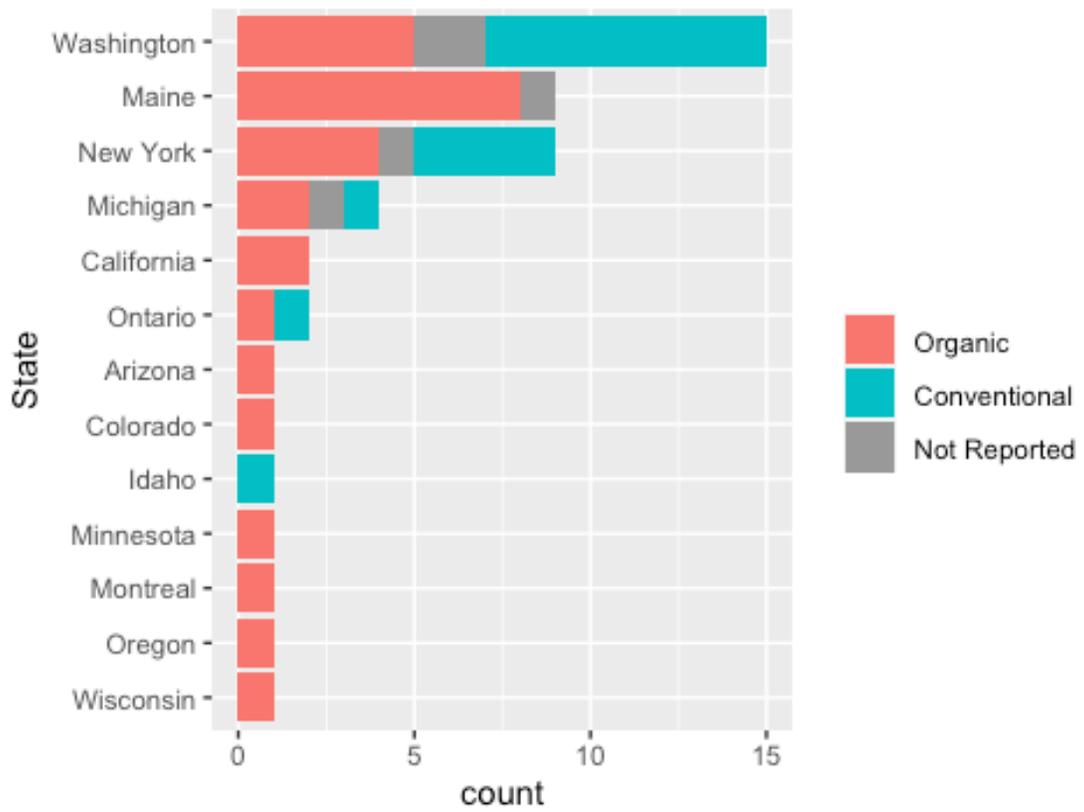
#### ***Survey Information***

The survey was online-only and was created using Qualtrics software. Both English and Spanish versions were made available. Survey questions were distributed through Cooperative Extension networks, public bean breeding programs, organic farming associations and bean growers' associations. The survey was intended to collect information from a wide range of operations across the continental US and Canada. Data subsets allowed targeting of specific grower groups, in particular those growing heirloom type dry beans vs. large scale organic producers.

#### ***Location and Types of Farms***

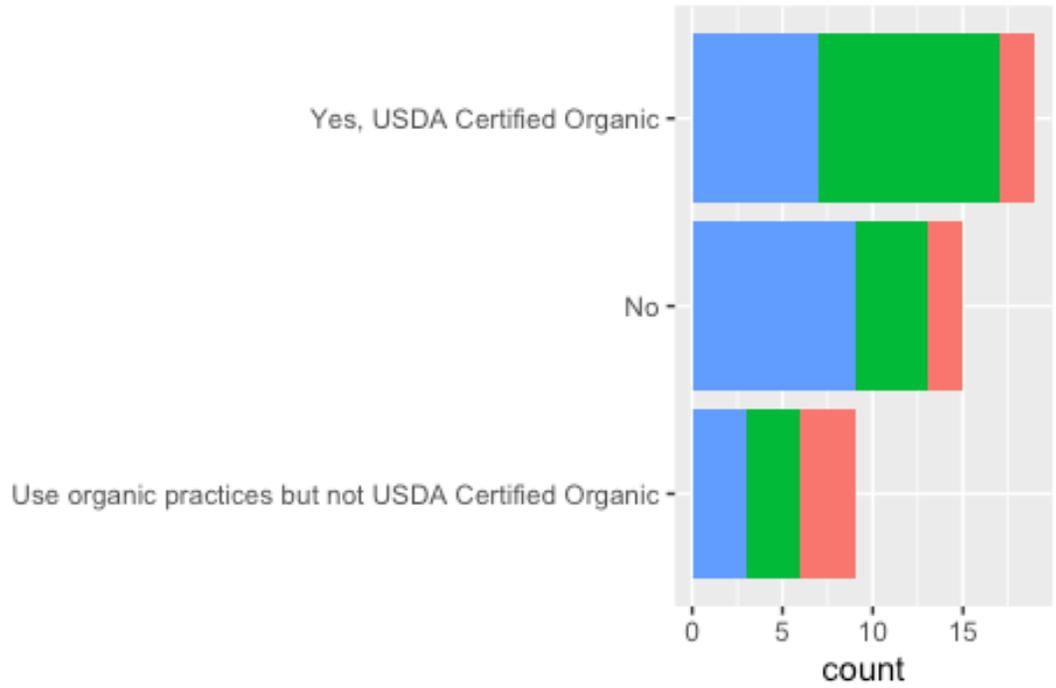
51 unique responses were collected from dry bean growers (organic or conventional) across the United States and Canada. Respondents included 17 Northeastern growers, 26 Organic growers (certified or not) and 15 growers producing heirloom types for commercial sale. Scale ranged from very small <1 acre to several thousands of acres. All respondents growing heirloom varieties were small-scale, though some were mechanized.

