LEVERAGING A GLOBAL SPRING, 2-ROW BARLEY POPULATION TO ACCELERATE THE DEVELOPMENT OF SUPERIOR FORAGE BARLEY VARIETIES FOR MONTANA GROWERS

by

Traci Janelle Hoogland

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Plant Sciences – Plant Genetics

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2021

DEDICATION

This work is dedicated to everyone who has mentored me along the way: Dr. Tony Jelsma, Dr. Jeff Ploegstra, Dr. Greg Vanden Heuvel, Dr. Madhulika Sharma, Dr. Sarah Pethybridge, Dr. Amara Dunn, Dr. Niloofar Vaghefi, and, of course, Drs Jamie Sherman, Jack Martin, Phil Bruckner, Mark Greenwood, and Jennifer Lachowiec.

ACKNOWLEDGEMENTS

A special thanks is rendered to the growers of Montana and the Montana Wheat and Barley Committee for providing funding which partially supported this work.

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

TABLE OF CONTENTS CONTINUED

LIST OF TABLES

Table Page

LIST OF TABLES CONTINUED

Table Page

LIST OF FIGURES

Figure Page

ABSTRACT

As more people around the globe escape poverty, they are eating more meat and dairy products. To support this increased demand for animal products there is an urgent need to develop more sustainable high-quality forage and hay crops for the livestock production industry. Barley (*Hordeum vulgare spp. vulgare L.*) is considered one of the most drought tolerant of the annual cereals and spring barley has been shown to out yield established perennial forages under drought conditions in central Montana (Cash, Surber, & Wichman, 2006). To accelerate the development of superior forage barley varieties for Montana, the following goals were identified 1) Utilize a genome wide association analysis to find genetic regions related to key forage and agronomic traits, 2) Use statistical modeling to a) examine the relationship between difficult to measure forage traits such as quality and yield, and easy to measure agronomic traits such as flowering time and plant height, b) identify agronomic traits that can be used as proxies for yield and quality in the earliest stages of the breeding program when genetic and phenotypic variability are at their greatest. Through these techniques the importance of variation in timing of plant maturity was identified. Statistical modeling showed that variability in forage yield and quality was observed to be closely related to variability in the timing of heading and soft-dough dates. Plant height was also determined to be of importance especially for biomass yield. Through genome-wide association analysis, novel QTL were discovered in relation to all studied traits. QTL were detected on all seven chromosomes and the majority of forage trait QTL co-located with QTL related to the timing and progression of plant development and maturity. This appeared to indicate that in a population of global barley accessions, the loci with the greatest impact on forage traits may be those containing genes regulating plant development and senescence. This further strengthened the evidence from the modeling study that a relationship exists between the two trait categories: traits for measuring the timing of plant development and forage traits.

CHAPTER ONE: PROJECT JUSTIFICATION AND GOALS

Importance of Breeding for Forage Barley in Montana

Value of forage breeding research

As more people around the globe escape poverty, they are eating more meat and dairy products. Economic development and population growth result in an increased demand on the animal food industry. Annual meat production is projected to increase from 218 million metric tonnes in 1997-1999 to 376 million tonnes by 2030 (Steinfeld & Otte, 2003). To support this increased demand for animal products there is an urgent need to develop more sustainable high-quality forage and hay crops for the livestock production industry. Diminishing resources, including land and water, require forage to be produced with improved efficiency, i.e. higher forage yield and quality with fewer inputs. These crop efficiency goals are even more challenging in light of the projected scenarios of global warming. It is predicted that by the end of the century, the area-weighted average yield of crops will decrease by 30-46% under the slowest warming scenario and decrease by 63-82% under the most rapid warming scenario (Schlenker & Roberts, 2009).

Plant breeders play a critical role in the ability of forage growers to meet the increasing demands of the livestock production industry. Making crosses, screening lines, and selecting individuals with improved traits has been critical to increasing food production over time. Given climate projections, the importance of plant breeding only increases as forage resiliency must increase. While plant breeders must consider the global climatic and product demand trends, public breeders at land-grant universities such

as Montana State, must often initially breed locally – first and foremost meeting the needs of their in-state constituents.

In order to effectively breed a resilient forage crop for a region, it is important to first understand the unique characteristics and constraints of the region. Primarily due to low rainfall, the Great Plains of Montana were historically a grassland. Reflecting this, 65% of Montana's agricultural lands are still pasture and rangeland. Of the four billion dollars comprising Montana's average annual agricultural revenues, about 2 billion is generated by the cattle industry. (USDA National Agricultural Statistics Service, 2020) During summer months, cattle are fed on range or pastureland, but supplemental forages are required to maintain herds during the winter. Most ranchers produce their own winter hay or grain, but whether the feed is produced by the ranch or purchased the cost is extremely burdensome. Forage experts such as Dennis Cash, a forage extension agent in Montana for 14 years, have stated the following:

> "Depending on location, producers need a 2 to 4 month supply of hay to get through the winters in the northern Great Plains. Aside from long periods of snow cover, high-quality forages are required to offset poor-quality roughages available on range. Winter feed is the largest cost on ranching operations, and slight improvements in forage production can significantly reduce costs." (Cash, 2006)

Indeed, in a 2017 census of Montana farmers the cost of livestock feed was the single largest production expense representing nearly 12% of the total annual cost of operation (USDA National Agricultural Statistics Service, 2020).

In Montana, alfalfa is currently the primary crop for hay with an annual production of about 3,000,000 tons per year; while all other forages combined make up about 2,000,000 tons per year (USDA National Agricultural Statistics Service, 2020).

However, alfalfa requires irrigation in Montana, a state where eighty percent of farming acres are rainfed. Other forages, such as annual cereals with arid tolerance, early biomass and high nutritional value; could provide a more sustainable hay option for growers. Given the predicted trends of rising temperatures and diminished water resources, the ability to rapidly produce biomass under low water input conditions will only grow more important. Breeding for more diverse forage crops would also provide a wider phenological range, giving growers technical advantages including flexible forage harvest time to synchronize with transportation, processing and storage limitations (Kennelly & Weinberg, 2015; Tsuberi *et al.*, 2016).

Barley (*Hordeum vulgare spp. vulgare L.*) is considered one of the most drought tolerant of the annual cereals and spring barley has been shown to out yield established perennial forages under drought conditions in central Montana (Cash *et al.*, 2006). Additionally, it is exceedingly salt tolerant (Maas, 1993; Steppuhn & Raney, 2005), another important consideration in a state where more than 300,000 acres are severely affected by high soil saline levels (Montana Salinity Control Association, 2017). Barley is well adapted to the cool, short growing seasons found in the state, but just as importantly, it is considered to be adaptable to a greater range of climate than any other cereal (Britannica, 2020) and is a promising model organism for how grasses adapt to climate change (Dawson *et al*., 2015). Taken together, barley's characteristics indicate that it would be an ideal target for forage breeders trying to fill the increased need of Montana growers for an adaptable and sustainable roughage crop.

Although forage yield is very important, forage quality is also a significant factor when breeding for a forage cultivar. Forage quality is a broad term that encapsulates how well a forage meets the nutritional requirements of the animal consuming it. When considering barley forage quality to meet the nutritional needs of cattle in Montana, two related traits should be of particular interest to breeders: intake and digestibility.

Intake is a measure of how much of a given forage an animal can consume before experiencing gut-fill and appetite suppression. This factor is primarily related to the concentration of plant biochemicals which are difficult (hemicellulose, cellulose) or essentially impossible (lignin) for even a ruminant animal to digest. If a forage is of poor quality and thus low intake, the animal may not be able to ingest a sufficient quantity of feed to meet its nutritional requirements.

Digestibility is a measure of how easily a forage is digested by an animal. Forage digestibility is primarily related to the percent of indigestible material, mostly lignin, present in a forage; it determines the nutritional value of the crop and can affect feeding efficiency. Breeding for increased forage digestibility is, in a manner, a way to breed for forage nutritional yield. As an example: if two forage lines, A and B, which each yield 3 tons of dry biomass per acre; where line A is 60% digestible and line B is 65% digestible – then line B out yields line A by 0.15 tons per acre in digestible yield (1.95 vs 1.8 ton/acre respectively). A rancher growing line B would meet the energy requirements of their herd using 8% less land than if they had grown line A. That would not only mean savings on inputs for growing the forage, but savings on fuel, time, equipment, and labor for both harvesting those acres and transporting the additional hay to the animals.

In addition to being more efficient on the production end, feeding a higher quality forage can also have a positive impact on animal gains. A conservative estimate is that a one percent increase in forage digestibility can lead to a three percent increase in the average daily gains of steers, a three percent increase in milk production for dairy cows, and an overall increase in animal production per unit area (Casler & Vogel, 1999; Mohammed, 1967).

Montana's hay production from 2018 can be used to present a crude illustration of the potential impact that increasing forage digestibility could have on Montana growers. In 2018, there were 2.9 million acres of hay production in Montana with an average yield of about 1.93 tons per acre for a total production of about 5.6 million tons of hay. The price of hay was, on average, $145 per ton and Montana's hay crop was valued at about $800 million. (USDA National Agricultural Statistics Service, 2020) Since, in general, a 1% increase in digestibility equates to between a $1.64 (Ward, 2000) and a $2.19 (Hopper, Peterson, & Burton, 2004) increase in price/ton; increasing the digestibility of the 2018 hay crop by 1% would result in a $9-12 million increase in the total value of the crop. This increased value represents the added monetary value of the hay crop, although it does not capture money saved by increasing the animal production per unit area, the increased feeding value of hay that is fed on-site, or the increased value of the calves weaned from cattle fed a higher quality hay.

If Montana growers would benefit from barley forage cultivars, and forage yield and quality are the two main focuses for a forage breeder, then next step would be to determine how best to accelerate breeding and releasing superior barley forage cultivars

for Montana. In 2015, the following dissertation work was begun to address that question. To accelerate the development of superior forage barley lines for Montana, four main goals were identified:

1) Since stage of maturity is known to impact yield and quality (Casler & Vogel, 1999), implement daily monitoring of forage trials to ensure experimental plots are harvested at the same stage of maturity, eliminating that key source of variation in forage quality and thus making forage lines comparable across plots, environments, and years.

2) Develop near-infrared reflectance (NIR) technology to make screening for forage quality faster and cheaper – allowing more samples, and thus more barley lines at earlier stages of the breeding program and across more environments, to be tested.

3) Identify exotic germplasm with superior digestibility and biomass yield for incorporation into the MSU barley breeding program

4) Utilize a genome wide association analysis to find genetic regions related to key forage and agronomic traits.

5) Use statistical modeling to a) examine the relationship between difficult to measure forage traits such as quality and yield, and easy to measure agronomic traits such as flowering time and plant height, b) identify agronomic traits that can be used as proxies for yield and quality in the earliest stages of the breeding program when genetic and phenotypic variability are at their greatest.

Breeding and selecting for forage traits

In order to understand each of these goals, it is helpful to understand the overall breeding process. Traditional breeding for barley typically requires 10-15 years from initial cross to the release of a superior variety. Barley cultivars are most commonly released in the United States as homozygous plants in homozygous populations. While breeding programs can use several different techniques to advance lines through a selection process, the generalized breeding scheme traditionally used in the spring barley breeding program at Montana State will be discussed here. Under this schedule, an initial cross is made between two parental lines and the seed from the cross is designated as the familial generation 1 or F1. In the F1 generation, the individuals are 100% heterozygous for each gene for which the parents carried different alleles. The plants grown from the F1 seed self-fertilize to produce the F2 generation (50% heterozygous) and this process is repeated to produce the F3 (25% heterozygous) and F4 (12.5% heterozygous) barley heads. The F1 through F3 generations are increased in the greenhouse and selection pressure is avoided. The F4 heads are harvested and each head is threshed separately for seed. The seed from each F4 head is then space-planted in the field as F4 “head-rows” to allow single plants to be selected from each cross in the field based on visual traits. Typically, about 100 heads per cross are planted each with 20 seeds per head. Thus, a total of 2000 F4 plants may be evaluated per cross. From these 2000 plants only about 10-20 F4 plants may be selected (0.5-1.0% of the initial 2000 plant population), or the cross may be dropped altogether. The F5 seed (6.25% heterozygosity) from each selected F4 plant is bulk threshed and planted in single-row field plots the following season. From

the 10-20 F5 “plant-rows” selected from a cross, 2-4 rows (0.1-0.2% of the initial 2000 plant population) per cross may be selected for harvest of F6 seed (3.125% heterozygosity) and advancement to the forage preliminary yield trial (PYT) the following season. At the F4 and F5 stages the plants are visually evaluated for traits such as awnlessness and general plant vigor, but no quantitative data had been collected on plants, head-rows, or crosses overall. The destructive nature of forage sampling makes testing at the F4 stage impossible, since only single plants are available. Forage sampling at the F5 stage is possible but, given the number of samples to be tested, quality data can only be produced if a fast and cheap assessment tool is available.

At key points in the earliest stages of the breeding process, heavy selection pressure is applied, and the genetic variability of the population is severely reduced. The heaviest selection pressure in the breeding pipeline has been applied when data on important forage traits such as quality and yield is the least robust. Prior to 2016, due to lack of cost-effective quality testing, forage quality data was not even a factor in selection of forage barley. Goals 2 and 5 of this work were thus identified.

Goal 2, developing NIR technology, would make quality testing possible at the F5 stage and beyond. With NIR, hundreds of forage quality samples could be analyzed each year, allowing selection for forage quality within the breeding pipeline for the first time. Determining digestibility can be expensive and labor intensive. Feeding studies are the gold standard for nutritional quality of a forage, but their expense limits their utility for breeding selections. Therefore, proxies have been developed. Important proxies are Neutral Detergent Fiber (NDF) and Acid Detergent Fiber (ADF) methods via the

ANKOM 2000® analysis system. However, these proxies too are expensive and time consuming. A more efficient proxy is near-infrared reflectance (NIR) where forage samples are dried and milled to a 1mm particle size using a Wiley® Cutting mill and then placed in a chamber, excited with infrared light and reflectance is measured. For NIR to be predictive of digestibility, a calibration curve must be constructed, requiring NIR, NDF and ADF to be measured on a representative set of materials. Relative to other methods, NIR is the only cost-effective way to measure forage quality on hundreds of forage samples in a timely manner.

Goal 5 would allow the identification of agronomic traits related to forage quality and yield. These traits could be assessed even at the F4 stage, when destructive forage sampling is not feasible, and could be used to increase the accuracy of selection at the earliest stage of the field breeding program. Currently, there is no means of selecting for forage quality at the F4 generation. Since this is a stage of high selection pressure, a morphological proxy that could be measured on single plants is required.

It is understood that the larger the genetic variability in a population under selection pressure, the greater the potential genetic progress (Moreno-González & Cubero, 1993). Being able to accurately select for forage traits early in the breeding pipeline is thus especially important because the early stages have greater genetic diversity.

Utilizing exotic germplasm to improve MSU germplasm

Genetic variability is crucial to a breeding program for several reasons. As previously stated, greater genetic variability allows for greater potential genetic progress (Moreno-González & Cubero, 1993). The introduction of genetic diversity into the breeding program allows for the continued, long-term improvement of forage traits. But, just as important to the sustainability of the breeding program, introducing genetic diversity is a key component to ensuring that the breeding program can adapt to changing environmental factors.