## Respondent location

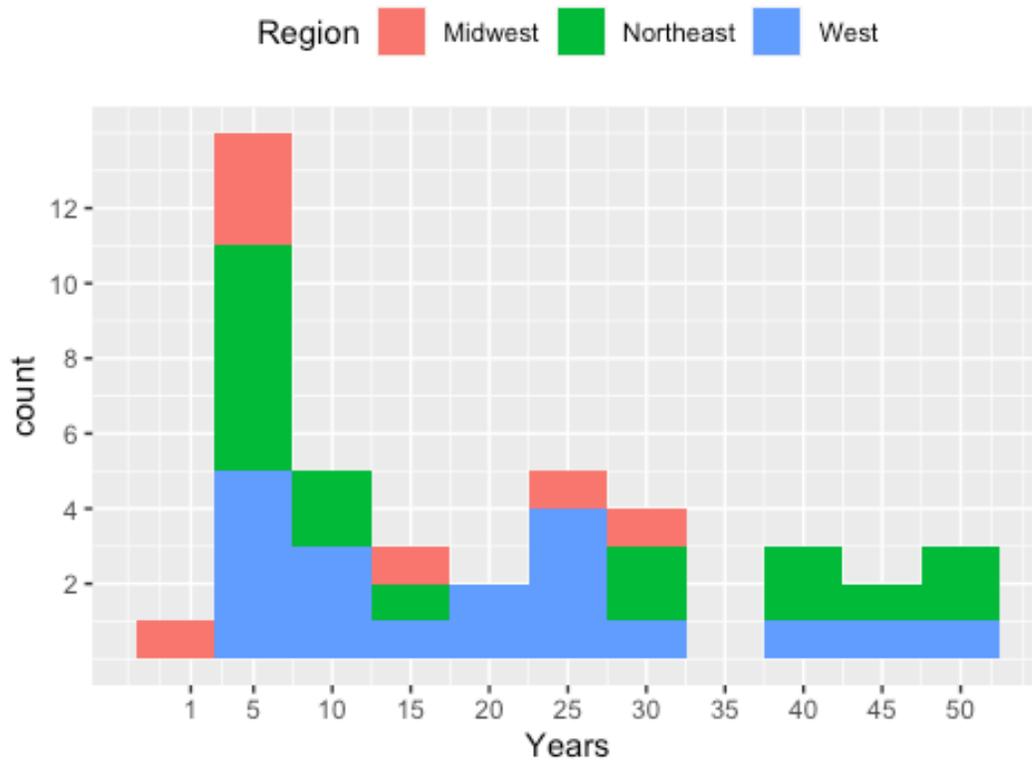


# Farm type

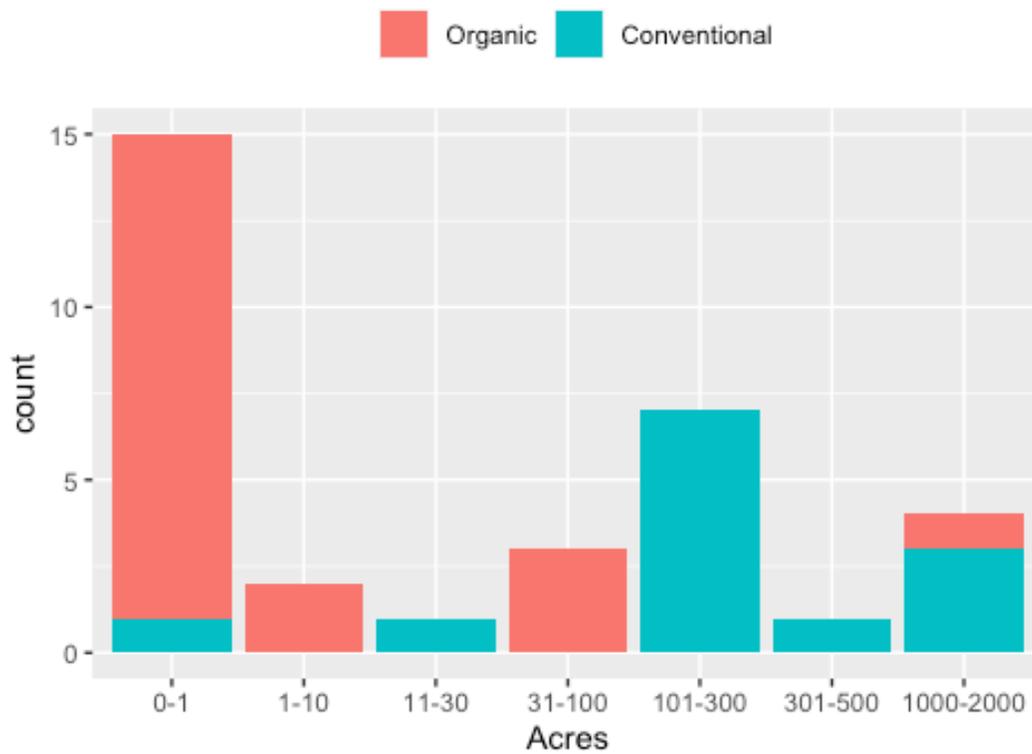
Region ■ Midwest ■ Northeast ■



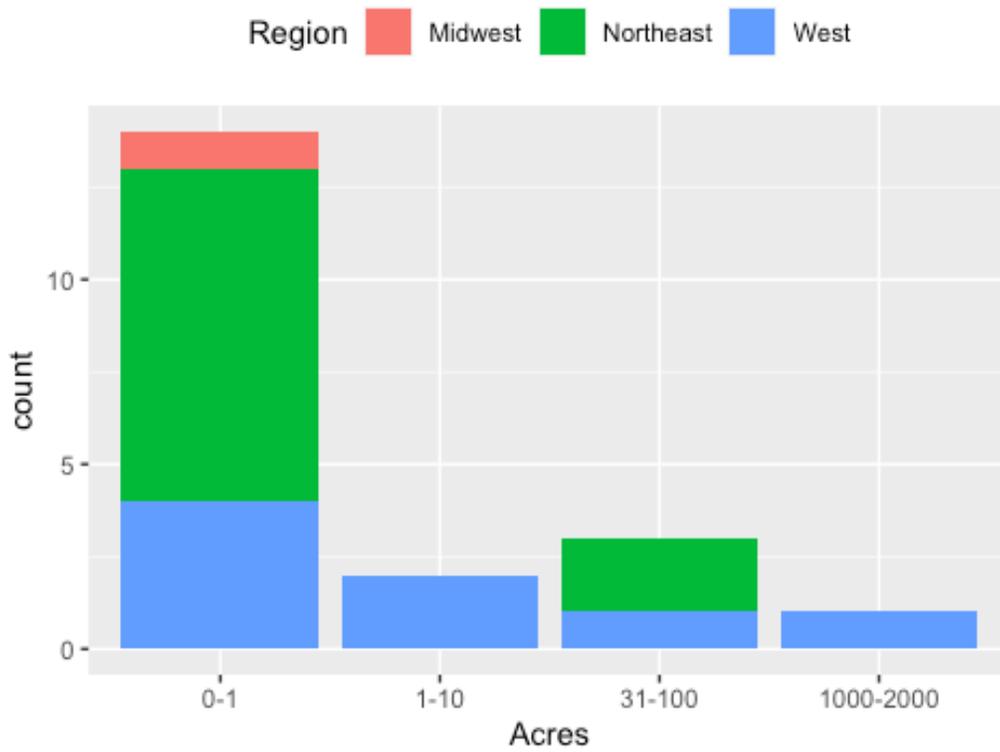
## Years Farming



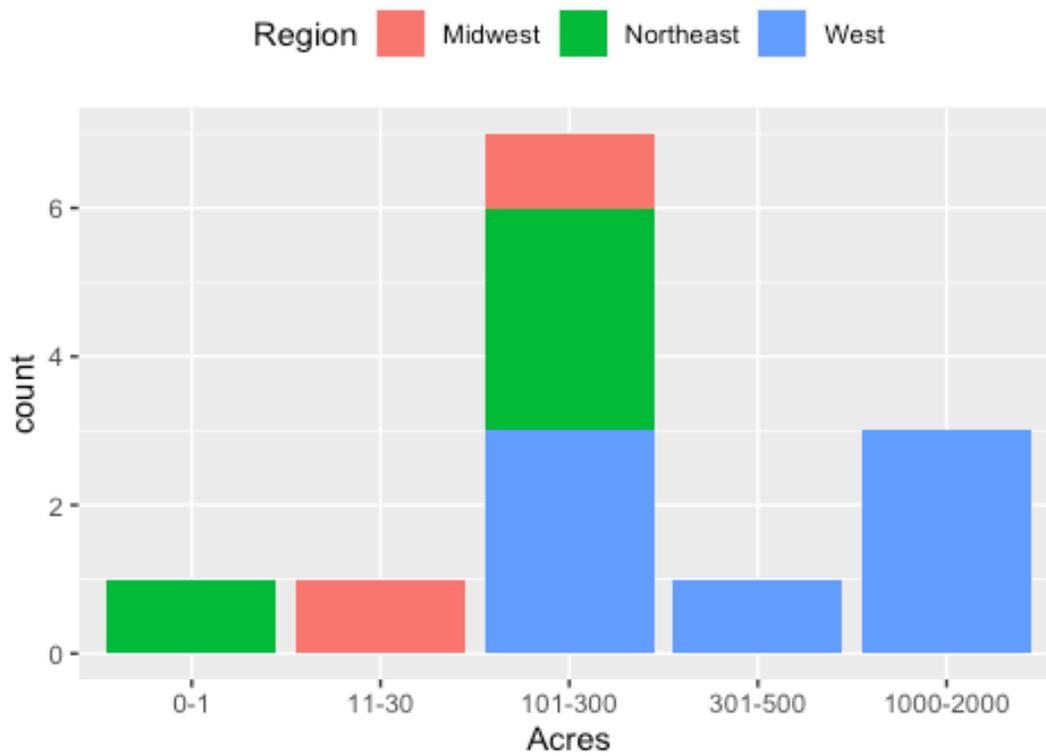
## Acreage of Respondent Farms



### Organic Acreage by Region



### Conventional Acreage by Region



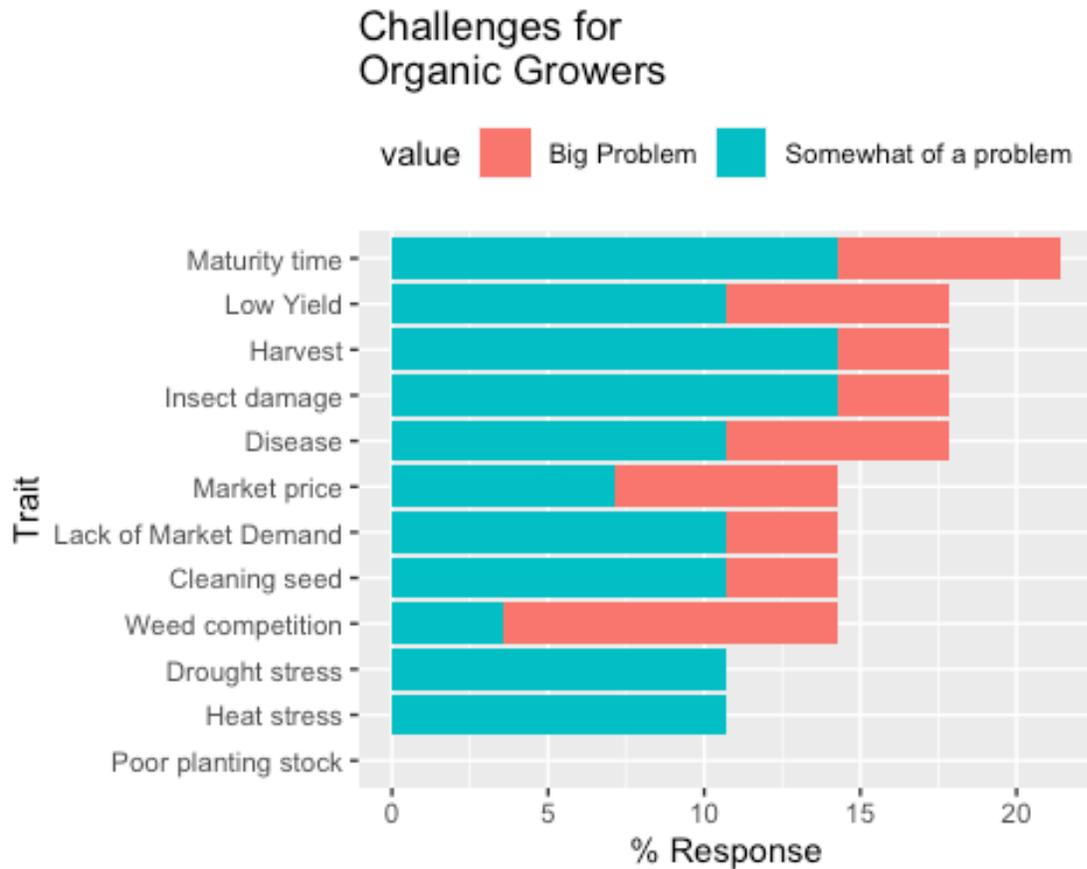
## Survey Results

### Organic Growers

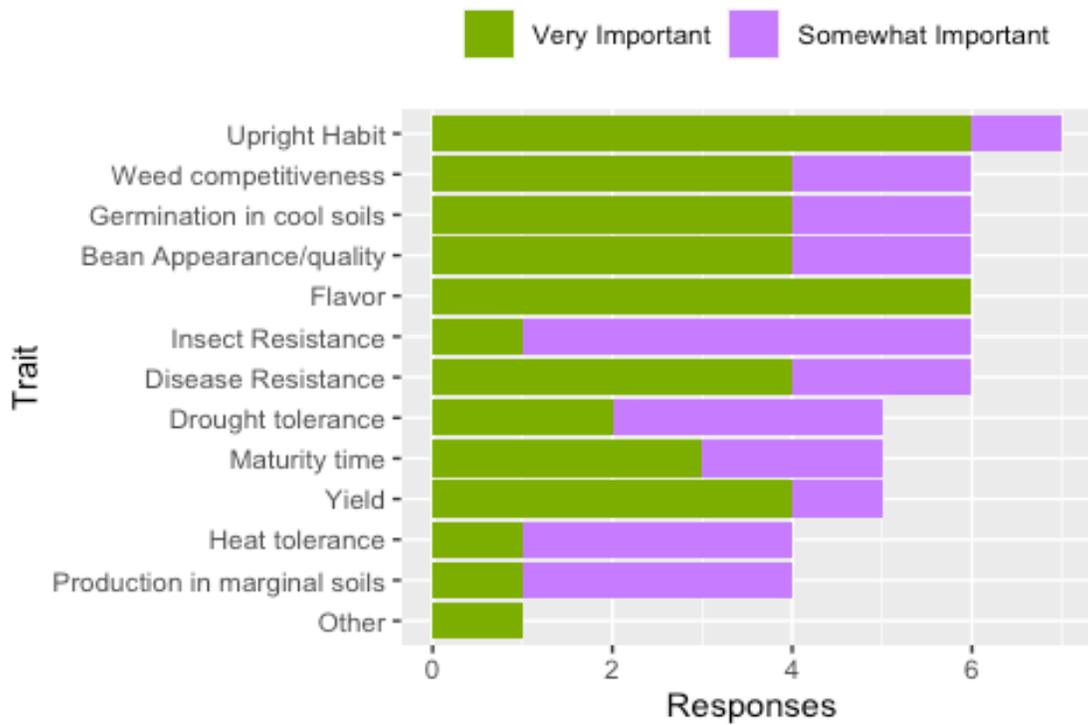
Top challenges for organic growers included issues with crop maturation, timely and successful harvest, low yields and insect damage to crop. Priority traits included upright growth habit, weed competitiveness, germination in cool soils, appearance/quality, flavor, and insect and disease resistance.