Barley was one of the first plant species to be domesticated (Von Bothmer & Komatsuda, 2011), with domestication centers in the Near East Fertile Crescent and the Tibetan Platea (Åberg, 1938; Bekele, 1983; Fuller, Willcox, & Allaby, 2011; Molina-Cano *et al.*, 1987; Molina-Cano *et al.*, 2002; von Bothmer *et al.*, 1995; Von Bothmer & Komatsuda, 2011; Wang *et al.*, 2015). Over time barley moved with humans until it had been cultivated on and adapted to every continent where crops are grown (Britannica, 2020). This has led to a wealth of genetic diversity in the species, a diversity that has been collected in global germplasm repositories and made available to researchers and plant breeders.

One such resource is the United States Department of Agriculture – Agricultural Research Service's (USDA-ARS) National Plant Germplasm System, which includes a National Small Grains Collection (NSGC) containing more than 33,000 *Hordeum vulgare L.* accessions. These lines have been collected from more than 100 countries and are categorized based on their level of genetic improvement, from unimproved landraces

to genetic resources and breeding program material, all the way to elite cultivars. As part of the USDA-NIFA funded Triticeae Coordinated Agricultural Project (TCAP), Muñoz-Amatriaín *et al.* selected 1,860 unique barley accessions from the NSGC to represent the total diversity of the collection and genotyped them with a barley SNP iSelect platform with 7,842 markers (Munoz-Amatriain *et al.*, 2014). Of the 1,860 lines in this "Core" collection, 621 lines were of the spring, 2-row morphology, which is the form of barley most commonly grown in Montana.

Spring, 2-row lines from the barley Core collection provide a valuable resource of exotic genetics to incorporate into existing MSU germplasm. Screening lines from around the globe for forage and agronomic traits would provide triple benefit. First, fulfilling goal 3, lines with superior traits could be identified and crossed into Montana adapted forage barley lines. These lines would contribute beneficial alleles for forage quality and yield, but as diverse lines collected from around the world, they would also be a source of critical genetic diversity. Second, through genome wide association analysis, the population could be used to investigate important questions regarding which genetic regions control important forage traits and how agronomic traits might relate to forage traits; goal 4. Finally, screening this population provides an ideal opportunity for collecting the very data needed to perform the statistical modelling necessary to achieve goal 5.

As previously discussed, forage digestibility is impacted by the timing of harvest (crop physiological developmental stage). As a rule, as crop physiological development advances, indigestible cellular components such as hemicellulose, cellulose, and lignin

increase, and forage quality decreases. Conversely plant biomass accumulates as the crop ages and biomass yield benefits from delays in harvest. These relationships, between timing of maturity and forage yield and quality would be of particular interest as a focus for modeling.

Summary

Forage and feed are where crop and livestock production meet – where the two largest pieces of the Montana's economy intersect. And because a relationship exists, as we would expect it to, between the digestibility of a forage and the performance of livestock – forage quality is of economic significance for both forage and livestock growers. With a relationship that is, in general, 1% increase in digestibility = ~3% increase in animal daily gains, even small gains in forage digestibility can have big impacts for growers.

In addition to breeding for forage traits, long term any plant breeding program must also consider the genetic diversity of the germplasm and the necessity to breed for adaptability to a changing climate. Breeding for increased digestibility can have a positive financial impact on growers and could enhance agricultural sustainability. Fewer acres of high-quality forage are needed to support a given number of animals; thus, a high-quality forage cultivar requires fewer acres and less water for equivalent feed production.

In order to accelerate the creation and release of improved forage barley lines, genetic diversity in the breeding program should be increased and selection for forage

traits should occur as early as possible in the breeding pipeline – when the genetic diversity is the greatest.

Taken together the five goals of this work should address each of these important considerations, leading to superior forage barley being released faster and the long-term health of the breeding program being ensured.

CHAPTER TWO: LITERATURE REVIEW

Importance of Forage Breeding for Montana

Of the four billion dollars comprising Montana's average annual agricultural revenues, about 2 billion is generated by the cattle industry (USDA National Agricultural Statistics Service, 2020). During summer months, cattle are fed on range or pastureland, but supplemental forages are required to maintain herds during the winter. Most ranchers produce their own winter hay or grain, but whether the feed is produced by the ranch or purchased the cost is extremely burdensome. Forage experts such as Dennis Cash, a forage extension agent in Montana for 14 years, have stated the following:

> Depending on location, producers need a 2 to 4 month supply of hay to get through the winters in the northern Great Plains. Aside from long periods of snow cover, high-quality forages are required to offset poor-quality roughages available on range. Winter feed is the largest cost on ranching operations, and slight improvements in forage production can significantly reduce costs. (Cash, 2006)

Indeed, in a 2017 census of Montana farmers the cost of livestock feed was the single largest production expense representing nearly 12% of the total annual cost of operation (USDA National Agricultural Statistics Service, 2020).

Plant breeders play a critical role in the ability of forage growers to meet the increasing demands of the livestock production industry. Making crosses, screening lines, and selecting individuals with improved traits has been critical to increasing food production over time. Given climate projections, the importance of plant breeding only increases as forage resiliency must increase.

For Montana, a forage species of particular interest is barley, *Hordeum vulgare subsp. vulgare L*. Barley is considered one of the most drought tolerant of the annual cereals and spring barley has been shown to out yield established perennial forages under drought conditions in central Montana (Cash *et al*., 2006). Additionally, it is exceedingly salt tolerant (Maas, 1993; Steppuhn & Raney, 2005), another important consideration in a state where more than 300,000 acres are severely affected by high soil saline levels (Montana Salinity Control Association, 2017). Barley is well adapted to the cool, short growing seasons found in the state, but just as importantly, it is considered to be adaptable to a greater range of climate than any other cereal (Britannica, 2020) and is a promising model organism for how grasses adapt to climate change (Dawson *et al.*, 2015). Taken together, barley's characteristics indicate that it would be an ideal target for forage breeders trying to fill the increased need of Montana growers for an adaptable and sustainable roughage crop.

In order to breed for improved forage varieties, researchers must learn as much as possible about the plant physical, biochemical, and developmental traits underlying key forage traits such as forage yield and quality. Screening genetically diverse materials to identify the genetic regions underlying these traits is an important step. An additional important area of research would be identifying traits that can be quantified early in the plant breeding selection process that are directly related to yield and quality. Such traits could function as proxies for forage yield and quality, allowing for selection for yield and quality in early generations of the breeding pipeline, when genetic variation is greatest.

Forage Quality

Vegetative plant biomass, or forage, provides the foundation of the diets of ruminant livestock such as cattle and sheep. As a major component of these animals' diets, forage quality is closely monitored by both forage growers and livestock producers. Forage quality is a term used to encompass a range of forage characteristics. When considering a forage, what the term "quality" implies is fitness for a particular end-use. Thus, forage quality is about how well a given forage meets the needs of the animal and, ultimately, how well the animal performs on that forage. While forage quality encompasses all the traits considered to be desirable in a forage, it is possible to summarize the major quality characteristics as falling into two, related categories: "ingestibility" and digestibility.

Forage consumption is dependent on an animal's willingness and ability to ingest a quantity of forage sufficient to meet their needs, or the forages' ingestibility. While digestibility can be defined as – from the ingested quantity of forage, what proportion of the plant material can the animal break down and utilize. Ingestibility and digestibility are two areas of forage quality where even small improvements can mean large gains for animals, and so are often the focus of forage improvement.

Ingestibility and digestibility are related to the biochemical make-up of the plant tissue. The main plant biochemical components responsible for low animal intake and poor digestibility are plant structural components such as hemicellulose, cellulose, and lignin. These complex structural components make up the majority of the plant cell wall (Figure 2-1).

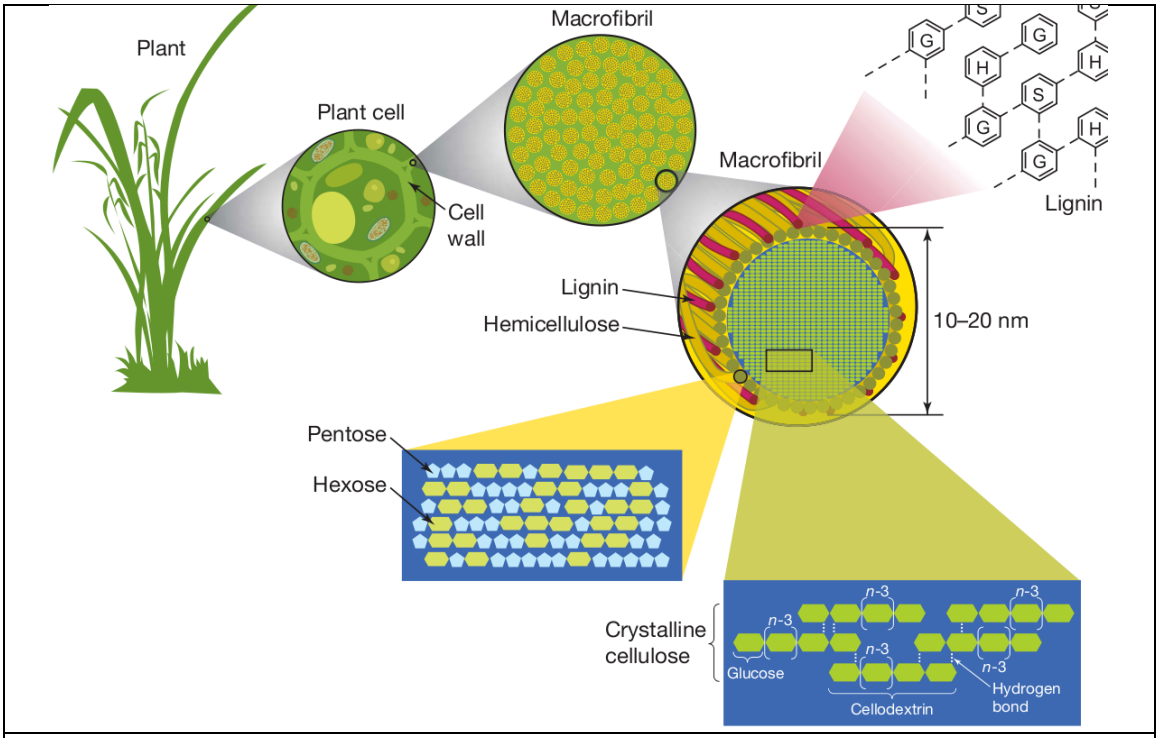


Figure 2-1 Depiction of plant cell walls and primary structural components. Modified with permission from (Rubin, 2008).

Impact of plant maturity and architecture on forage quality

The deposition, composition, and proportion of biochemical components such as hemicellulose, cellulose and lignin within a plant are influenced by many factors. Plant architecture characteristics such as leaf-to-stem ratio impacts quality since stem tissue, having a greater proportion of structural components, will have poorer quality. Plants also deposit more cell wall components over time and thus decrease in forage quality as they mature (Cherney & Marten, 1982a, 1982b). However, in cereals the decrease in digestibility generally levels off during grain fill as the highly digestible grain head becomes a greater proportion of the total biomass yield (Cherney & Marten, 1982a, 1982b; Garnsworthy & Stokes, 1993). When grain proportion becomes substantial, i.e. at

dough stage (Zadok stage 80), digestibility could even increase (Ben-Ghedalia, Kabala, & Miron, 1995; Crovetto *et al.*, 1998; Helsel & Thomas, 1987).

Biotic stresses such as insect pressure and abiotic stresses like drought can increase lignification, cause leaf senescence, and in general decrease green plant tissue; all of which decrease forage quality. Each of these factors - plant architecture, plant maturity and leaf senescence, plant responses to biotic and abiotic factors - have genetic components that interact with environmental factors to determine the final plant phenotype of interest, forage quality.

The Biochemistry, Candidate Causative Genes, and Plant Physiological Pathways Underlying Fiber Digestibility

Forage digestibility is a complex quantitative trait. It is the phenotypic result of the interaction of many different genes, most of which have yet to be identified. Although a few mutations (mostly enzymes in the lignin pathways) with significant implications for digestibility have been discovered or generated in species such as alfalfa and maize (Brenner *et al.*, 2010), very little work has been done to study the genetics of digestibility in barley. Although it is possible to breed for forage digestibility without knowing anything about the underlying genetics – simply by making and screening many, many barley crosses – such an approach is expensive, time consuming and requires as many as 15-20 years to develop a single adapted variety.

Correlating fiber digestibility with a molecular mechanism, single gene, or even a few genes is difficult. Three major biochemical compounds are closely linked to the digestible fiber available in a plant: cellulose, hemicellulose, and lignin. Any gene related

to the deposition of lignin, cellulose and hemicellulose may be tied to the forage quality of a species. Based on the genome of the model organism *Arabidopsis thaliana*, even just considering the lignin biosynthesis pathway alone, there are about 70-80 genes encoding enzymes directly involved in lignin formation (Lubberstedt, 2007).

Mutations in Caffeic Acid O-Methyltransferase (COMT) in *Medicago sativa* and implications for forage quality

Although all three cell wall components; hemicellulose, cellulose, and lignin, are considered to be of poor digestibility, lignin is the only compound that is essentially indigestible in the animal. Lignin is a three-dimensional phenolic structure resulting from the polymerization of 4-coumaryl, coniferyl, and sinapyl alcohols (the monolignol H, G and S units respectively) within the plant cell wall (Figure 2-2a). The resulting polymer is structurally highly variable, making enzymatic breakdown exceedingly difficult.