The most problematic disease for organic growers was white mold, followed by BCMV, rust and root rots. The most problematic insecters were Mexican Bean Beetle and Potato Leafhopper.

The most cited research need for organic bean growers was variety development and improvement, followed by agronomic research. Specific suggestions included insect and disease resistance as well as management strategies to mitigate damage. Other suggestions included equipment scaled to help mechanization and labor saving on small to medium sized operations.



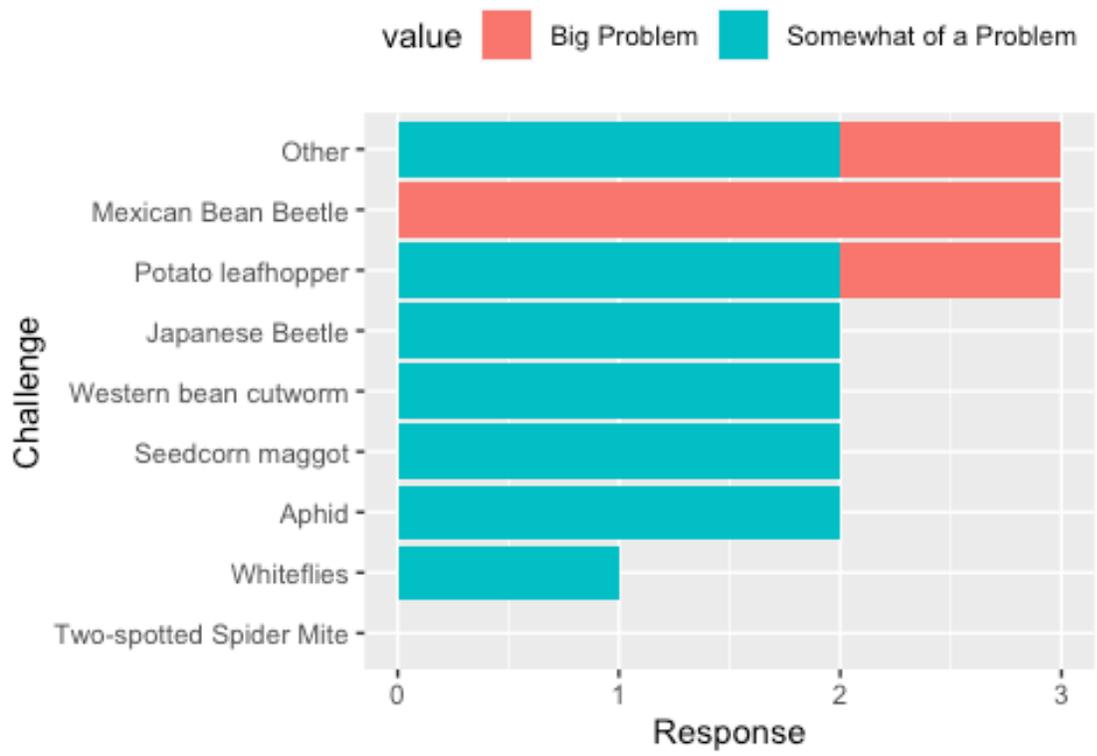
## Trait Priorities for Organic Growers



## Most Problematic Diseases for Organic Growers

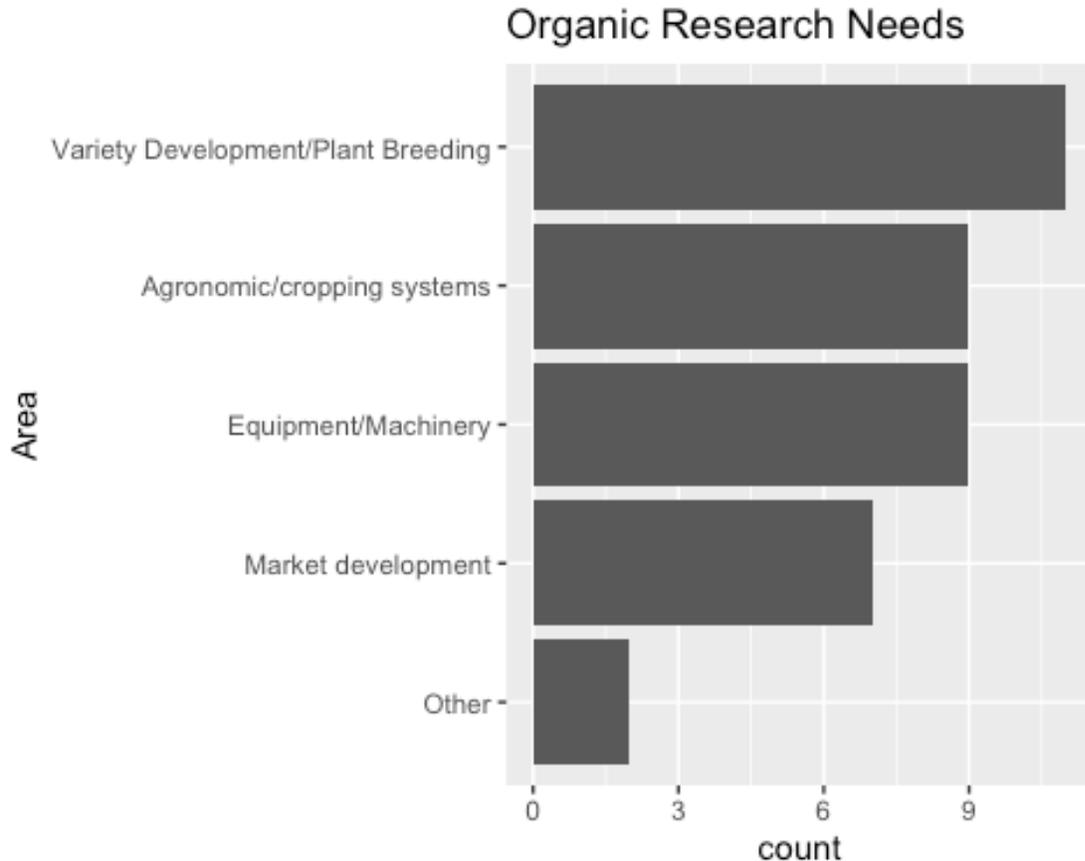


## Most problematic insect pests for Organic Growers



### 'Other' Responses

- Big Problem - Cucumber beetle
- Somewhat of a Problem - deer
- Somewhat of a Problem - Flea beetle
- Somewhat of a Problem - Brown snail



#### Elaborate: Organic Grower Needs

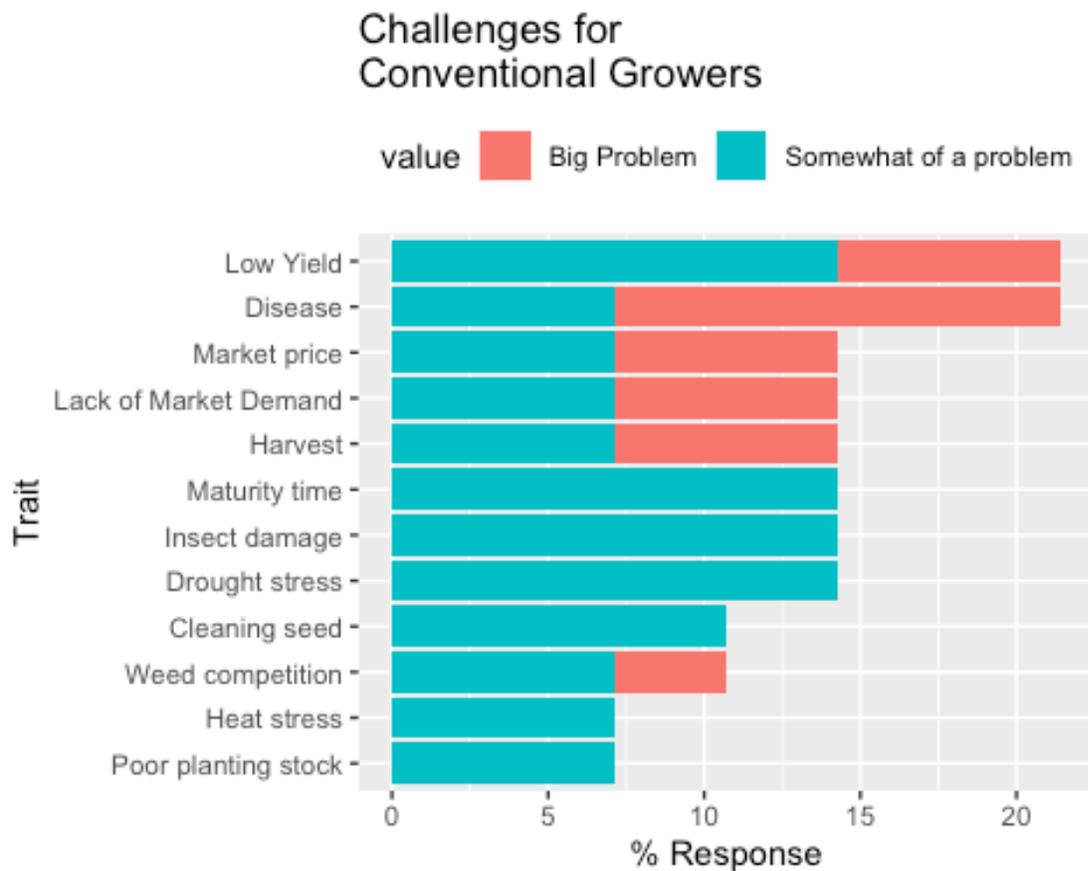
- insect/disease resistance and management
- Get clientele interested in the value of this incredible food source.
- More aggressive variety that will perform consistently under organic management
- More information and research for small scale mechanization
- We need higher yields to make organic beans profitable enough.
- harvest and processing capacity, and access to reasonably priced equipment is definitely the biggest barrier. disease testing for seed would also be helpful, and more readily available growing advice for og producers would be cool. but really my biggest limitation is access to functional scale appropriate equipment.
- Looking for small scale bean harvesting tools. Ability to mechanize harvest of 1000bf. Like a small potato digger, but for beans.
- Marketing
- Producing more BCMV varieties, providing varieties that can out compete weeds, using more drought resistant beans- combining the traits of teparies in other common beans