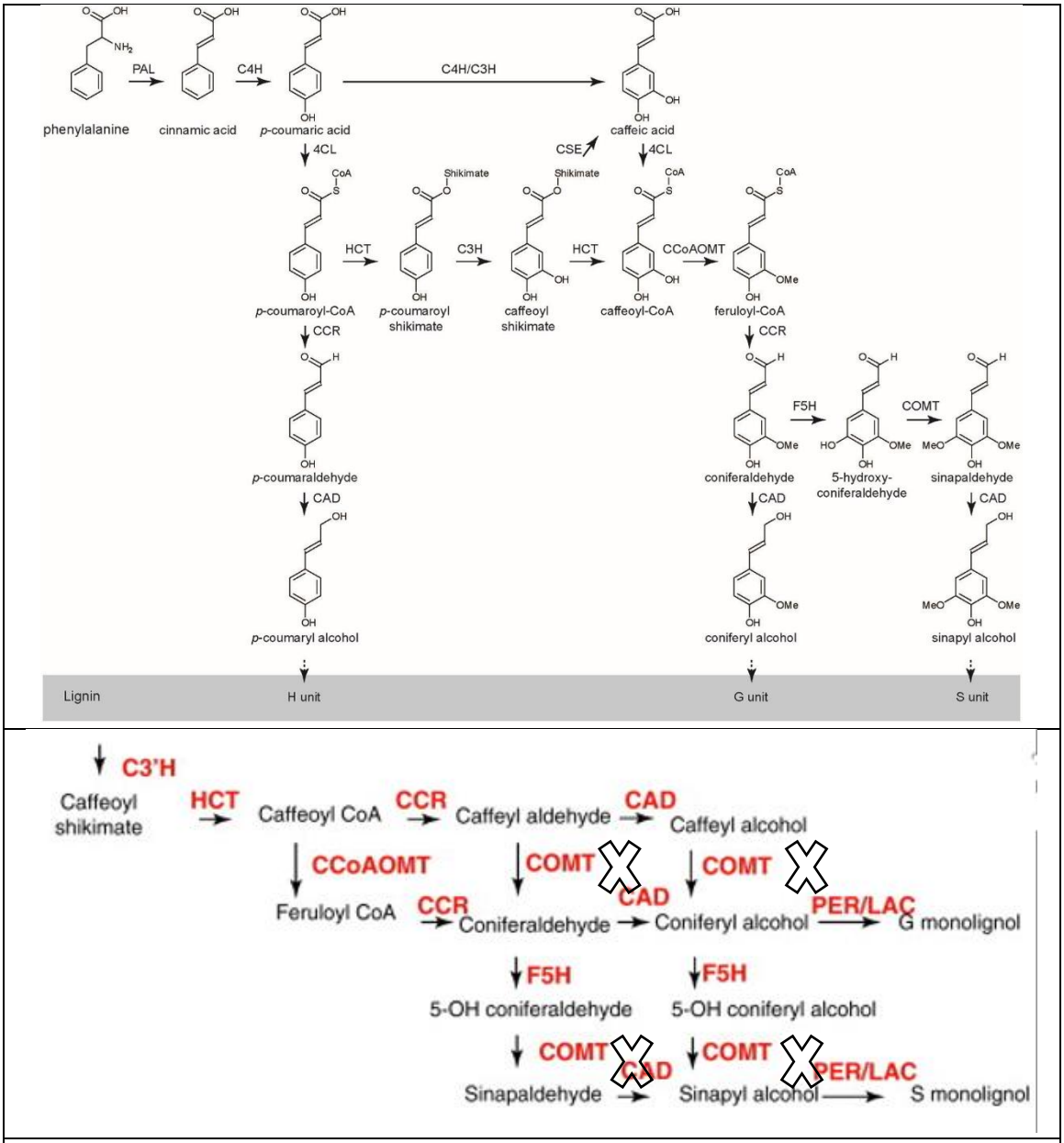


Figure 2-2 **a)** An overview of the monolignol biosynthetic pathway in the model organism *Arabidopsis*. The major enzymes are as follows: phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate:CoA ligase (4CL), hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase (HCT), *p*-coumarate 3-hydroxylase (C3H), caffeoyl shikimate esterase (CSE), caffeoyl-CoA *O*-methyltransferase (CCoAOMT), cinnamoyl-CoA reductase (CCR), ferulate 5-hydroxylase (F5H), caffeic acid *O*-methyltransferase (COMT), and cinnamyl alcohol dehydrogenase (CAD). Modified with permission from (Miedes, Vanholme, Boerjan, & Molina, 2014) **b)** A closer look at G and S unit synthesis, the role of COMT and the implications of a COMT knock-out. Note the reduction in pathways leading to G unit synthesis and the elimination of S unit synthesis. Modified with permission from Zhao & Dixon (2011)

In addition to being indigestible itself; as a three-dimensional structure, lignin interacts with the other cell wall components – reducing an animal's ability to digest them as well. This characteristic makes lignin reduction of particular interest for forage quality improvement.

A mutation that has previously been linked to forage digestibility in *Medicago sativa* (alfalfa) is a mutation in the Caffeic Acid O-Methyltransferase (COMT) gene. This gene encodes for an enzyme which catalyzes the penultimate step in monolignol biosynthesis. This is one step in the complex biosynthesis pathway for lignin.

This mutation causes non-functional caffeic acid O-methyltransferase and results in a COMT-deficient phenotype. COMT functions in the final stages of lignin biosynthesis. Non-functional COMT impedes the production of G units and essentially eliminates the production of S units (Figure 2-2b). Thus, a non-functional COMT mutation not only reduces the lignin content of a plant, it changes the lignin composition (Figure 2-2b).

Such changes cause improved forage quality characteristics even over an extended harvest window. Alfalfa with a COMT mutation that renders the enzyme non-functional has been shown to have an 8-10% reduction in lignin content and a subsequent 3% increase in digestibility (Mertens & McCaslin, 2008). Although further testing is required, it is possible that this mutation may have negative impacts on other alfalfa agronomic traits including "standability", longevity and disease resistance (Casler, 2013; Casler, Buxton, & Vogel, 2002; Miedes *et al.*, 2014).

The lignin biosynthesis pathways and related enzymes are highly conserved. Between alfalfa and aspen the amino acid sequence of the COMT gene is 86% conserved (Gowri *et al.*, 1991). Thus, discoveries made in alfalfa can be used to increase our understanding of other forage crops. As an example, the COMT alfalfa gene mutation is in the same gene as the sorghum *bmr12* mutation and the *bm3* mutation in maize (Sattler *et al.*, 2010). These are two of the mutations responsible for the brown mid-rib phenotype (BMR) in sorghum and maize. BMR is a phenotype associated with reduced lignin and increased digestibility. Although these mutations have been found naturally occurring in maize, the alfalfa mutation was induced (Gorthy *et al.*, 2013; Vignols, 1995).

On a molecular level: The COMT gene has been mutated in a number of ways. In alfalfa these mutations are largely in proprietary genetic material and have not been well-characterized in published literature. More is known concerning mutations in the same gene in sorghum. More than thirty different single base pair mutations have been reported in this gene in sorghum. The majority are mutations that create a premature stop codon, but there are also a number of missense mutations (Sattler *et al.*, 2012). Regardless of the type of mutation, the result is a non-functional protein product

Lignin is a complex polyphenolic propanoid compound that is synthesized not by a single pathway, but by a "metabolic grid" (Casler *et al.*, 2002). Lignin has a three-dimensional structure in the plant cell wall. It is a polymer composed of three monolignol subunits. Mutations in the enzyme catalyzing any particular step could have implications in plant lignin concentration – of course, but also in lignin composition. In alfalfa, a mutation in the COMT gene resulting in a non-functional enzyme eliminates the S

monolignol pathways and reduces the pathways available for G monolignol synthesis. In addition to these changes in lignin composition, there is an overall reduction in lignin concentration (Guo *et al*., 2001; Marita *et al.*, 2003).

Additional lignin mutations and implications for forage quality

While COMT mutations are the most common lignin mutations confirmed to significantly impact digestibility, at least eleven enzymes are utilized in the monolignol biosynthesis pathway from phenylalanine (Figure 2-2a). Mutations in any of these eleven enzymes would be expected to impact lignin synthesis. Currently, mutations that impair 4-coumarate:coenzyme A ligase (4CL), cinnamyl alcohol dehydrogenase (CAD) and caffeoyl-CoA O-methyltransferase (CCoAOMT) have also been found and are also being explored for lignin-reduction potential in prominent forage crops (Gorthy *et al*., 2013; Miedes *et al*., 2014; Zhao & Dixon, 2011).

BMR phenotypes in maize (*bm1*) and sorghum (*bmr6*) (Sattler *et al.*, 2010), and the orange lemma (*rob1*) phenotype in barley (Stephens & Halpin, 2008) are all caused by mutations in a gene encoding for CAD. In maize and sorghum, plants with this mutation exhibited reduced H-, G- and S-subunits (Sattler *et al.*, 2010) and in all instances these mutations resulted in an overall reduced lignin concentration in the plant (Klevenhusen *et al.*, 2019; Sattler *et al*., 2010). Studies in maize and alfalfa have also linked polymorphisms in the CCoAOMT genes with improved fiber digestibility (Brenner *et al.*, 2010; Guo *et al.*, 2001; Li *et al*., 2013; Marita *et al*., 2003).

Forage Varieties Implementing Lignin Pathway Mutations

Low-lignin alfalfa varieties, as the market has branded them, have only just been released in limited availability to growers in the last few years. Currently only two companies have low-lignin alfalfa varieties: Alforex Seeds™ has released two versions of their Hi-Gest lines. These lines have non-transgenic mutations in the COMT gene. While HarvXtra™ is a transgenic variety being released by Forage Genetics International, in which CCoAOMT has been down regulated using RNA silencing techniques. These varieties have only been commercially available since 2015-16 and current yield/quality impacts are only from preliminary studies.

HarvXtra™ -008 is a transgenic reduced lignin (RL) variety created by Forage Genetics International. In a multistate trial in the northern United State the RL line was found to have 8.4% lower ADL and 3.5 to 7.5% better NDF when compared to two traditional alfalfa lines (Arnold *et al.*, 2019). Perhaps even more importantly, it was found that the RL line had similar nutritive value and dry matter yield to non-RL line harvested 5 to 10 days prior. Thus, RL lines could present an extended harvest window for growers relative to traditional cultivars (Arnold *et al.*, 2019).

It is important to note that lignin quantities change over time. As a plant matures, additional cell wall thickening and lignin deposition occurs. Thus, the maturity of a forage impacts its quality – once again related to the proportion of cell wall components.

As plants mature, increased amounts of structural fiber and lignin are deposited. This means that as a percentage of biomass; protein, energy, and minerals decrease. Thus, plant maturity is generally negatively related to forage quality. Already low digestibility

cell wall components such as hemicellulose and cellulose become even less nutritionally available as they are obstructed by lignin deposition.

This means that growers have limited windows in which to harvest a forage at a given quality stage – allowing a forage to mature past a certain point (as a way to increase harvestable yield) will negatively impact the forage digestibility. The ability to have an extended harvest window with lignin pathway mutation lines is perhaps even more important than the increased alfalfa quality of lines at a set maturity. The ability to allow the stand to mature an additional 7-10 days before harvesting means that it is possible to achieve a higher yield while maintaining adequate quality. (Figure 2-3)

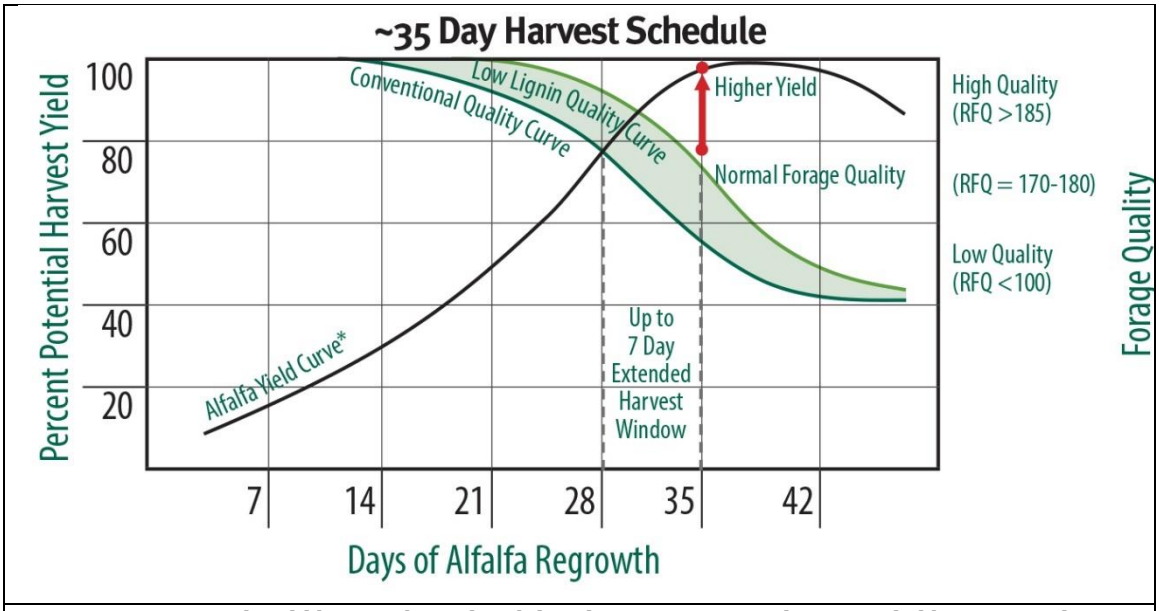


Figure 2-3: Estimated yield boost achieved with low-lignin varieties and an extended harvest window. RFQ = Relative feed quality. Modified with permission from Alforex Seeds™ (Alforex Seeds, 2015).

The COMT, and related mutations could mean significant economic gains for growers. While it remains to be seen how economically viable the newly released varieties will be; reduced lignin in forage is a quality with huge potential impacts on

animal gains and both forage breeders and growers will continue to strive to achieve it. Current varietal releases may have standability or survival issues, but if the gains are there, these issues will be addressed by future breeding work.

Beyond the genes encoding enzymes directly involved in lignin formation, it is also important to consider the genes controlling plant architecture, maturation, secondary cell structure deposition, and senescence. The complex genetic pathways impacting these characteristics also indirectly impact forage traits.

Flowering Time and Photoperiod Genes.

Unlike forage quality, flowering time and photoperiod have been extensively studied in barley. Vernalization, photoperiod, and ambient temperature are known to influence cereal plant development (Ellis *et al.*, 1988; Ellis *et al.*, 1989; Roberts *et al.*, 1988). While the focus has generally been on the impact to head development and grain yield, forage quality and yield are, of course, also impacted by these important physiological pathways and genes.

Stages of barley development have been classified based on reproductive development into pre and post-anthesis, ie fertilization. Pre-anthesis is further broken down into vegetative, early reproductive, and late reproductive phases (González, Slafer, & Miralles, 2002; Slafer & Rawson, 1994). Each developmental phase is regulated by different environmental cues with the responses determined based on genotype (González, Slafer, & Miralles, 2003; Whitechurch, Slafer, & Miralles, 2007).

The major genes known to influence barley maturity can be divided into three main response pathways: photoperiod genes, circadian clock genes, and vernalization

genes. These pathways interconnect to form a complex grid of mechanisms which regulate and coordinate the morphological changes which result in plant flowering.

Barley has been categorized as having two major growth types: winter and spring. Winter barley is sensitive to vernalization and periods of cold are needed for, or will speed up, passage from vegetative to reproductive states. Spring barleys, conversely, do not respond to periods of cold treatment and do not require them for flowering. While barley is generally thought of as being either winter or spring type, genotypes actually fall on a spectrum of the two types (Rollins *et al.*, 2013).

For a typical spring barley, which does not require vernalization, day length and temperature are the two main environmental cues used by the plant to determine when the plant will shift from the vegetative to the reproductive states. Most barley genotypes are facultative long-day, but short-day genotypes also exist (Roberts *et al.*, 1988). For all genotypes, photoperiod and temperature effects interact with each other – the critical photoperiod of a line can vary based on ambient temperature (Ellis *et al.*, 1988; González *et al.*, 2002).

Barley is a genetically diverse crop. While it is typically considered to be a day-length sensitive crop, with changes in day-length needed to trigger flowering, day neutral and early flowering genotypes also exist. Early flowering genotypes are said to contain *early maturity* (*eam*) or *earliness per se* (*eps*) genes. Lines with early flowering traits were favored in regions with short growing seasons or where terminal drought was an issue (Lundqvist, 2009).