### Conventional Growers

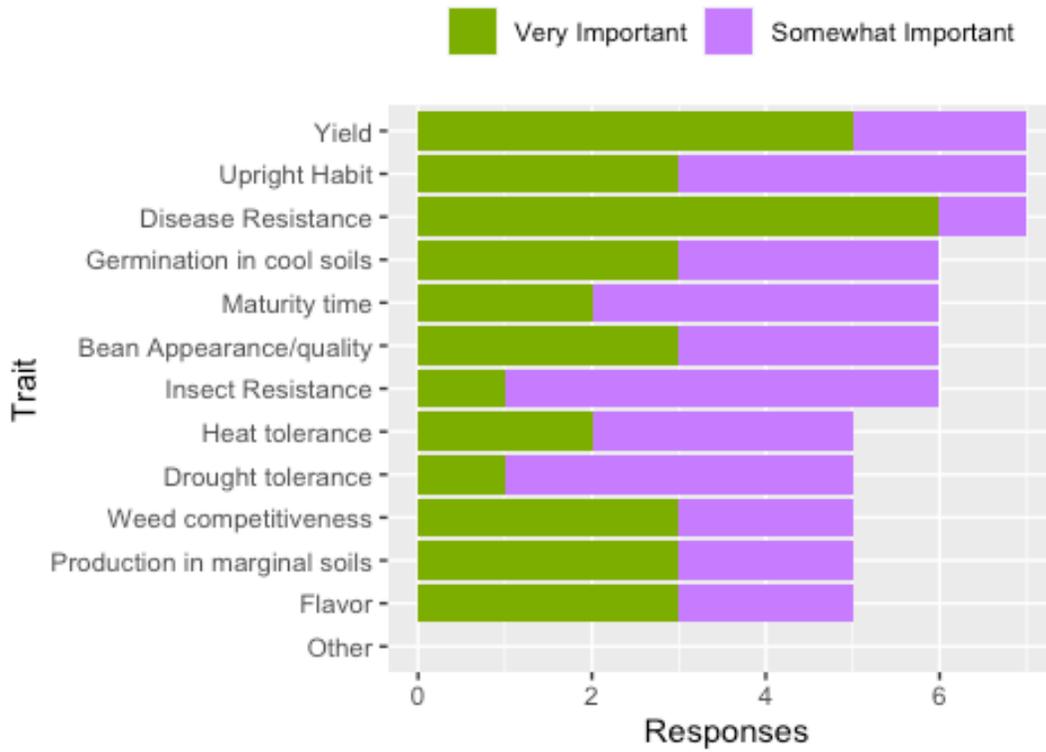
Biggest challenges cited by conventional growers were low yields and disease pressure, followed by market price, market demand and harvest challenges. Traits of top priority for conventional growers were disease resistance, yield and upright growth habit.

White mold and root rots were the most problematic diseases cited. Aphids and the two-spotted spider mite were the most problematic insect pests.

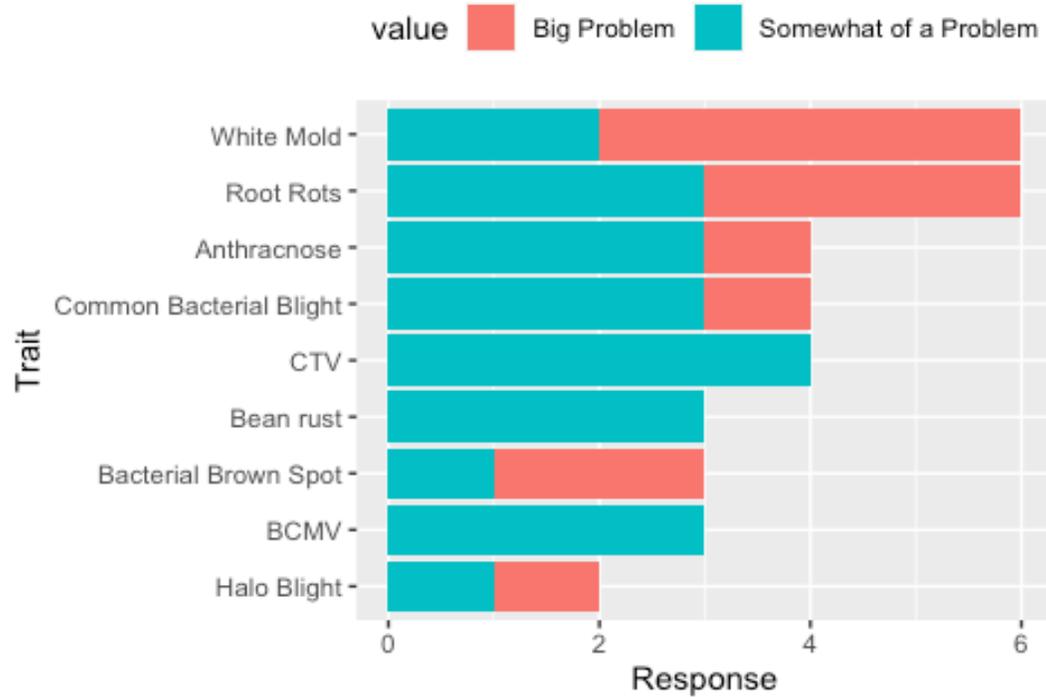
Research needs of top priority were agronomic research and variety development and improvement.



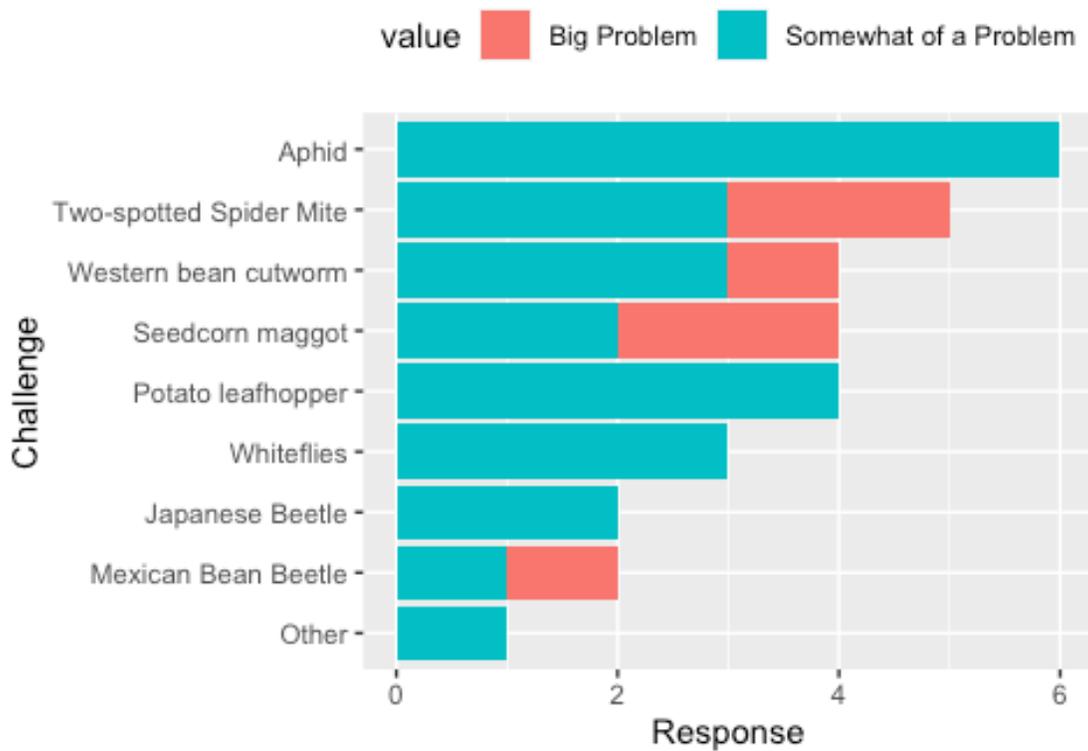
### Trait Priorities for Conventional Grower:



### Most Problematic Diseases for Conventional Growers



## Most problematic insect pests for Conventional Growers



### 'Other' Responses

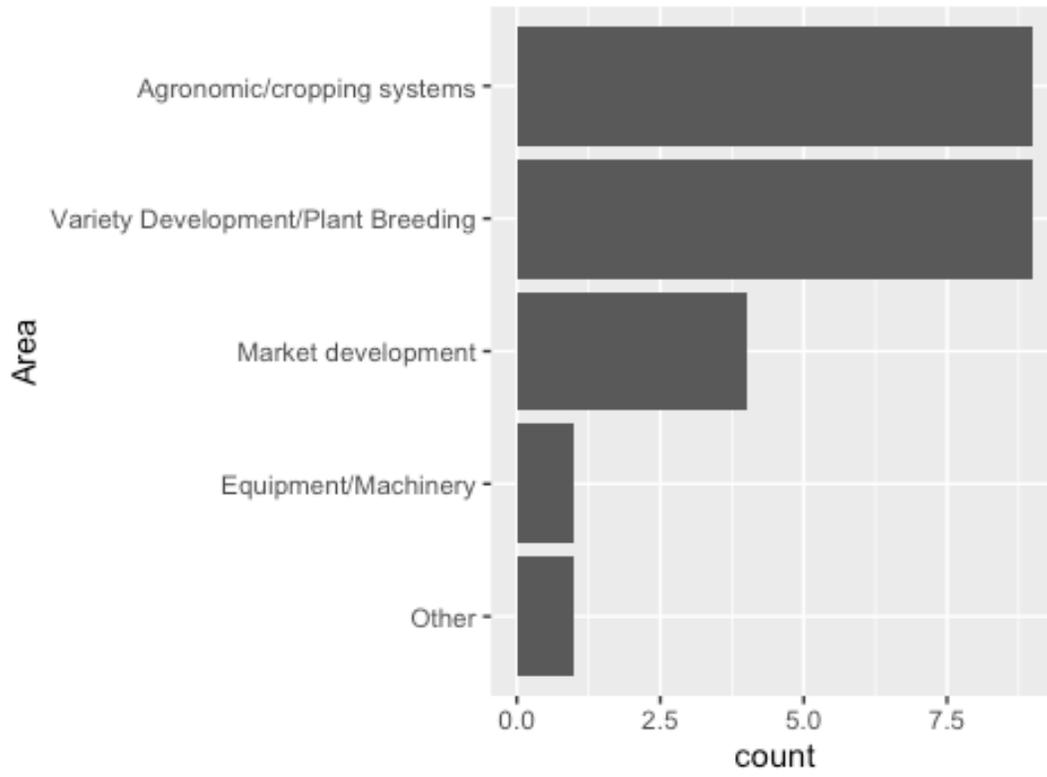
Big Problem - Cucumber beetle

Somewhat of a Problem - deer

Somewhat of a Problem - Flea beetle

Somewhat of a Problem - Brown snail

## Conventional Research Needs

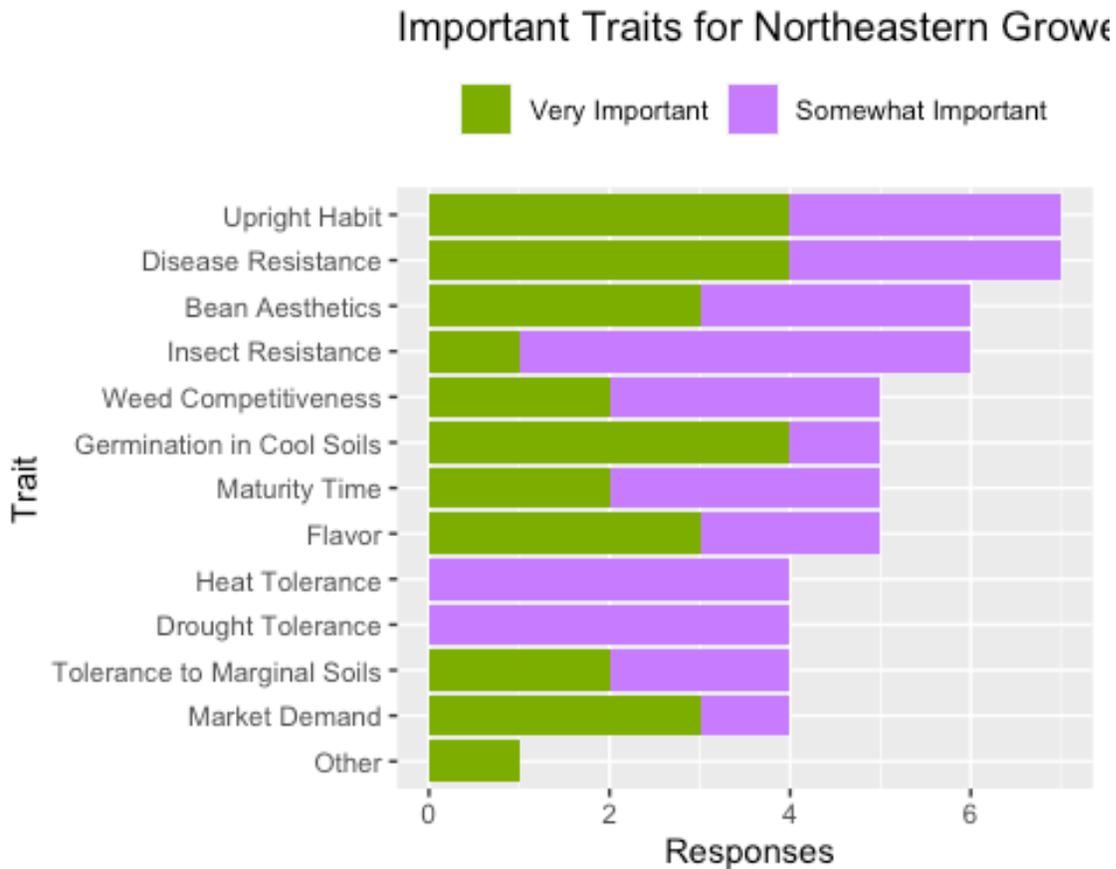


### Northeastern Growers

Most important traits for Northeastern growers were upright growth habit and disease resistance, followed by bean aesthetics and insect resistance. Biggest challenges to production were crop maturation timing and difficulty of harvest, presumably due to inclement fall weather. Price and yields were next reported challenges.

White mold was the most problematic disease reported, followed by root rots. Mexican Bean Beetle and Western bean cutworm, and potato leafhopper were most problematic pests.

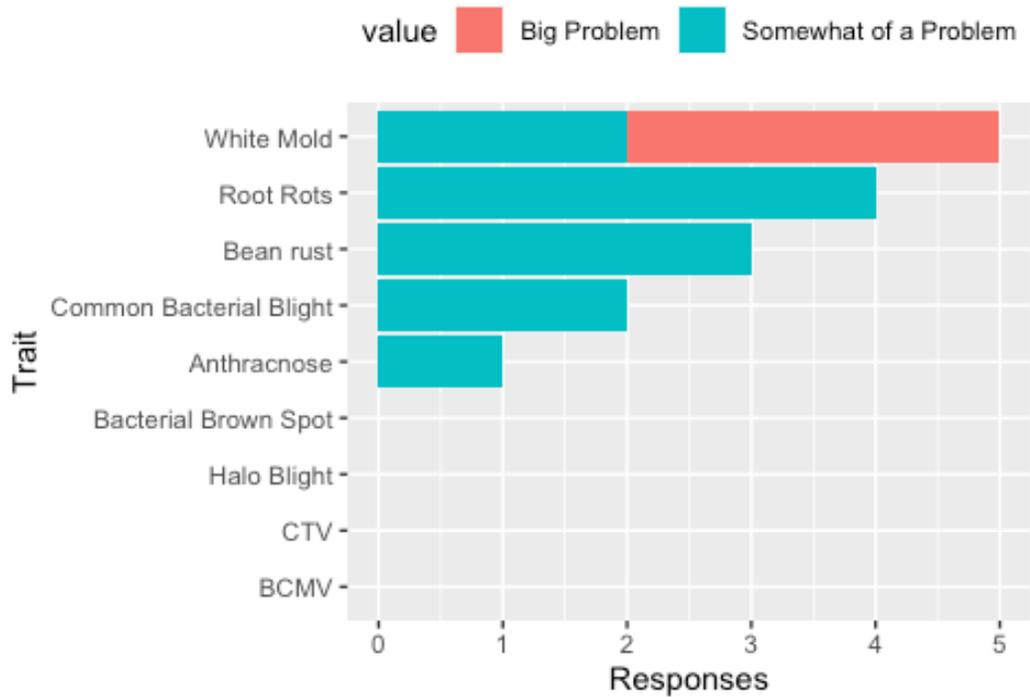
Variety development and market development were most reported research needs. Specifically, affordable and appropriate scale equipment was cited, as well as varieties that yield and perform well in organic systems.