Plant development is known to impact the deposition of lignin and other poorly digestible plant cell components and thus impact forage quality directly. However, vernalization, photoperiod, and flowering time genes are also known to indirectly impact forage quality – by impacting plant architecture traits such as height, tillering, and leaf number.

Photoperiod-H1 (Ppd-H1) controls leaf size (Digel *et al.*, 2016): In a genome wide association study of European winter barley cultivars, the *PHOTOPERIOD-H1* (*Ppd-H1*) gene was identified as a candidate gene underlying variation in leaf width and length. Digel *et al.* (2016) hypothesized that Ppd-H1 mediated the induction of *BARLEY MADS BOX* genes *BM3* and *BM8*, which correlated with reduced leaf number and size.

The genetic architecture of barley plant stature (Alqudah, Koppolu, Wolde, Graner, & Schnurbusch, 2016): 218 worldwide spring barley accessions were utilized to investigate the impact of row-type classes (two-row, *Vrs1* vs six-row, *vrs1*) and photoperiod response (photoperiod sensitive, *Ppd-H1* vs reduced photoperiod sensitivity, *ppd-H1*) on tillering and plant height. Two-row plants (*Vrs1*) were found to be taller and had more tillers than six-row plants. Plants with reduced photoperiod sensitivity (*ppd-H1*) had more tillers than wildtype plants.

QTL controlling biomass yield and forage quality in perennial wildrye (S. R. Larson, Jensen, Robins, & Waldron, 2014): While the genetics of forage traits in barley has not been well studied, research exploring this topic does exist for other grass forages. Looking beyond barley, in perennial wildrye, in post-flowering plants QTL related to

genes controlling plant height and flowering time have been linked to fiber digestibility and biomass yield. Larson *et al*. found that plant height QTL co-located with dry matter yield and that flowering time QTL colocalized with QTL controlling NDF, ADF, and acid detergent lignin (2014). In contrast to post-flowering wildrye, in pre-flowering trials the detected forage quality traits were affected by the *COMT-1* locus rather than flowering time or plant height (Larson *et al.*, 2012).

Estimating Forage Digestibility

It is possible to estimate the ingestibility/digestibility characteristics of a forage using analytical proxies – numbers we can obtain from laboratory, chemical analyses. The two most widely utilized tests are Neutral Detergent Fiber (NDF) and Acid Detergent Fiber (ADF). These tests provide an estimate of intake and digestibility, respectively. NDF provides an estimate of the proportion of an analyzed forage that is made up of hemicellulose, cellulose and lignin. ADF is a subsequent analysis in which the hemicellulose is removed, leaving only cellulose and lignin (Figure 2-4). For both NDF and ADF, a higher quality forage will have lower values.

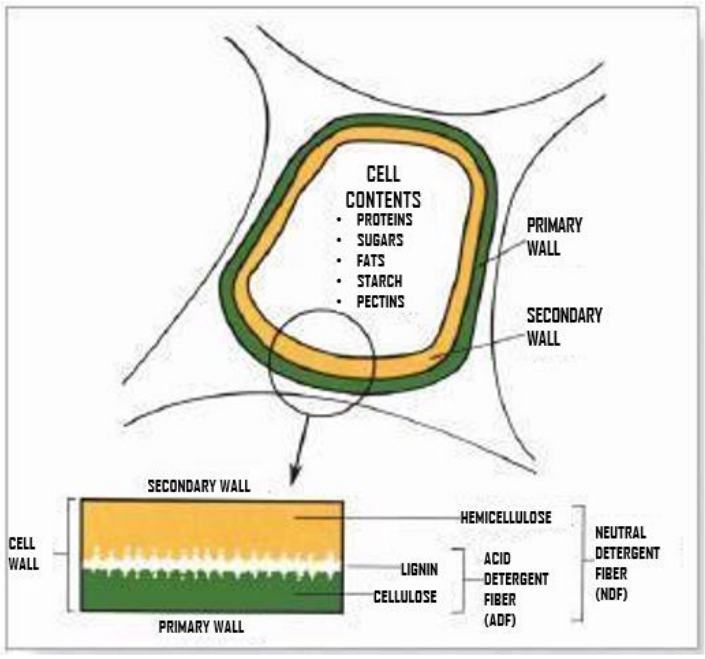


Figure 2-4 The primary locations of major plant structural components and the analytical tests used to measure each. Modified from Bittman (1999).

While true digestibility can only be measured in terms of animal performance and feed utilization efficiency – and thus requires feeding trial experiments – in practice this trait is approximated by a number of laboratory tests. All of the commonly used laboratory tests approximate digestibility through chemical and/or enzymatic breakdown of a small, finely ground forage sample. The higher the degree of breakdown, the better the digestibility of the forage. This breakdown directly correlates to the animals' ability to breakdown a given forage via digestion and utilize the nutrients present in the plant biomass.

Phenotyping hundreds of lines for forage quality is cost prohibitive for the majority of public forage breeders. Using wet chemistry techniques, testing a single sample for NDF/ADF alone can cost anywhere from $20-50, depending on the accuracy

of the technique. Because of the expense, breeders cannot incorporate digestibility testing in the early stages of the breeding pipeline – when selection would be the most beneficial. Most breeders simply have too many potential lines at the F4 stage to test them all for forage quality. Instead, a breeder will likely have to wait to test a line's potential fiber value until that variety is close to being considered for release as a potential forage variety. Many excellent forage lines are likely lost in the early stages of line selection simply because we do not have all of the relevant phenotypic data.

Until suitable molecular markers are available, or modeling can be used to identify easily measured proxies, it will be difficult for breeders to meet the demands of growers for the release of improved forage varieties. As noted above, a one percent improvement in forage digestibility correlates to a three percent improvement in animal average daily gain – a significant economic gain for growers and producers alike.

Modeling and Predicting Forage Yield and Quality

As noted above, due to the underlying molecular complexity of this trait, no single barley gene or genetic markers are currently available to aid breeders in selecting for high barley forage digestibility. Currently, breeders must screen all potential lines via laboratory proxy measurements to estimate the digestibility of any given line. Phenotyping hundreds of lines for forage quality is cost prohibitive for the majority of forage breeders. Because of the expense, breeders cannot incorporate digestibility testing in the early stages of the breeding pipeline – when selection would be the most beneficial. Until suitable molecular markers are available for marker-assisted-selection or modeling

can be used to identify easily measured proxies, it will be difficult for breeders to meet the demands of growers for the release of improved forage varieties.

Alfalfa is, by far, the most intensely studied forage crop. In pure stands of alfalfa some proxy-based modeling has been successful in predicting first cutting forage quality. Models for alfalfa have used growing degree-days, plant height, and stage of maturity to predict forage quality (Fick & Onstad, 1988; Hintz & Albrecht, 1991). These models were developed as tools to help growers know when to time the harvest of their crop, but the success of the models suggests a connection between the terms of the model and the quality trait that could also be leveraged by a forage breeder.

The models are generally considered to be of two types: Predictive Equations for Alfalfa Quality (PEAQ) and Growing Degree Days (GDD) models, described in detail below.

PEAQ Models

In 1991, Hintz and Albrecht published the first PEAQ model for predicting forage chemical composition based on morphological proxies. They studied fifteen morphological traits on five alfalfa cultivars across three environments in Wisconsin. Depending on the year, they collected as many as four cuttings per location. The physical traits measured included nodes per plant, plant height, mean stage count, mean stage weight, and a number of additional stage-of-development descriptors. Root mean square error (RMSE) values were used to assess the accuracy of multiple regression models that utilized two or three of the morphological terms. Based on R^2 values, the best model for NDF and ADF included the forage sample's maximum stage of development (MAX),

mean stem height weighted by plant mass (MHW), and the fraction of the sample's dry weight that was stems and stipules (STEM). With these terms the models explained 92 and 91 percent of the variation for NDF and ADF respectively. Both MHW and STEM require time consuming steps where the collected forage sample is broken up into fractions or categories, and thus both traits are labor intensive to measure. Ultimately, the models selected by the researchers included only the two most easily measured morphological traits: height of tallest stem (MAXHT) and maturity stage of most mature stem (MAX). With just these two traits, the models still had R^2 values of 0.88 and 0.89 for ADF and NDF respectively (Hintz & Albrecht, 1991).

The findings of Hintz and Albrecht have been validated for pure stands of alfalfa in additional environments in Wisconsin (Owens, Albrecht, & Hintz, 1995; Sulc *et al*., 1997), New York, Pennsylvania, Ohio, California (Sulc *et al*., 1997), New Mexico, Northern Mexico (Santillano-Cázares *et al*., 2014), Central Europe (Hakl *et al.*, 2019), and Northern Europe (Andrzejewska, Contreras-Govea, & Albrecht, 2014). The PEAQ equation was shown to perform well across a wide range of environments. Sulc *et al.* (1997) found that NDF values estimated using PEAQ were within 4 percent of the wet chemistry values for 84 to 89% of samples tested (1997).

More recently researchers have also shown that the PEAQ model appeared to be robust for alfalfa cultivars that have been developed with reduced lignin (RL) (Arnold *et al.*, 2019) and for mixed alfalfa-grass stands (Parsons, Cherney, & Gauch, 2006; Parsons, Peterson, & Cherney, 2013; Wood *et al*., 2018; Wood *et al.*, 2019). In alfalfa-grass

mixtures, growing degree days and the grass contribution to total yield were also required in the PEAQ model.

GDD Models

Growing degree-days are an index that describes the accumulated amount of heat that plants have experienced. GDD is based on the minimum and maximum temperatures at which a plant can operate and so are calculated differently for different species. For alfalfa, GDD in ^{o}F is calculated by averaging the 24-hour maximum and minimum temperatures and subtracting a base temperature of 41 ^{o}F. These values are then summed to find a seasonal total GDD.

Models using GDD to predict forage quality of alfalfa have been developed (Allen & Beck, 1996; Allen *et al.*, 1992; Cherney, 1995; Fick & Onstad, 1988; Sanderson, 1992), but have limitations. GDD has been shown to be a moderately good predictor of forage quality when soil moisture is not a limiting factor, but not under moisture limiting conditions (Van Soest, 1996). Thus, GDD models can be useful in spring or high moisture environments but should not be used when soil moisture may be limiting.

Another factor that limits the usefulness of GDD models is a lack of robustness across years and environments. Model results have been shown to be variable depending on the state (Cherney, 1995; Sanderson, 1992), or across years (Allen & Beck, 1996). The lack of model stability is a serious flaw of GDD models and demonstrates the potential issues with extrapolating a model built in one time and place to use in another time and place.

Future Models and Modeling Goals

The majority of forage quality modeling has been developed for use in pure stands of alfalfa and no models currently exist for forage barley. These models were developed as tools to help growers know when to time the harvest of their crop, but the success of the models suggests a connection between the terms of the model and the quality trait. This connection could be leveraged by the forage breeder. If terms such as stem height and maturity stage are predictive of forage quality, these or similar terms could be promising as proxies for selection in a forage breeding program.

Current State of Barley Forage Quality Research

Relative to alfalfa, maize, or any of the more commonly studied forage species, little work has been done on barley as a forage. Since barley has been shown to be a useful model organism for genetic studies, the extent of forage research in the crop has largely been limited to genetic studies in bi-parental mapping populations.

QTL Mapping of Steptoe-Morex Population

In barley, the only genetic diagnosis papers that have been published focus on QTL discovery. Siahsar *et al.* and Taleei *et al.* found eleven unique, statistically significant QTLs related to forage digestibility in 72 dihybrid lines derived from a Steptoe X Morex dihybrid mapping population (Siahsar *et al.*, 2009; Taleei, Siahsar, & Peighambari, 2009). Five QTLs were significant for dry matter digestibility, four QTLs for acid detergent fiber (ADF), and five QTLs for neutral detergent fiber (NDF). Of these fourteen QTLs, three were found to be related to multiple fiber values, leaving eleven

unique fiber QTLs. Six of these QTLs were significant in two different environments, while the remaining five QTLs were significant in only one environment. In these studies, the 74 lines (72 dihybrid lines + 2 parental lines) varied significantly in fiber digestibility. Between the values of the most and least digestible lines, there was a 14.8% difference in dry matter, a 33.9% difference in ADF and a 28.6% difference in NDF (Siahsar *et al*., 2009; Taleei *et al.*, 2009).

In the same Steptoe X Morex dihybrid mapping population, Suber *et al.* examined 145 dihybrid lines in two environments. In this experiment, the researchers found fourteen forage quality QTLs, five loci were previously mapped by Siahsar and Taleei and nine unique QTLs were recorded (Surber *et al.*, 2011).

QTLs related to forage digestibility in barley have been found in the Steptoe-Morex bi-parental mapping population (Siahsar *et al.*, 2009; Surber *et al.*, 2011; Taleei *et al.*, 2009). However, the significance of these QTLs across populations is yet to be determined and no work has yet been published concerning MAS or modeling for this trait.

Outside of barley – in more intensely studied forage crops – more is known about at least some of the genes related to forage digestibility. Still, although some genes have been related to digestibility, these few genes are far from being the whole story and may, in fact, have only a small role in the phenotypic variation within and between most populations.

Conclusions

Correlating fiber digestibility with a molecular mechanism, single gene, or even a few genes is difficult. Because of their low digestibility, three major biochemical compounds are closely linked to the digestible fiber available in a given plant: cellulose, hemicellulose, and lignin. Any gene related to the deposition of lignin, cellulose, and hemicellulose may be tied to the forage quality of a species. Based on the genome of the model organism *Arabidopsis thaliana*, even just considering the lignin biosynthesis pathway alone, there are at least 70-80 genes encoding enzymes directly involved in lignin formation (Lubberstedt, 2007).

As noted above, due to the underlying molecular complexity of this trait, no single gene or genetic markers are currently available to aid breeders in selecting for high forage digestibility in barley. Currently, breeders must screen all potential lines via laboratory proxy measurements to estimate the digestibility of any given line.

Although there are QTLs published as being related to this digestibility, as of yet they have only been found in a single, bi-parental mapping population (Siahsar *et al*., 2009; Surber *et a.*, 2011; Taleei *et al.*, 2009). The significance of these QTLs across populations is yet to be determined and no work has yet been published concerning MAS for this trait.

Modeling forage quality in barley could be an alternative to MAS for early selection of forage traits within a breeding program. Extensive modeling in alfalfa and alfalfa-grass stands have demonstrated that models utilizing easily measured field traits can be used to predict forage quality. Such models have been shown to be robust across a

number of different cultivars, locations, and years. Although these models were developed for prediction, they are evidence of an underlying connection between the measured traits and forage quality. Such a connection could provide useful to forage breeders. It is possible that traits that important in predicting forage quality could be leveraged as proxies for forage quality.