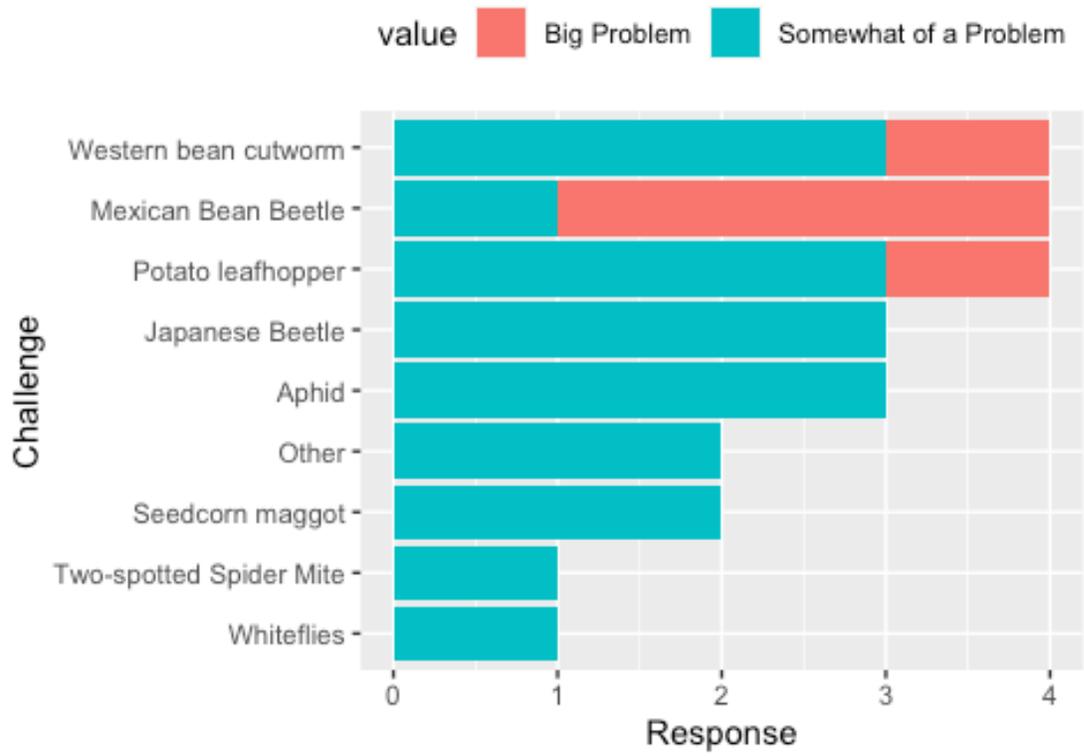
## Challenges for Northeastern Growers

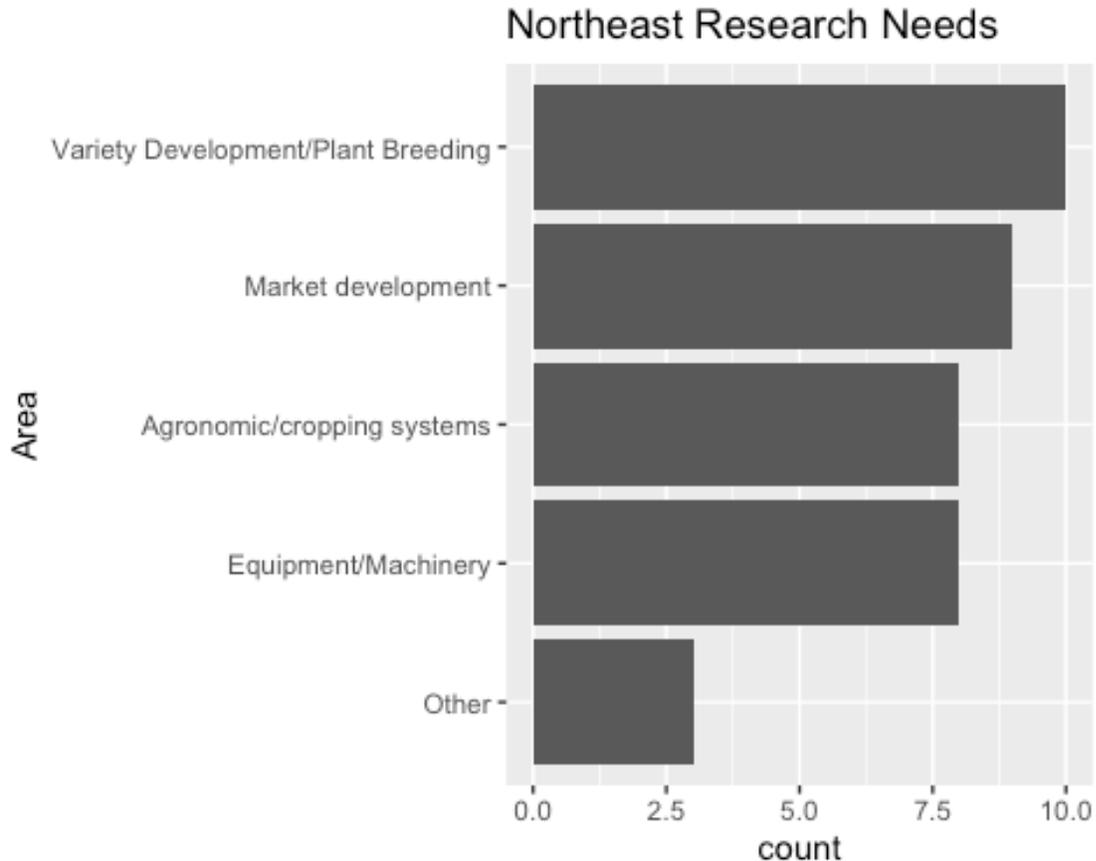


## Most problematic diseases: Northeast



## Most problematic pests: Northeast





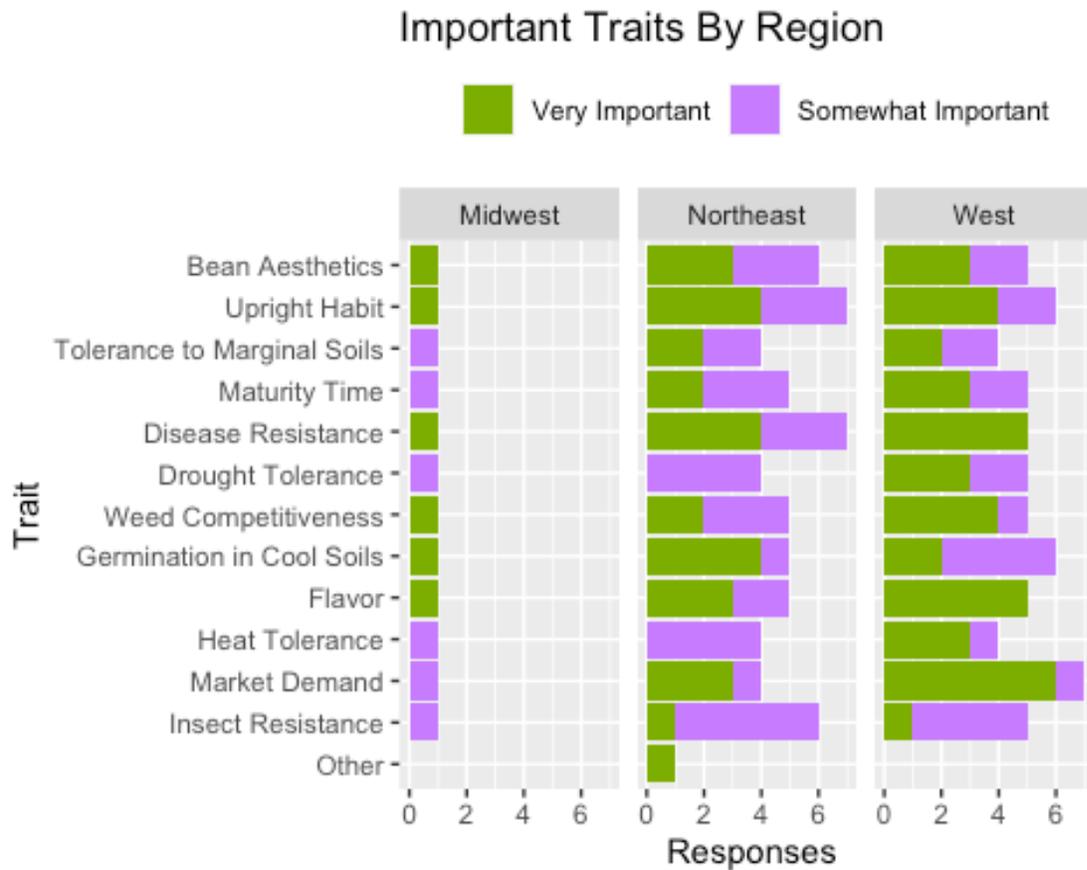
Elaborate: Organic Northeast Grower Needs

- Get clientele interested in the value of this incredible food source.
- More aggressive variety that will perform consistently under organic management
- More information and research for small scale mechanization
- We need higher yields to make organic beans profitable enough.
- harvest and processing capacity, and access to reasonably priced equipment is definitely the biggest barrier. disease testing for seed would also be helpful, and more readily available growing advice for og producers would be cool. but really my biggest limitation is access to functional scale appropriate equipment.
- Looking for small scale bean harvesting tools. Ability to mechanize harvest of 1000bf. Like a small potato digger, but for beans.

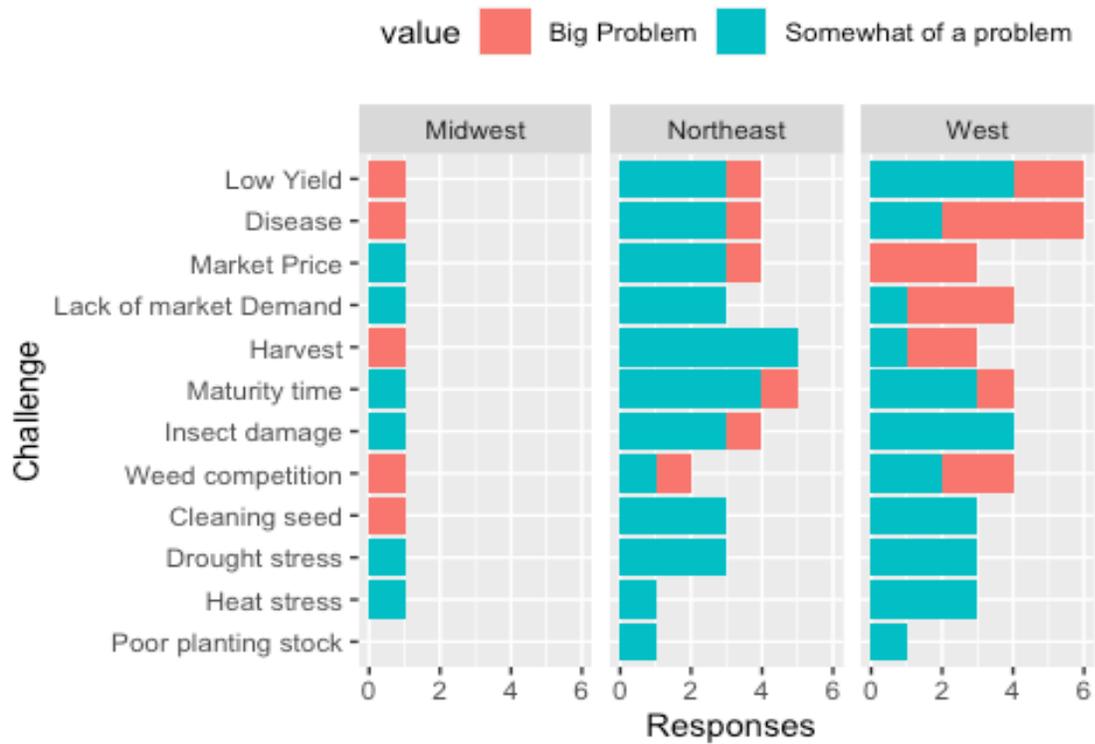
### Grower Responses by Region

When responses were compared across regions, some differences were apparent, though few responses from the Midwest limited comparison. First, Northeast growers were more likely to experience challenges with crop maturation and harvest, likely due to wet fall weather. Western growers cited low yields and disease comparatively more than other regions.