CHAPTER THREE: GENOME WIDE ASSOCIATION MAPPING OF FORAGE YIELD AND QUALITY TRAITS IN A GLOBAL POPULATION OF SPRING, 2-ROW BARLEY

Introduction

Cow-calf operations account for about 50% of the annual GDP of the state of Montana and winter feeding costs are the largest expense for most operations (USDA National Agricultural Statistics Service, 2020). While alfalfa hay remains a popular option for forage production in Montana, annual cereal forages such as barley provide a valuable alternative. Where alfalfa requires a year to establish, annual cereals grow rapidly and yield returns in a single season, providing flexibility and affordability for growers. Additionally, barley is more drought and saline tolerant than alfalfa (Maas, 1993; Steppuhn & Raney, 2005), making it well adapted to areas without access to irrigation.

Forage yield and digestibility are of the utmost importance to livestock producers and forage growers, and thus they are major traits of interest to forage breeders. Increasing biomass yield is of obvious benefit as more forage can be produced on fewer acres, but several studies have shown that there is also a direct, positive relationship between digestibility and animal productivity. A conservative estimate is that a one percent increase in forage digestibility can lead to a three percent increase in the average daily weight gains of steers. (Casler and Vogel, 1999, Mohammed, 1967)

Forage digestibility is a complex quantitative trait. It is the phenotypic result of the interaction of many different genes, most of which have yet to be identified. Although a few mutations (mostly enzymes in the lignin pathways) with significant implications for digestibility have been discovered or generated in species such as alfalfa and maize (Brenner, 2010), little work has been done to study the genetics of forage digestibility in barley. Animal intake and forage digestibility are negatively impacted by high fiber content (Mertens, 1987; Reid, 1988) and are commonly approximated by measuring neutral detergent fiber (NDF, animal intake), and acid detergent fiber (ADF, digestibility).

Early quantitative trait loci (QTL) mapping in barley for forage traits was performed using the Steptoe x Morex biparental mapping population with a low-density linkage map of 327 markers (Hayes, 1993; Kleinhofs, 1993). Although Steptoe is a feed barley and Morex a malting barley, this population has been utilized to successfully identify loci influencing forage traits. Harvesting at dough stage, Siahsar *et al.* (2009) found QTL for ADF on chromosomes 1H, 2H, 3H, and 5H; and for NDF on 1H, 2H, 3H, 5H, and 6H in 72 lines from this population. Using 145 of the lines from the same population, Suber *et al.,* (2011) harvested forage samples at two different maturity points, anthesis and "peak forage yield". At anthesis, QTL for ADF were identified on chromosomes 1H, 2H, and 7H, while QTL for NDF were found on 2H, 4H, and 7H (Suber *et al.*, 2011). At the later stage of maturity however, QTL for both ADF and NDF were only found on chromosome 2H (Suber *et al.*, 2011). The findings of these QTL studies are limited by several factors: the relatively narrow genetic diversity of the

biparental population, large linkage groups due to a lack of genetic recombination, and a small number of markers with, at the time, no physical map.

Technological advances have since made possible a barley 9K single nucleotide polymorphism (SNP) genotyping chip, which, as part of the USDA-NIFA funded Triticeae Coordinated Agricultural Project (TCAP), was used to genotype 1,860 unique barley accessions from the USDA-ARS National Plant Germplasm System (NSGC) (Munoz-Amatriain *et al.*, 2014). The NSGC is a global germplasm repository which contains more than more than 33,000 *Hordeum vulgare L.* accessions collected from over 100 countries and categorized based on their level of genetic improvement. The 1,860 genotyped lines were chosen to represent the total diversity of the NSGC barley collection and were designated the USDA Barley World Core Collection informative Core (iCore). Unlike biparental mapping, genome-wide association mapping studies (GWAS) can be conducted on a population of unrelated individuals, allowing the selection of genetically diverse and recombinant rich populations. Demonstrating the utility of the approach, GWAS have been utilized successfully to map QTL for forage quality and yield in alfalfa (Biazzi *et al.*, 2017; Jia *et al.*, 2017; Lin *et al.*, 2020; Liu *et al.*, 2017; Sakiroglu *et al.*, 2017; Yu, 2017), maize (Wang *et al.*, 2016; Lopez-Malvar *et al.*, 2019), sorghum (Li *et al.*, 2018; Zhao *et al.*, 2016), and perennial ryegrass (Arojju *et al.*, 2016).

The objectives of this study were (i) to identify iCore lines with superior forage traits and (ii) to use an association mapping approach to identify QTL related to key forage traits.

Materials and Methods

Plant Materials

The iCore is made up of 1,860 lines that have been genotyped with a barley SNP iSelect platform with 7,842 markers (Munoz-Amatriain *et al.*, 2014). Only 621 lines in the iCore are spring, two-row barley, which is the type most commonly grown in the region. From the spring, two-row subset of the iCore, 260 barley lines were randomly selected for testing and this new population was designated WC1. The WC1 lines originated from 69 different countries and the breakdown of their genetic improvement classification was 52% cultivar, 38% landrace, and 10% genetic and breeding accessions.

To validate the QTL findings from WC1, a second set of 260 lines was randomly selected from the remaining untested spring, two-row iCore and designated WC2. Lines from 57 different countries were included in WC2; 51% of the lines were classified as cultivars, 29% as landrace, and 20% as genetic and breeding accessions.

Field Trials

The 2016-17 WC1 trials were planted between May 4-6, and the 2018 WC2 trial was planted on April 28. Field plots in 2016 were planted as two rows plots, 8 ft in length and planted on 9 ft centers. Plots in 2017 and 2018 were planted as three row plots, 15 ft in length and planted on 18 ft centers. All plots were seeded at a rate of 1.1 g/ft^2 with a row spacing of 1 ft. The WC1 was grown under both irrigated and rainfed (dryland) treatment conditions in field trials at the Arthur H. Post Research Farm in Bozeman, MT (Latitude: 45.6729, Longitude: -111.1547, Elevation: 1455 m) on Amsterdam silt loam soil in both the 2016 and 2017 field seasons. The irrigated treatment received 5 inches of

irrigation applied to the field in two, 2.5 inch increments immediately prior to the heading stage of development.

Each WC1 field was planted in an augmented randomized complete block design (Federer 1956; 1961) with 10 blocks and 4 check plots per block. The check varieties, Lavina, Hays, Conlon, and Stepford, were randomly assigned plots within each block and the 260 experimental lines were randomly assigned to the remaining plots. This design resulted in 30 plots per block and 300 plots per treatment: 260 plots for the genetically diverse lines and 40 plots for the checks. Across both dryland and irrigated there were a total of 600 plots per year and 1200 plots across both years.

Agronomic Data Collection

Plots were monitored daily throughout the growing season and data were collected on plot heading date (Zadoks stage 51), soft-dough date (Zadoks stage 85), maturity date (Zadoks stage 92), and mature plant height. Zadok stages were defined as when 50% of the main tillers in a plot reach the designated stage. The Zadok stage 92 definition was modified to be assessed visually, where maturity was defined as when no green color remained in 50% of the main tiller heads, including the awns and glumes. From heading, soft-dough, and maturity dates, two additional parameters were calculated: the number of days between heading and soft-dough, and the number of days between soft-dough and maturity. Plant height was taken on mature plants from ground level to the top of the plant excluding awns.

When a plot reached the soft-dough stage of maturity, six 6-inch forage samples were collected from representative areas of the plot and bulked. Samples were cut one

inch from the soil surface. Collected forage samples were dried, weighed, milled to pass through a 2mm screen, and analyzed. From each forage sample, samples were analyzed for ADF and NDF and an estimated dry ton/acre biomass yield was calculated. It is important to note that although digestibility is the forage quality variable of interest, the laboratory proxies ADF and NDF measure % non-digestibility – thus, as ADF or NDF values decrease, digestibility and intake respectively increase. In addition to forage trait data, grain yield was also collected on each plot. If they are to be successfully produced by seed growers for the forage market, forage barley varieties must have sufficient grain yield and selection for grain yield is requisite for cultivar success. Best linear unbiased prediction (BLUP) values of each trait of each line were obtained and used as phenotypic data for mapping analysis (see Statistical Analysis below).

The WC2 population was also grown at the Arthur H. Post Research Farm but only under rainfed (dryland) during the 2018 and 2019 field seasons. Experimental design, data collection, and sampling were the same as described for WC1.

Marker and Map Data

The SNP data and genetic map used for this study were downloaded from the Triticeae Toolboox public repository (https://triticeaetoolbox.org). The lines were genotyped with the barley iSelect SNP chip as previously described (Munoz-Amatriain *et al.*, 2014). The Morex 2016, IBSC physical map was utilized (Beier *et al.*, 2017). Markers with a minor allele frequency of less than 5% or which were missing more than 50% of allele data were removed, resulting in 6,585 remaining markers. Of the remaining markers, 6,383 had known marker positions while 202 markers were unaligned.

Statistical Analysis

All data processing and statistical modeling were performed in the open-source statistical platform R (R Core Team, 2020) through the graphical user interface RStudio (RStudio Team, 2020). To adjust WC1 agronomic data for across field variation within an environment, a mixed model with factors for block, check, and accession was used to estimate best linear unbiased predictors (BLUPS) using the restricted maximum likelihood (REML) method in the 'lme4' package (Bates *et al*., 2015). BLUPs were adjusted with environment trait grand means to restore traits to original scales. Agronomic trait BLUPs for each accession were then averaged within and across treatments to generate accession level BLUP data for the Dryland, Irrigation, and Average treatments. Pearson's correlation coefficient was calculated to estimate the correlation between agronomic trait BLUPs across treatments. All statistical analyses were then repeated for the WC2 population.

To identify WC1 lines with superior forage traits the lines were ranked by total digestible biomass yield and grain yield across treatments. Total digestible biomass yield was estimated as dry biomass yield * % Total Digestible Nutrients (%TDN), where %TDN was calculated as (98.625 – (%ADF * 1.048)) (Catling, McElroy, & Spicer, 1994). Based on trait rankings, with high digestible biomass yield and high grain yield being desirable, a summary of lines with superior forage traits was generated.

Population Structure

The population structure of both WC1 and WC2 was inferred by conducting a principal component (PC) analysis. Missing marker data was imputed separately for each population using the expectation maximization (EM) algorithm in the "A mat" function of the rrBLUP package. The EM algorithm replaces missing marker data based on the multivariate normal distribution of the population (Endelman, 2011). The resulting genotypic data including imputed marker values were then scaled and centered using the "scale()" function and the "svd()" function was applied to the scaled data. Multiplying a matrix of the scaled data by the "v" vector created by the svd function produced the principal component values and a screeplot was utilized to assess the variation accounted for by the first 10 PCs. By the elbow criterion, the screeplots for both WC populations indicated that the first 2 PCs should be utilized for subsequent statistical procedures. In the WC1 population, the first 2 PCs accounted for 12.8 and 5.2% respectively of the original variation while in the WC2 population PC1 and PC2 accounted for 11.5 and 4.1%.

To further assess possible relatedness within the WC1 and WC2 populations, separate cluster analyses were performed. For each population the scaled genotypic data with imputed marker values were used to generate a Euclidean distance matrix and the results were then clustered using the "ward.D2" method in the "hclust" function of the cluster package (Maechler *et al*., 2019). Ward's clustering method is a hierarchical, agglomerative clustering method, which does not require prior specification of the number of clusters. The underlying algorithm attempts to minimize the variability within

clusters while maximizing the variability between clusters. Ward's method tends to create clusters of similar sizes, avoiding a pit-fall of clustering called "chaining".

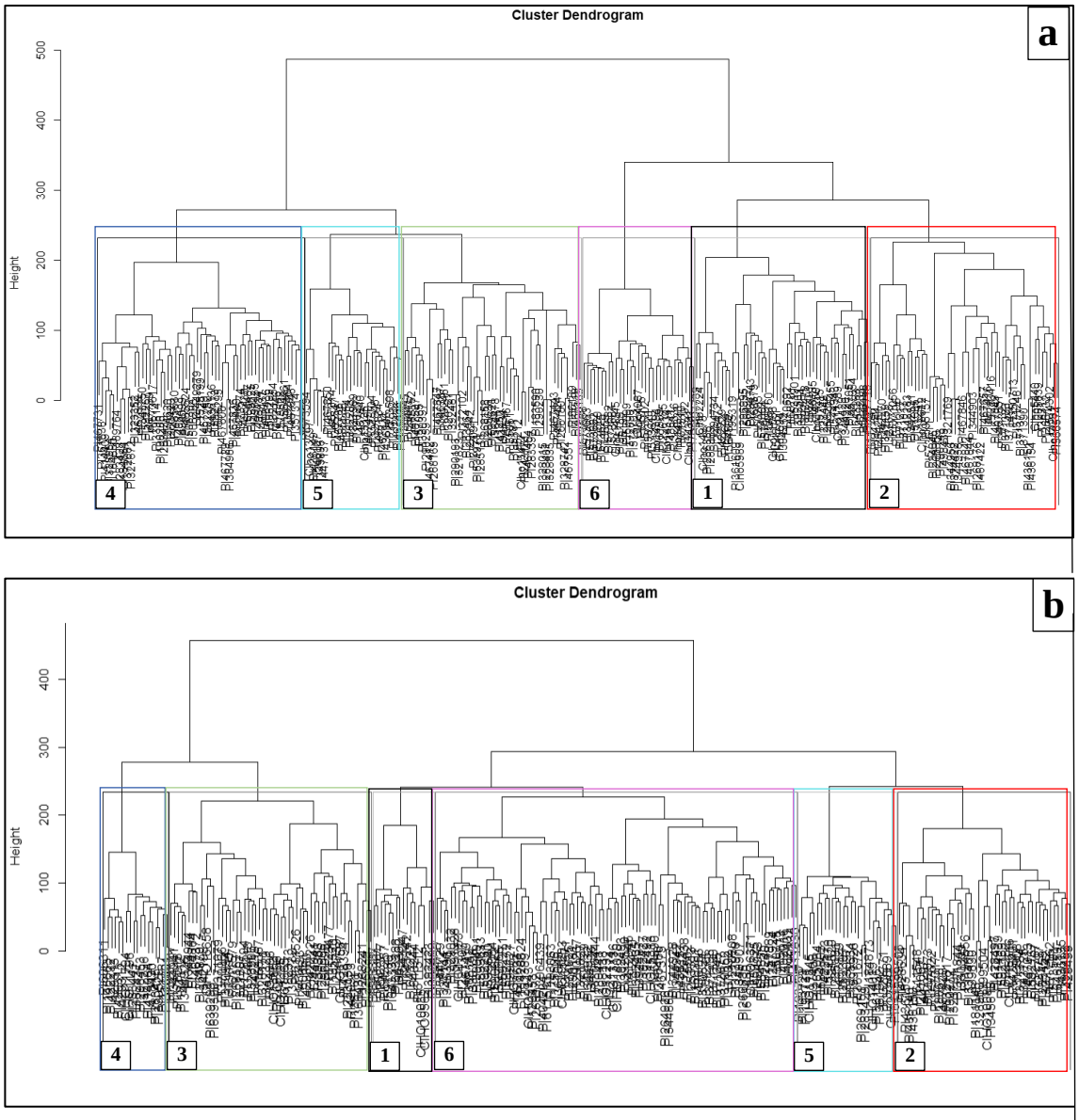


Figure 3-1: Dendrogram of hierarchical cluster analysis of Euclidean distances of WC1 (a) and WC2 (b) genotypic data via Ward's clustering method. The subpopulation numbers between the WC1 and WC2 are determined by the dendrogram cutting order and do not correspond between populations. A six-cluster solution is depicted for each of the separate cluster analyses.