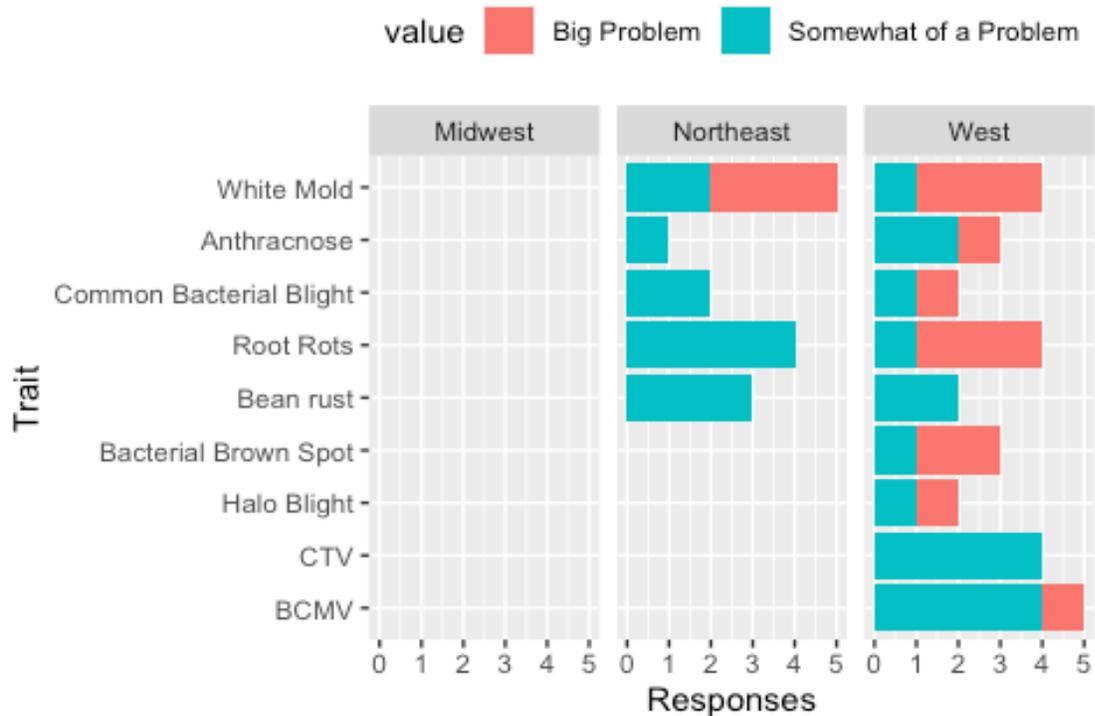
Northeastern growers reported more problems with White Mold, compared to greater issues with BCMV and CTV. Aphids, two-spotted spider mite and seedcorn maggot were bigger challenges in the West compared to comparatively bigger issues with Mexican Bean Beetle, Potato Leafhopper and Western bean cutworm in the Northeast.



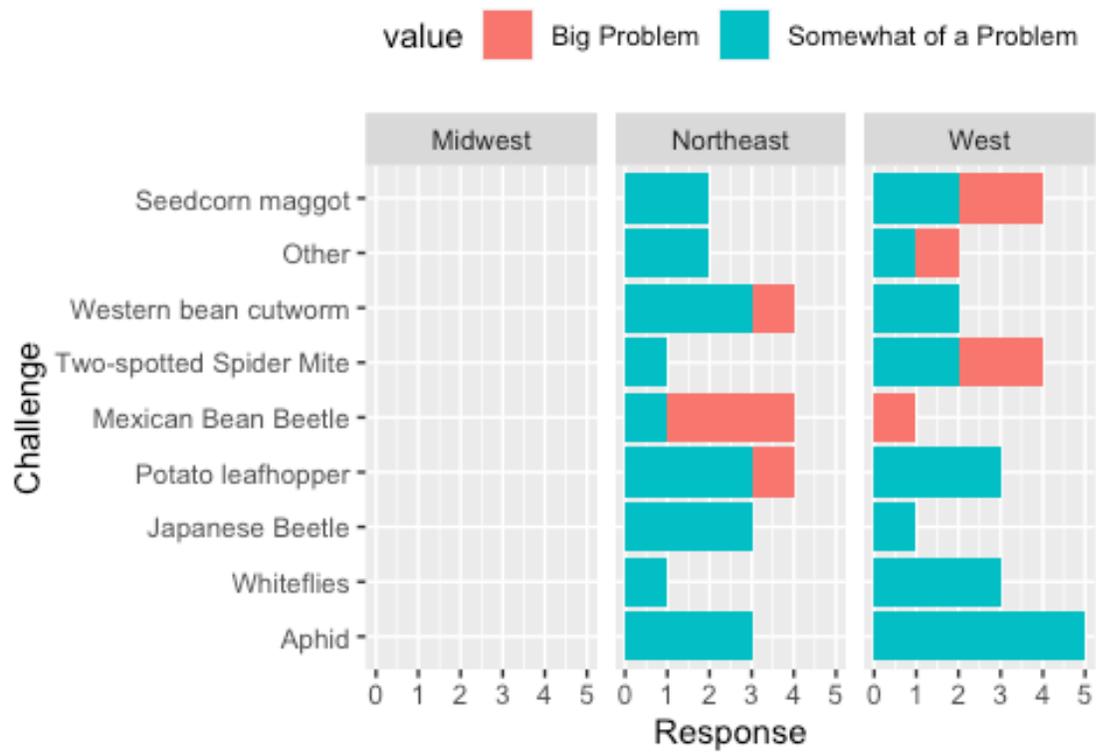
## Challenges by Region



## Most Problematic Diseases by Region



## Most Problematic Pests by Region



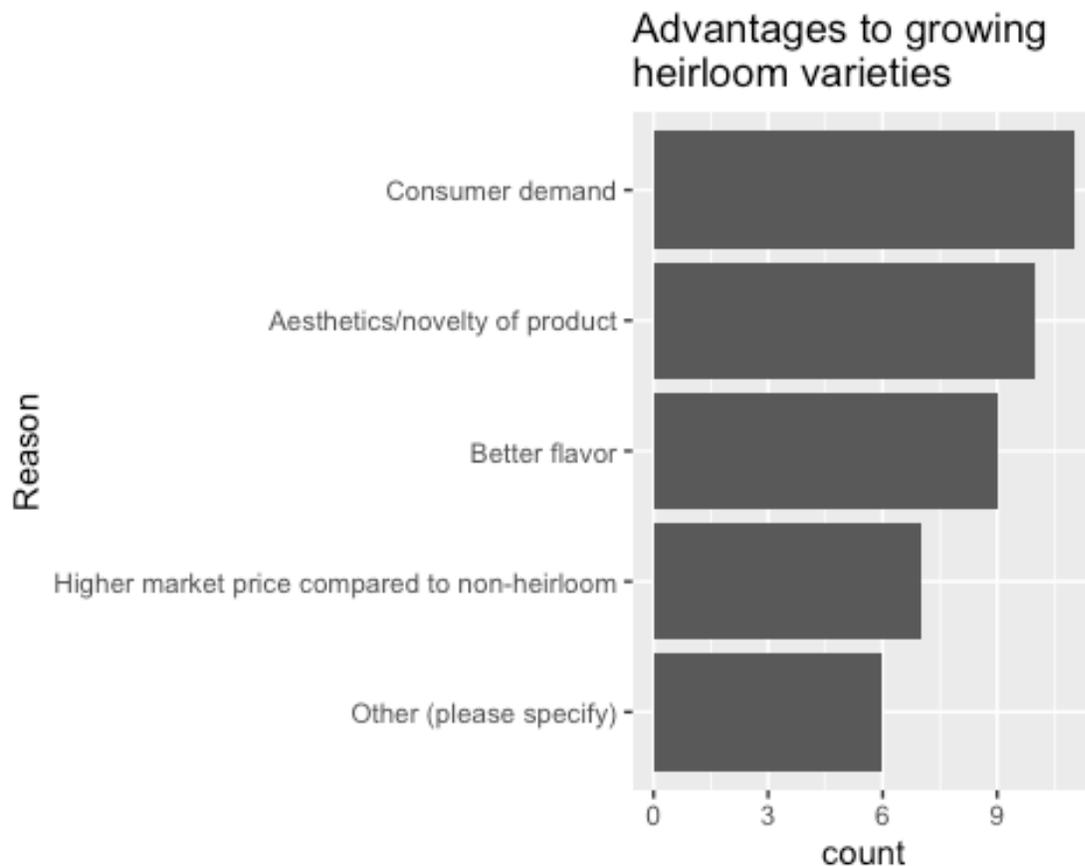
### Heirloom Bean Growers

Heirloom bean growers cited consumer demand and uniqueness of product as the top reasons to grow heirloom varieties. Aesthetics and flavor, followed by germination in cool soils, maturity time and disease resistance were cited as the most important traits for heirloom or heirloom-like bean varieties.

The biggest challenges to growing heirloom beans were difficulty of harvest, cleaning seed, appropriate maturation time, low yield and poor planting stock.

White mold, anthracnose and CTV were the most significant diseases for heirloom bean growers. Significant pesters were Japanese beetle, potato leafhopper and aphids.

Biggest research needs were variety development and improvement and appropriate equipment and machinery. A specific variety improvement suggestion was standability/upright growth and disease resistance.

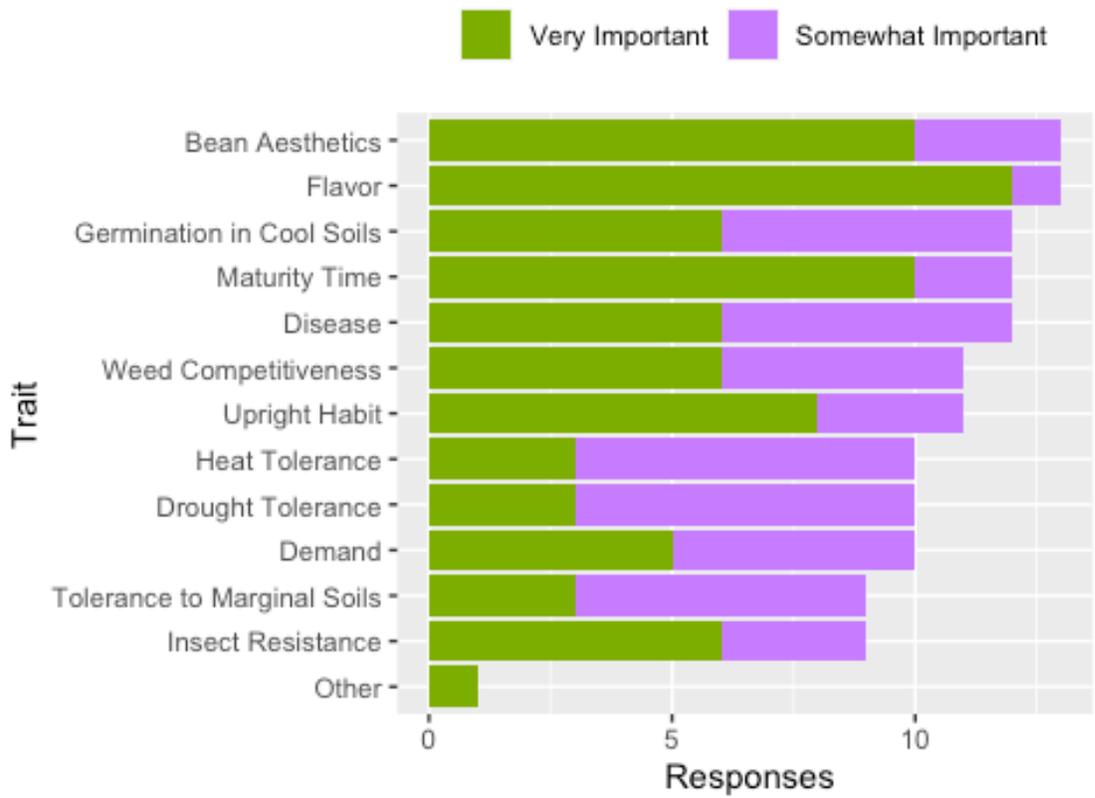


#### 'Other' Responses

- self-enjoyment
- Bring back old varieties. Diversity in food chain
- Brings a unique mix to our product offering.
- preserving part of our collective heritage
- Timing/early finish

complete diet for customers

### Heirloom Bean Important Traits



'Other' Responses

Very Important - not prone to shattering, amenable to hand harvest

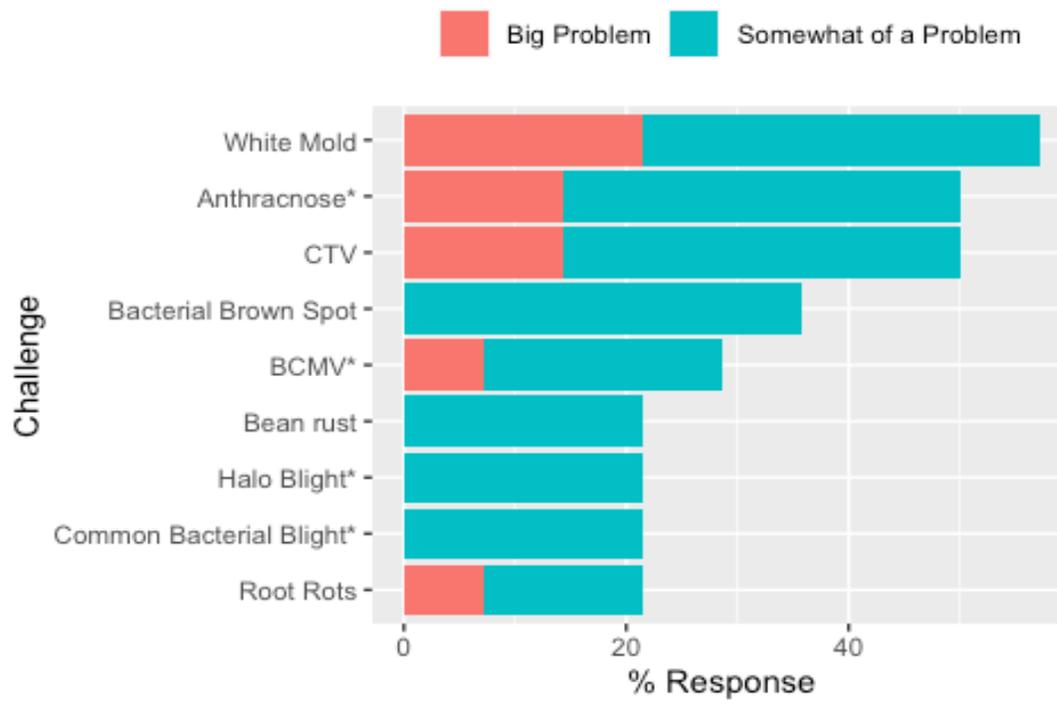
## Heirloom Bean Challenges



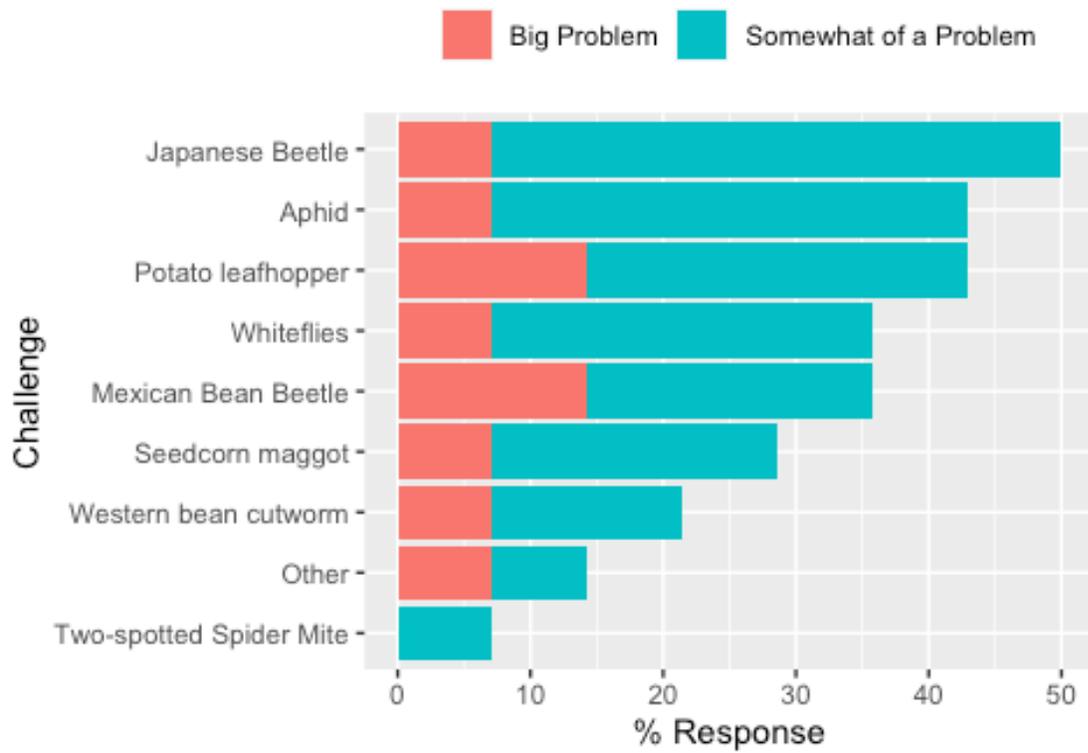
'Other' Responses

Big Problem - Waiting for soil to be warm enough to ensure germinations

## Most problematic diseases in heirloom varieties



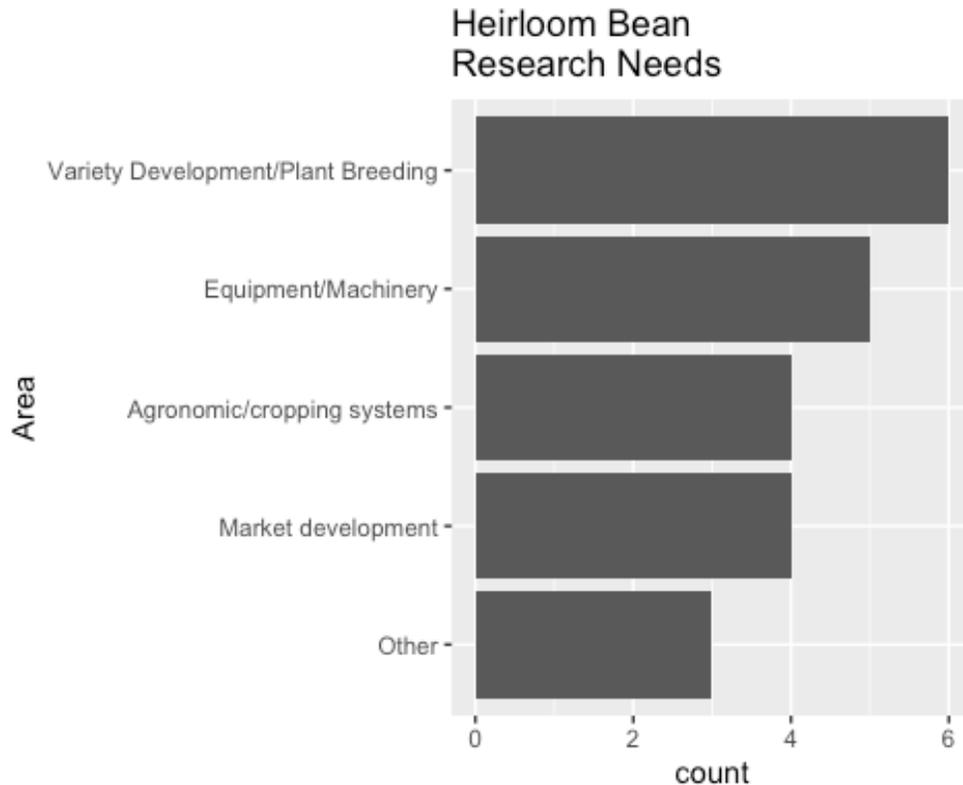
## Most problematic pests in heirloom beans



'Other' Responses

Big Problem - Deer

Somewhat of a Problem - tarnish plant bug



#### 'Other' Responses

Culinary reserach - to determine the best varieties for differnet recipes. Then getting the word out to develop markets

Understanding the history behind the beans and sharing profits with the communities that they belong to.

earlier maturing varieties

Elaborate on research needed for heirloom dry bean production

varieties that have been improved such as silver cloud cannellini have greatly improved yields over their heirloom counterparts

stand ability, pod integrity, disease resistance

Showing which beans hsvc evolved from heirloom beans and have this on record  
same stuff

More research on crossing tepary beans with common beans, more effort to increase plant architecture and BCMV resistance in heirloom beans, develop models that work with organic heirloom dry beans

develop sturdy, upright plants



Funded by Northeast SARE Graduate Grant GNE 19-208