Both the pseudo F-statistic (Calinski & Harabasz, 1974) and Fang and Wang's (Fang & Wang, 2012) bootstrapping approach were utilized in R to determine the appropriate number of clusters for the clustering solution for each population, but in each case both tests proved inconclusive. For both populations the pseudo F-statistic, implemented in the "clusterSim" package (Walesiak & Dudek, 2017), indicated cluster solutions of either one or two clusters while the bootstrapping method, implemented in the "fpc" package (Hennig, 2015), minimized at the maximum value considered. Figure 3-2 depicts the plotted results of these two cluster assessments.

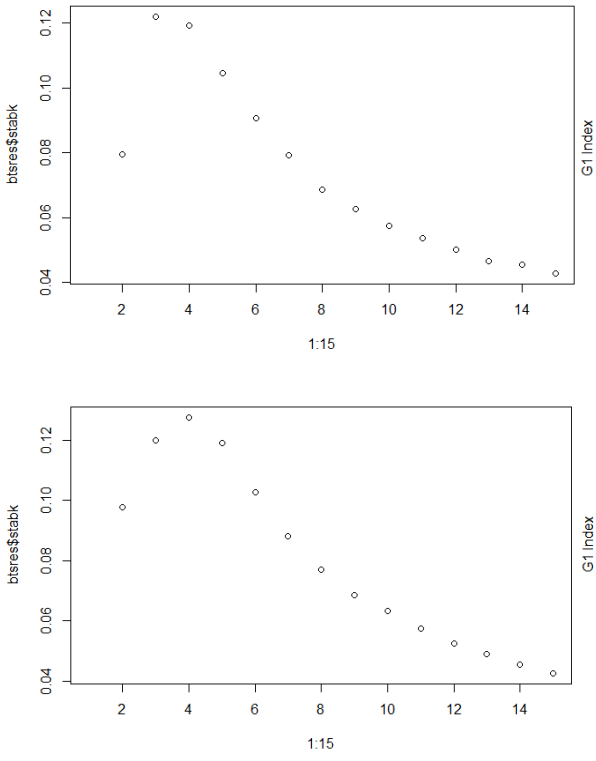


Figure 3-2a: Plot of bootstrapping stability results vs number of clusters showing a minimized value at the maximum cluster value examined. WC1 top, WC2 bottom.

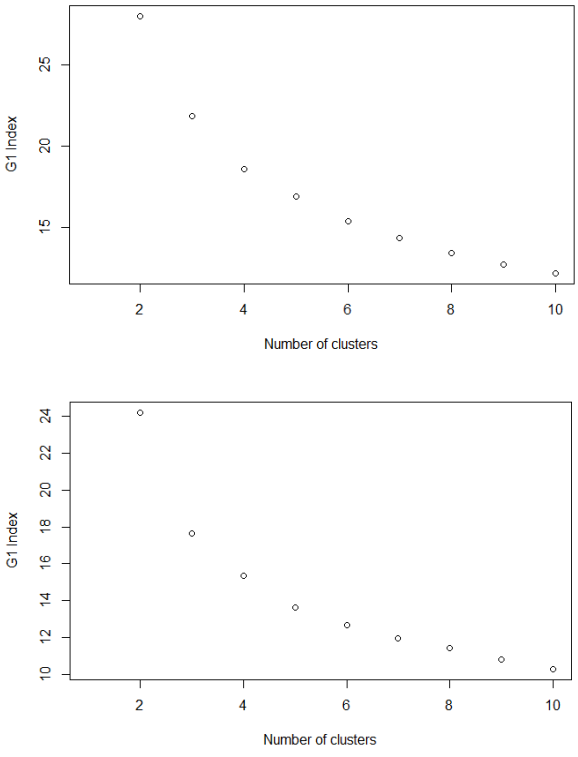


Figure 3-2b: Plot of G1 index vs number of WC1 clusters showing one or two clusters as optimal. WC1 top, WC2 bottom.

Because the results from the numerical measures were inconclusive, the clustering results were plotted with dendrograms and, based on the height of the clustering cuts, a cluster number (k) of 6 was selected for each. The heights of the cluster cuts were used to determine the optimal number of clusters as described in Everitt and Hothorn (Everitt & Hothorn, 2011). With 260 lines being clustered in each population, it is not possible to distinguish most of the accession labels in the dendrogram, however the dendrogram is still useful for illustrating the overall structure demonstrated by the clustering technique (Figure 3-1).

The results of the cluster analysis can also be used to better understand the PCA results. By coloring the PC coordinates based on the cluster analysis, the WC1 sub-populations within the PCA are evident in the first two dimensions. This can be seen in Figure 3-3.

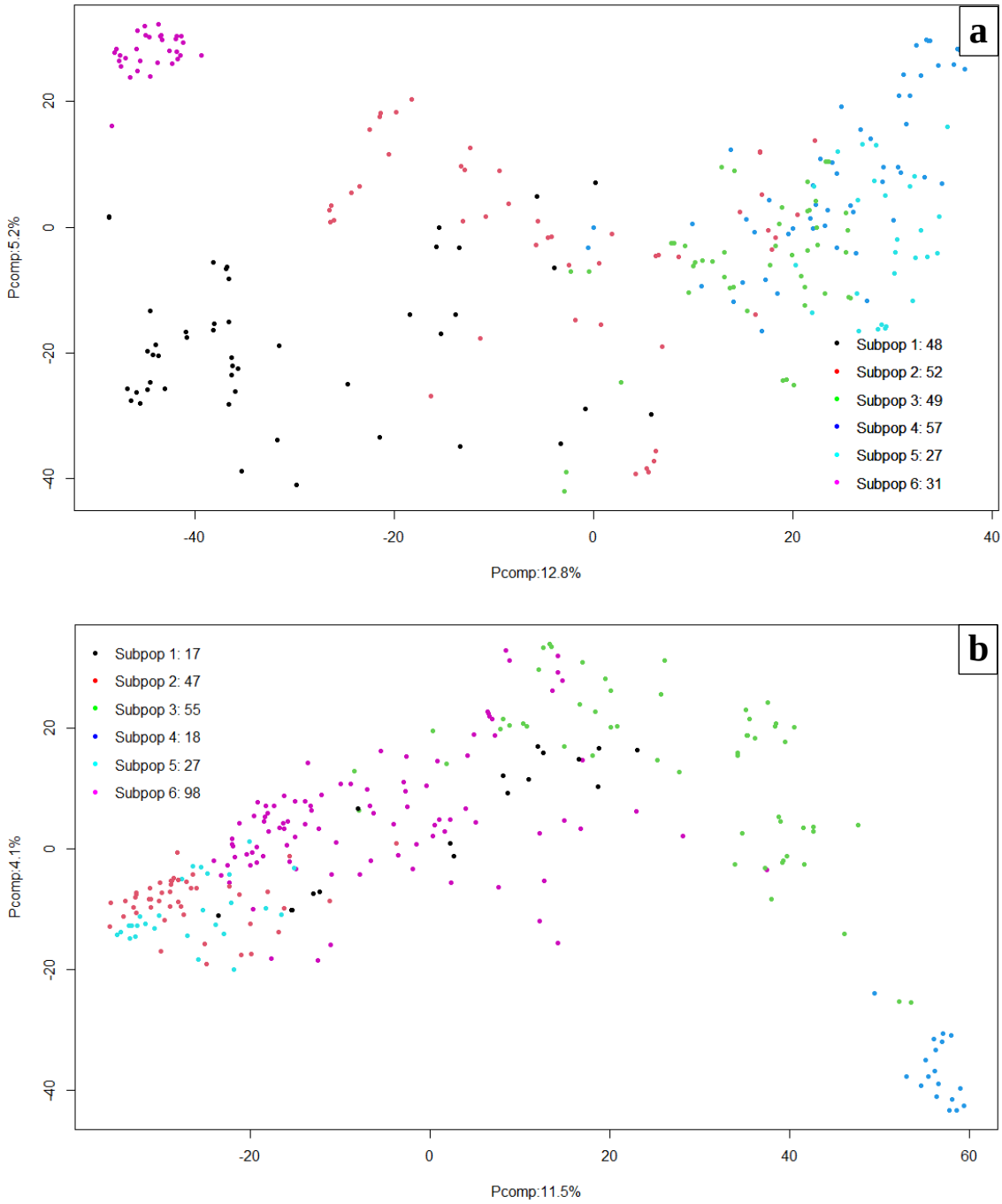


Figure 3-3: Scatterplot of PC coordinates of PC1 vs PC2, colored by clustering group (Subpop). WC1 shown in a), WC2 shown in b).

Genome-Wide Association Mapping

Genome-Wide Association Mapping (AM) was conducted using the marker dataset and BLUP adjusted phenotypic datasets from WC1 field trials as follows: line trait values averaged across dryland environments (Dryland), line trait values averaged across irrigated environments (Irrigated), and line trait values averaged across all environments (Average). The WC2 field trial was analyzed as a single dryland environment.

The AM analysis was conducted in the R (R Core Team, 2020) package Genome Association and Prediction Integrated Tool (GAPIT) (Lipka *et al.*, 2012) using the Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) method (Huang *et al*., 2019). Population structure was incorporated into the model with the inclusion of the first two principal components as model fixed effects while cryptic relationships between individuals was accounted for through a kinship matrix calculated with the VanRaden algorithm (VanRaden, 2008). BLINK tests markers for linkage disequilibrium (LD) and eliminates those that are in LD with the most significant markers. GAPIT allows for the number of groups in the kinship matrix to be optimized on a trait-by-trait basis. Optimal group kinship matrices are defined as those which maximize genetic variance and total variance explained while minimizing the 2*log likelihood statistic. Models excluding population structure, kinship, or both were also tested but were discarded based on the observation of non-normality of residuals in quantile-quantile plots. Marker score $-\log_{10}(p)$ values were generated and a false discovery rate (FDR) of 0.05 was used as a threshold for significance in all analyses.

Possible pleiotropic effects of forage quality marker-trait associations on biomass yield, grain yield, heading date, soft-dough date, and maturity date were also explored. For SNPs determined by BLINK to be significant for ADF or NDF in dryland, irrigated, or across treatments, the impact of the minor allele on all quality, yield, and timing of maturity traits were determined by a Type III ANOVA test.

Results

Field Weather Data

The climate was variable over the course of the study. The 2016 and 2017 field seasons were about 33% drier than the long-term average (LTA) for the location (~5.5 inches compared to 8.4 inches, Table 3-1). Five additional inches of irrigation were applied to the irrigated environments prior to heading for a total of ~10.5 inches of total in-season water. Compared to the LTA, July 2017 was particularly hot and dry with only 0.1 inches of precipitation and nearly twice as many days with a high temperature above 90 ºF (Table 3-1). Hot and dry temperatures in July 2017 shortened the duration between heading and soft-dough dates by a full week when compared to 2016 (data not shown). Precipitation conditions in 2018 and 2019 were more similar to the LTA and temperatures were cooler with few days above 90 ºF in July or August.

Table 3-1 Summary of field season precipitation and temperatures at Arthur H. Post Farm, Bozeman, MT from 2016 to 2019. LTA = Long term average.

Year	May				June				July				August				Total Season Precip (in) (May-Aug)
	Ave high temp (Fº)	Ave low temp (Fº)	Total precip (in)	# Days above 90 (Fº)	Ave high temp (Fº)	Ave low temp (Fº)	Total precip (in)	# Days above 90 (Fº)	Ave high temp (Fº)	Ave low temp (Fº)	Total precip (in)	# Days above 90 (Fº)	Ave high temp (Fº)	Ave low temp (Fº)	Total precip (in)	# Days above 90 (Fº)	
LTA	66	38	2.9	0	74	45	2.8	1	84	51	1.4	8	84	49	1.3	7	8.4
2016	64	39	2.7	0	80	48	0.8	3	83	51	1.2	6	84	49	0.9	7	5.6
2017	67	39	2.6	0	76	46	2.2	2	88	55	0.1	15	83	51	0.6	1	5.5
2018	68	44	2.9	0	72	47	3.6	0	82	51	0.2	3	80	49	1.3	5	8.0
2019	60	39	1.7	0	73	45	2.1	0	80	50	2.7	1	82	50	0.7	0	7.2

Trait Variation by Population and Year

The 2016-17 WC1 irrigated treatment received 5 inches of irrigation prior to barley heading while the 2016-17 dryland treatments relied solely on precipitation. The WC2 confirmation population was grown in 2018-19 under only dryland conditions. The high genetic diversity of the populations resulted in lines which were highly variable for all the traits measured. Within an environment ADF, NDF, and forage yield varied by as much as 23%, 21%, and 360% respectively. Likewise heading, soft-dough, and maturity

dates varied by as many as 30 days, 31 days, and 23 days respectively from the earliest to the latest maturing lines within an environment (Data not shown).

The irrigation treatment in the WC1 population had a significant impact on all measured traits with the exception of soft-dough date. Irrigation was related to an increase in forage and grain yield, a reduction in forage quality, and an overall delay in plant maturation (Table 3-2). Compared to 2016-2017, the 2018-19 field seasons were cooler and wetter, and this is reflected in the higher grain and forage yield values of the WC2 (Table 3-2).

Table 3-2: Trait estimates by experimental population and treatment. [a] Significance of irrigation treatment across WC1 years, *** indicates that the trait was impacted by irrigation at the 0.0001 probability level, ns is not significant

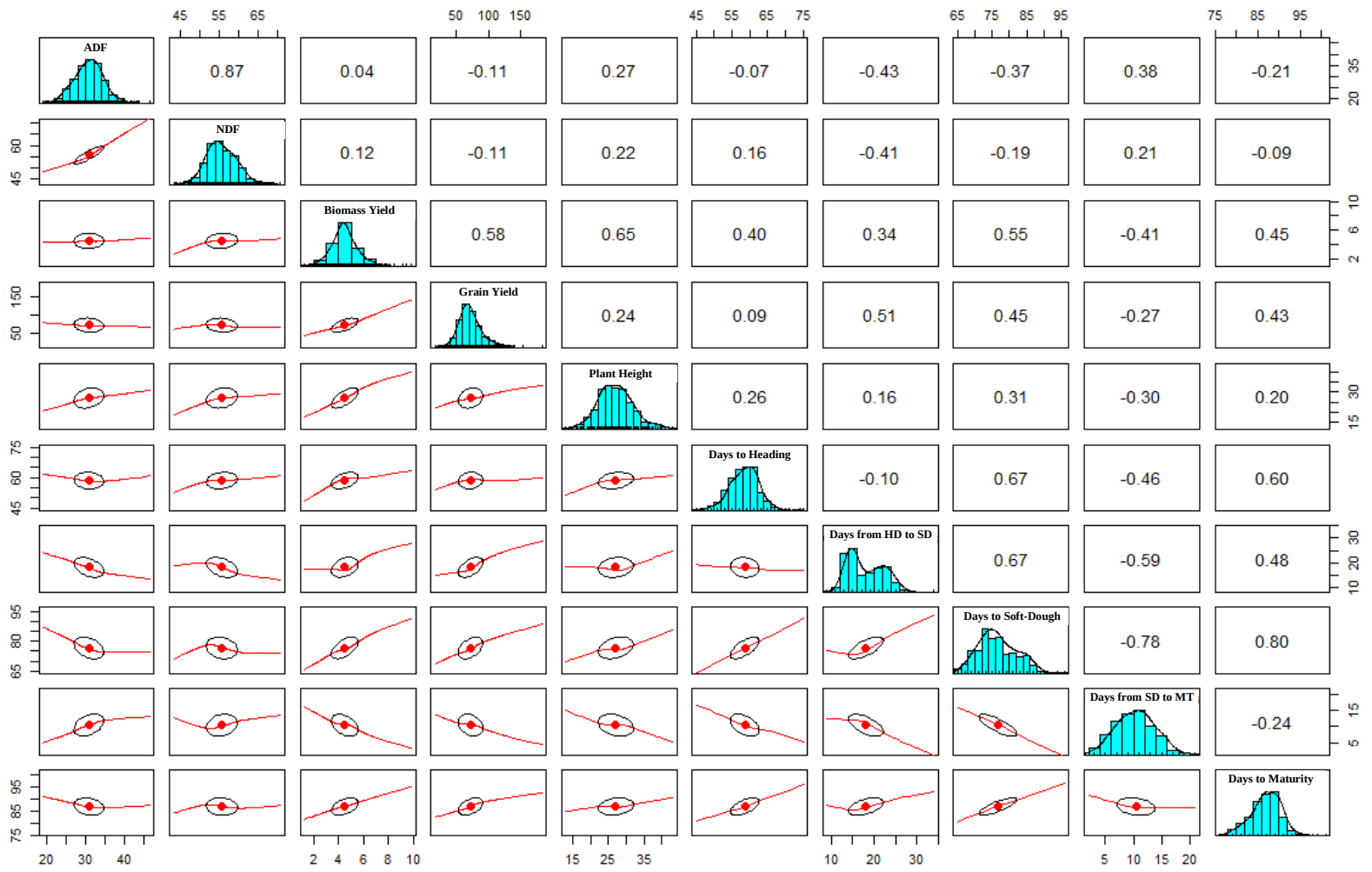
Trait	**Across WC1 treatments[a]**	**WC1 Dryland**	**WC1 Irrigated**	**WC2**
Days to heading	58.7***	59.0	58.3	62.0
Number of days between heading and soft-dough	18.1***	17.7	18.6	17.4
Days to soft-dough	76.8[ns]	76.7	76.9	79.3
Number of days between soft-dough and maturity	10.4***	9.8	11.0	18.3
Days to maturity	87.1***	86.4	87.9	97.7
Acid detergent fiber (%)	30.8***	29.5	32.0	35.9
Neutral detergent fiber (%)	55.7***	54.8	56.6	57.0
Forage yield (ton/ac)	4.5***	4.0	5.0	5.4
Grain yield (bu/ac)	72.1***	66.4	77.8	84.2
Plant height (in)	27.1***	24.8	29.3	33.3

Correlations among traits

Agronomic trait correlations were assessed using accession BLUPS across treatments and years (Figure 3-4). Biomass yield was positively correlated with all agronomic and plant maturity timing traits except for a negative correlation with the number of days between soft-dough and maturity (-0.41). The strongest correlations with

biomass yield were with plant height (0.65), grain yield (0.58), and days to soft-dough (0.55) (Figure 3-4). As expected, the forage quality parameters ADF and NDF were highly correlated (0.87). The number of days between heading and soft-dough was negatively correlated with ADF and NDF (-0.43 and -0.41 respectively), while the number of days between soft-dough and maturity was positively correlated with both quality traits (0.38 and 0.21 respectively) (Figure 3-4).

Figure 3-4: Trait correlations across all WC1 environments.



Identification of Lines with Superior Forage Traits

Extracted BLUP values for digestible biomass yield and grain yield were utilized to rank accessions across all WC1 environments. The ten highest ranking lines came exclusively from WC1 clusters 4 and 5 and were predominantly cultivars from the United Kingdom, New Zealand, and Poland (Table 3-3). A single landrace accession was included in the top ten lines, PI 349896, which was collected from Serbia. The highest ranked line, PI 190200, was a cultivar collected in 1950 from an unknown country in Europe.

Table 3-3: Selection of WC1 lines exhibiting superior forage traits based on highest ranking of BLUPS for digestible biomass yield and grain yield. Top ten ranked accessions compared to standard check varieties. Subpopulation region is based on UN M49 classification.

Overall Rank	Digestible Yield Rank	Grain Yield Rank	Accession	WC1 Subpopulation	Origin, Region	Status	Dry Matter Yield (ton/acre)	%ADF	Digestible Dry Matter Yield (ton/ac)	Grain Yield (bu/ac)
1	1	1	PI190200	4	Unknown, Europe	Cultivar	6.65	32.67	4.28	110.05
2	3	2	PI498433	4	New Zealand, Oceania	Cultivar	5.88	28.81	4.02	98.99
3	5	3	PI498434	4	New Zealand, Oceania	Cultivar	5.78	29.64	3.91	96.75
4	8	25	PI414409	4	Belarus, Europe	Cultivar	5.60	29.56	3.79	85.55
5	9	12	PI467840	5	Poland, Europe	Cultivar	5.51	28.95	3.76	88.80
6	11	16	PI592250	4	United Kingdom, Europe	Cultivar	5.06	25.24	3.65	87.94
7	13	17	PI274620	4	Poland, Europe	Cultivar	5.33	29.08	3.63	87.47
8	17	18	PI349896	5	Serbia, Europe	Landrace	5.39	30.52	3.59	87.38
9	24	15	PI338354	5	Belgium, Europe	Cultivar	5.32	30.68	3.54	87.96
10	28	8	PI485530	4	United Kingdom, Europe	Cultivar	5.16	29.45	3.50	90.90
NA	92	9	Hays	Check	United States, Americas	Cultivar	4.77	30.10	3.20	89.99
NA	172	19	Lavina	Check	United States, Americas	Cultivar	4.42	32.16	2.87	87.20
NA	189	91	Stepford	Check	United States, Americas	Cultivar	4.27	32.30	2.77	75.85
NA	208	140	Conlon	Check	United States, Americas	Cultivar	3.99	31.17	2.63	70.72

Population Structure

Ward's hierarchical clustering based on a Euclidean distance matrix generated from scaled genotypic data was used to group WC1 accessions based on SNP relatedness. The Ward clustering results were plotted with a dendrogram plot and 6 clusters were selected based on the heights of the cluster cuts as described in Everitt and Hothorn (2011, pg 171) (Figure 3-1a). Subpopulations generally clustered by accession status and geographic region of origin. Subpopulation 6 contained only landrace lines and had 54% of all the Asian landrace accessions (Table 3-4). The remainder of the landrace lines were largely split between subpopulation 1 (26%) and subpopulation 2 (28%). 27% and 22% of the Asian lines were in subpopulations 1 and 2 respectively and subpopulation 1 also contained most of the African lines (58%). Subpopulations 3, 4, and 5 were predominantly made up of cultivar and breeding lines from Europe: 65%, 74%, and 74% of each of these populations are European, non-landrace lines (Table 3-4). All the top ten performing accessions identified from WC1 came from subpopulations 4 and 5 (Table 3-3).

A principal component analysis was also utilized to interrogate population structure. The first 2 PCs accounted for 12.8 and 5.2% respectively of the genotypic variation in the WC1 population. Graphing the first 2 PCs and coloring the accessions by their subpopulation gives some additional insight into the relationships between the groups (Figure 3-3a). Subpopulation 6 is a unique cluster, separate from the other groups in the first two PCs. As expected, based on the composition of the groups, subpopulation 6 is closest to subpopulations 1 and 2 in the first PC dimension (100%, 52%, and 55%

landrace lines respectively). Likewise, subpopulations 3, 4, and 5 are also grouped in the first dimension (81%, 97%, and 82% non-landrace lines respectively).

Table 3-4: Breakdown of WC1 subpopulations by geographic region. Number of landraces shown in parenthesis. Subpopulation region is based on country UN M49 classification.

	WC1 Subpopulation						
Region	1	2	3	4	5	6	All
Africa	15 (9)	3 (0)	6 (2)	2 (0)	0	0	26 (11)
Americas	7 (2)	11 (2)	2 (0)	5 (1)	4 (2)	0	29 (7)
Asia	15 (10)	12 (9)	3 (2)	0	0	25 (25)	55 (46)
Europe	13 (5)	22 (16)	33 (2)	43 (1)	23 (3)	2 (2)	136 (29)
Oceania	0	2 (0)	1 (0)	7 (0)	0	0	10 (0)
Unknown	0	1 (1)	3 (3)	0	0	3 (3)	7 (7)
Total	50 (26)	51 (28)	48 (9)	57 (2)	27 (5)	30 (30)	263 (100)

Table 3-5: Mean trait values within each WC1 subpopulation across all treatments. Superscript numbers following averages are subpopulations which are significantly different by Tukey's honestly significant difference at $P < 0.5$

	WC1 Subpopulation						
Trait	1	2	3	4	5	6	All
Days to heading	$56.0^{2,3,4,5}$	$59.0^{1,4,6}$	$58.5^{1,4,5,6}$	$61.9^{1,2,3,5,6}$	$60.3^{1,3,4,6}$	$56.6^{2,3,4,5}$	58.7
Number of days between heading and soft-dough	$17.1^{3,4,5}$	17.6	18.4^{1}	18.2^{1}	18.4^{1}	17.1	17.8
Days to soft-dough	$73.1^{2,3,4,5}$	$76.6^{1,4,5,6}$	$77.0^{1,4,5,6}$	$80.0^{1,2,3,6}$	$78.7^{1,2,3,6}$	$73.7^{2,3,4,5}$	76.5
Number of days between soft-dough and maturity	$11.9^{2,3,4,5}$	10.3^{1}	10.2^{1}	$9.7^{1,6}$	9.8^{1}	11.0^{4}	10.5
Days to maturity	$85.0^{2,3,4,5}$	$86.9^{1,4,5,6}$	$87.1^{1,4,5,6}$	$89.7^{1,2,3,5,6}$	$88.5^{1,2,3,4,6}$	$84.7^{2,3,4,5}$	87.0
ADF (%)	31.0^{4}	31.6^{4}	30.5	$29.6^{1,2,6}$	30.6	31.4^{4}	30.8
NDF (%)	$54.8^{2,5}$	56.4^{1}	55.5	55.6	56.6^{1}	55.3	55.7
Forage yield (ton/ac)	$3.8^{2,3,4,5,6}$	4.6^{1}	4.6^{1}	4.8^{1}	4.8^{1}	4.5^{1}	4.5
Grain yield (bu/ac)	$61.8^{3,4,5}$	$67.8^{4,5}$	$71.1^{1,4,5}$	$78.2^{1,2,3,6}$	$77.9^{1,2,3,6}$	$68.2^{4,5}$	70.8
Plant height (in)	$26.0^{2,3,5}$	$28.7^{1,4,6}$	$29.1^{1,4,6}$	$25.3^{2,3,5}$	$27.9^{1,4}$	$26.3^{2,3}$	27.2

The first PC dimension appears to roughly break the accessions into landrace lines and non-landrace lines (Figure 3-3a). Overall subpopulations 4 and 5 had the highest grain

and forage yields and latest soft-dough date (Table 3-5). Subpopulation 4 also had the lowest ADF values (Table 3-5).

Genome-Wide Association Mapping

Using an FDR-adjusted P-value of 0.05, seventeen major effect QTL were detected in the WC1 population (Table 3-6). Fourteen of these QTL were confirmed with the more limited phenotypic data available on the WC2 population (Table 3-6) and six additional, new QTL were discovered in WC2 for a total of twenty-three QTL (Table 3-6). QTL were discovered associated with each measured trait. Twenty-seven WC1 QTL were found in the dryland environment, 30 in the irrigated, and 27 in the averaged. Detected WC1 loci had adjusted $-\log_{10}$ p-values from 1.1 to 20.0 and R^2 values ranging from 0.0002 to 0.33 (Table 3-6).

Seven WC1 loci were associated with forage or grain yield, seven with forage quality and fifteen with plant developmental stage traits. QTL were detected on all seven chromosomes and the majority of forage trait QTL co-located with QTL related to the timing and progression of plant development and maturity (Table 3-6). This seems to indicate that in a population of global barley accessions, the loci with the greatest impact on forage traits may be those containing genes regulating plant development and senescence.

Table 3-6: Summary of genetic loci significantly (FDR-adjusted P ≤ 0.05) associated with barley traits within and across treatments. Traits ending in "2" indicate QTL found in WC2. Minor allele is bolded and underlined. Bolded forage quality SNPs are further explored for possible pleiotropic effects.

Locus Index	Chr	Range (Mbp)	SNP	Trait	Dryland			Irrigated			Average			MAF	Alleles	Previously reported in
					FDR-adjusted -$\log_{10}(P)$	Effect of minor allele	R^2	FDR-adjusted -$\log_{10}$(P)	Effect of minor allele	R^2	FDR-adjusted -$\log_{10}$(P)	Effect of minor allele	R^2			
1	**1H**	504.1 - 553.7	**SCRI_RS_236623**	ADF	3.38	-0.25	0.0020	x	x	x	x	x	x	0.30	C/T	HEA: (Laurie, Pratchett, Snape, & Bezant, 1995); (Marquez-Cedillo *et al.*, 2001), (Pauli *et al.*, 2014)
			SCRI_RS_143790	ADF2	x	x	x	x	x	x	1.84	-1.60	0.0982	0.46	A/G	
			SCRI_RS_143790	NDF2	x	x	x	x	x	x	2.53	-1.95	0.1230	0.46	A/G	
			SCRI_RS_6824	PHT	3.70	-3.77	0.1237	x	x	x	x	x	x	0.11	C/A	
			12_30532	SOF	x	x	x	2.04	-1.97	0.0523	3.40	-1.99	0.0547	0.45	C/T	
			SCRI_RS_106142	MAT	x	x	x	x	x	x	2.33	-0.81	0.0107	0.16	T/G	
2	**2H**	29.1 - 29.8	**BK_14**	NDF	4.35	-2.65	0.1342	x	x	x	x	x	x	0.17	T/C	NDF (Siahsar *et al.* 2009), HEA: (Munoz-Amatriain *et al.*, 2014), (Gordon *et al.*, 2020), PHT: (Gordon *et al.*, 2020)
			12_30871	BMY	x	x	x	x	x	x	1.45	-1.03	0.3039	0.17	T/C	
			BK_15	BMY2	x	x	x	x	x	x	2.03	-0.86	0.2126	0.19	T/C	
			BK_12	PHT	2.10	-3.21	0.1291	3.82	-5.66	0.2132	1.88	-4.47	0.1911	0.17	T/G	
			12_30872	PHT2	x	x	x	x	x	x	15.74	-4.37	0.1387	0.17	C/T	
			12_30871	HEA	18.70	-6.16	0.3218	12.25	-6.53	0.3407	20.01	-6.32	0.3346	0.17	T/C	
			BK_12	HEA2	x	x	x	x	x	x	33.24	-8.62	0.4821	0.18	T/G	
			BK_12	SOF	8.35	-6.32	0.2949	9.78	-6.45	0.3168	8.68	-6.38	0.3185	0.17	T/G	
			BK_16	SOF2	x	x	x	x	x	x	13.46	-5.55	0.3861	0.19	G/C	
			SCRI_RS_207399	MAT	6.90	-2.35	0.0940	1.77	-2.22	0.1008	2.28	-2.25	0.1000	0.21	T/G	
3	**2H**	197.4 - 226.8	SCRI_RS_192676	SOF	x	x	x	x	x	x	2.30	-5.28	0.3177	0.28	A/G	
			SCRI_RS_192676	SOF2	x	x	x	x	x	x	3.66	2.96	0.3957	0.33	A/G	
			11_10624	MAT	3.93	6.69	0.2211	x	x	x	x	x	x	0.23	C/G	

Locus Index	Chr	Range (Mbp)	SNP	Trait	Dryland			Irrigated			Average			MAF	Alleles	Previously reported in
					FDR-adjusted -$\log_{10}(P)$	Effect of minor allele	R^2	FDR-adjusted -$\log_{10}(P)$	Effect of minor allele	R^2	FDR-adjusted -$\log_{10}(P)$	Effect of minor allele	R^2			
4	**2H**	494.3 - 549.1	11_11430	PHT	x	x	x	x	x	x	1.79	-2.19	0.0698	0.31	T/C	HEA: (Backes *et al.*, 1995); (Laurie *et al.*, 1995); (Marquez-Cedillo *et al.*, 2001); (Mesfin *et al.*, 2003), (Pauli *et al.*, 2014)
			12_20861	HEA	4.83	-6.21	0.1627	x	x	x	3.68	-6.43	0.1726	0.07	A/G	
			11_20585	HEA2	x	x	x	x	x	x	5.21	3.35	0.1653	0.07	A/G	
			11_20374	SOF	x	x	x	1.50	-5.46	0.1847	x	x	x	0.13	G/C	
			11_10436	MAT	x	x	x	2.97	-3.10	0.1610	1.81	-3.01	0.1466	0.16	C/T	
5	**2H**	613.6 - 614.2	SCRI_RS_198848	PHT	x	x	x	1.99	-1.14	0.0133	2.10	-0.68	0.0068	0.32	C/T	
			SCRI_RS_163975	HEA	1.56	-1.37	0.0199	x	x	x	x	x	x	0.23	C/T	
6	**2H**	637.0	SCRI_RS_172667	GYD	x	x	x	4.37	5.92	0.0374	x	x	x	0.42	T/C	
7	**2H**	723.7 - 743.2	**SCRI_RS_159212**	NDF	x	x	x	x	x	x	1.16	0.38	0.0062	0.28	G/A	
			11_10383	PHT	x	x	x	4.25	-5.55	0.1464	5.49	-4.75	0.1536	0.11	A/G	
			SCRI_RS_186925	HEA2	x	x	x	x	x	x	1.74	-1.45	0.0189	0.28	T/C	
			SCRI_RS_114969	MAT	1.90	-0.12	0.0002	x	x	x	x	x	x	0.12	T/C	
8	**3H**	197.4 - 201.2	SCRI_RS_128706	NDF2	x	x	x	x	x	x	5.97	1.73	0.0789	0.28	T/C	
			11_21129	PHT2	x	x	x	x	x	x	2.18	4.36	0.2094	0.32	G/A	
			11_21129	HEA2	x	x	x	x	x	x	4.01	2.16	0.0450	0.32	G/A	
			11_21129	SOF2	x	x	x	x	x	x	4.05	1.23	0.0281	0.32	G/A	
9	**3H**	499.4	SCRI_RS_179275	PHT	x	x	x	1.85	0.95	0.0817	x	x	x	0.45	G/A	
			SCRI_RS_168665	PHT2	x	x	x	x	x	x	2.87	0.60	0.0043	0.37	A/T	
10	**3H**	595.9-634.9	SCRI_RS_172730	HEA	x	x	x	x	x	x	1.45	-4.08	0.0792	0.09	C/G	HEA: (Pan *et al.*, 1994), (Pauli *et al.*, 2014), PHT: (Barua *et al.*, 1993).; (Bezant, Laurie, Pratchett, Chojecki, & Kearsey, 1996); (Pauli *et al.*, 2014) (Laurie *et al.*, 1995)
			12_31238	PHT2	x	x	x	x	x	x	3.74	2.42	0.0731	0.43	G/A	
			SCRI_RS_142438	GYD2	x	x	x	x	x	x	2.25	14.56	0.0579	0.08	C/T	
11	**4H**	13.8	SCRI_RS_134182	GYD2	x	x	x	x	x	x	1.10	16.83	0.0629	0.07	C/T	PHT: (Pillen, Zacharias, & Leon, 2003), (Pauli *et al.*, 2014)
12	**4H**	60 - 103.9	SCRI_RS_169389	PHT	x	x	x	1.85	-3.10	0.0937	2.10	-2.62	0.0961	0.29	T/C	
			11_20853	HEA2	x	x	x	x	x	x	3.15	-1.00	0.0026	0.06	T/G	

Locus Index	Chr	Range (Mbp)	SNP	Trait	Dryland			Irrigated			Average			MAF	Alleles	Previously reported in
					FDR-adjusted -$\log_{10}(P)$	Effect of minor allele	R^2	FDR-adjusted -$\log_{10}(P)$	Effect of minor allele	R^2	FDR-adjusted -$\log_{10}(P)$	Effect of minor allele	R^2			
13	**4H**	580.2 - 633.4	**11_11470**	ADF	1.43	-2.89	0.1068	x	x	x	x	x	x	0.36	G/A	HEA: (Laurie *et al.*, 1995), (Pauli *et al.*, 2014)
			12_30146	NDF	1.27	-3.36	0.0292	x	x	x	x	x	x	0.07	C/G	
			11_20765	PHT	2.33	-2.24	0.0858	8.39	-4.18	0.1577	7.96	-3.21	0.1341	0.25	G/T	
			SCRI_RS_119486	PHT2	x	x	x	x	x	x	3.38	1.05	0.0127	0.35	G/A	
			11_20178	GYD	x	x	x	1.47	-14.24	0.0440	x	x	x	0.05	G/A	
			11_20119	HEA	x	x	x	2.06	-0.71	0.0065	x	x	x	0.35	A/C	
			12_30385	HEA2	x	x	x	x	x	x	1.45	-1.30	0.0098	0.15	A/G	
14	**5H**	409.6 - 442.8	SCRI_RS_186111	BMY2	x	x	x	x	x	x	2.03	0.19	0.0137	0.25	A/G	
			12_11512	PHT2	x	x	x	x	x	x	2.87	-5.87	0.1239	0.08	G/C	
15	**5H**	505.7	11_10183	SOF	x	x	x	x	x	x	3.70	-4.95	0.1722	0.15	C/T	
16	**5H**	573.4 - 597.1	**SCRI_RS_12723**	ADF	x	x	x	4.44	-1.55	0.0764	x	x	x	0.22	T/C	NDF: (Siahsar *et al.* 2009), PHT: (Pauli *et al.*, 2014)
			SCRI_RS_135254	NDF	x	x	x	1.17	-1.48	0.0809	1.16	-1.44	0.1055	0.35	A/G	
			SCRI_RS_8256	PHT	4.26	-2.58	0.1017	1.85	-3.38	0.0919	5.70	-3.01	0.1048	0.22	A/G	
			SCRI_RS_51000	PHT2	x	x	x	x	x	x	2.59	2.34	0.0241	0.10	C/A	
			SCRI_RS_142618	GYD	x	x	x	1.96	5.50	0.0200	x	x	x	0.19	C/G	
			12_20045	HEA	x	x	x	x	x	x	1.93	-2.91	0.0371	0.08	C/G	
			SCRI_RS_201248	HEA2	x	x	x	x	x	x	4.84	-0.33	0.0010	0.29	T/C	
			12_20045	MAT	x	x	x	3.96	-3.06	0.0801	1.54	-2.60	0.0587	0.08	C/G	
17	**5H**	605.4 - 618.2	**12_30400**	ADF	3.38	-1.43	0.0772	2.61	-0.46	0.0088	1.47	-0.96	0.0552	0.35	A/C	
			12_30400	NDF	3.23	-2.11	0.1391	x	x	x	x	x	x	0.35	A/C	
			SCRI_RS_128407	BMY	x	x	x	x	x	x	1.29	-1.25	0.1749	0.06	G/C	
			SCRI_RS_175848	PHT2	x	x	x	x	x	x	2.87	-3.11	0.0370	0.09	G/C	
			SCRI_RS_128407	MAT	x	x	x	2.33	-4.03	0.1135	1.54	-3.89	0.1020	0.06	G/C	
18	**5H**	653	SCRI_RS_142656	GYD	2.74	5.36	0.0120	2.55	-0.51	0.0001	x	x	x	0.05	A/G	PHT: (Spaner *et al.*, 1999); (Tinker *et al.*, 1995), (Pauli *et al.*, 2014)
			SCRI_RS_199298	HEA	1.46	0.31	0.0010	x	x	x	x	x	x	0.23	A/C	
			SCRI_RS_199298	MAT	3.06	0.64	0.0073	x	x	x	x	x	x	0.23	A/C	

Locus Index	Chr	Range (Mbp)	SNP	Trait	Dryland			Irrigated			Average			MAF	Alleles	Previously reported in
					FDR-adjusted -log10(P)	Effect of minor allele	R^2	FDR-adjusted $-\log_{10}(P)$	Effect of minor allele	R^2	FDR-adjusted $-\log_{10}(P)$	Effect of minor allele	R^2			
19	**6H**	5.2 - 16.1	SCRI_RS_237782	HEA2	x	x	x	x	x	x	4.01	2.24	0.0438	0.27	T/C	
			12_30843	MAT	2.77	-0.73	0.0125	x	x	x	x	x	x	0.35	T/A	
20	**6H**	115.8 - 165.7	SCRI_RS_236999	PHT2	x	x	x	x	x	x	2.87	0.78	0.0077	0.48	T/C	
			SCRI_RS_124850	MAT2	x	x	x	x	x	x	6.19	2.21	0.0717	0.20	C/T	
21	**6H**	518.7 - 571.9	**SCRI_RS_144448**	ADF	x	x	x	2.34	-1.32	0.0791	x	x	x	0.42	C/T	
			SCRI_RS_144448	NDF	x	x	x	1.17	-0.72	0.0203	x	x	x	0.42	C/T	
			SCRI_RS_234548	PHT	1.73	2.07	0.0825	x	x	x	x	x	x	0.31	G/A	
			11_20488	GYD2	x	x	x	x	x	x	2.40	-13.79	0.1308	0.23	C/G	
			SCRI_RS_111434	MAT	x	x	x	5.16	1.86	0.1048	x	x	x	0.44	T/G	
			SCRI_RS_184811	MAT2	x	x	x	x	x	x	2.66	-3.09	0.2155	0.43	C/T	
22	**7H**	37.8 - 59.7	SCRI_RS_220780	PHT	x	x	x	1.38	2.08	0.0466	x	x	x	0.35	T/C	HEA: (Backes *et al.*, 1995); (Laurie *et al.*, 1995); (Marquez-Cedillo *et al.*, 2001), (Pauli *et al.*, 2014), PHT: (Bezant *et al.*, 1996), (Pauli *et al.*, 2014)
			12_31305	HEA	1.49	3.07	0.1419	x	x	x	x	x	x	0.46	G/T	
			12_31305	SOF	x	x	x	x	x	x	2.98	3.63	0.1830	0.46	G/T	
			12_20803	MAT	4.34	-3.39	0.1459	3.60	-3.01	0.1378	x	x	x	0.15	A/G	
			11_10772	MAT2	x	x	x	x	x	x	2.39	1.36	0.0420	0.45	G/T	
23	**7H**	400.8 - 505.4	11_11445	PHT2	x	x	x	x	x	x	3.38	1.12	0.0086	0.16	A/T	
			SCRI_RS_133777	MAT2	x	x	x	x	x	x	7.19	0.08	0.0001	0.19	G/A	
24	**7H**	576.3 - 650.0	11_21201	PHT	2.07	2.39	0.1032	1.99	3.07	0.0910	3.87	2.73	0.1027	0.28	A/G	
			12_30974	GYD	x	x	x	2.85	-5.89	0.0137	x	x	x	0.10	A/G	
			SCRI_RS_146157	SOF	1.71	3.05	0.1172	x	x	x	x	x	x	0.42	A/G	
			11_21448	MAT	4.34	2.35	0.1308	x	x	x	x	x	x	0.36	G/A	
			SCRI_RS_152144	MAT2	x	x	x	x	x	x	3.67	0.32	0.0021	0.36	C/T	

HEA is heading date, SOF is soft-dough date, MAT is maturity date, BMY is biomass yield, GYD is grain yield, PHT is plant height, ADF is ADF, NDF is NDF. Traits followed by 2 indicate trait QTL detected in the WC2 population

Seven WC1 SNPs associated with forage quality were tested for possible pleiotropic effects. Lines were grouped by major or minor allele and the impact of the minor allele on the average ADF, NDF, forage yield, grain yield, heading date, soft-dough date, and maturity date was tested using a Type III ANOVA (Table 3-7). Table 3-7 shows that all the selected forage quality SNPs had a highly significant impact on average ADF, average NDF or both. All the selected loci also had highly significant pleiotropic effects on timing of plant maturity and all but one SNP had significant pleiotropic effects on forage and grain yield. However, the direction of the pleiotropic effect varied between the loci. (Table 3-7)

Table 3-7: Pleiotropic effects of forage quality SNPs on agronomic and maturity traits. Trait averages displayed by major (Maj.) and minor (Min.) allele.

LI	SNP	Alleles	ADF (%)		NDF (%)		Forage Yield (ton/ac)		Grain Yield (bu/ac)		Heading Date		Soft-Dough Date		Maturity Date	
			Maj.	Min.	Maj.	Min.	Maj.	Min.	Maj.	Min.	Maj.	Min.	Maj.	Min.	Maj.	Min.
1	SCRI_RS_236623	C/**T**	30.8	30.6^{ns}	56.0	54.9***	4.65	4.14***	71.8	66.4***	59.6	57.0***	77.7	74.4***	87.8	85.5***
2	BK_14	T/**C**	30.9	30.2*	56.0	53.9***	4.67	3.65***	71.8	62.7***	59.9	53.6***	77.8	71.4***	87.6	84.8***
6	SCRI_RS_159212	G/**A**	30.4	31.6***	55.7	55.9^{ns}	4.52	4.44^{ns}	70.7	68.7^{ns}	59.4	57.3***	77.5	74.7***	87.7	85.6***
12	11_11470	G/**A**	31.1	30.1***	55.7	55.6^{ns}	4.38	4.72***	68.9	72.6*	57.9	60.5***	75.7	78.6***	86.5	88.3***
15	SCRI_RS_12723	T/**C**	31.0	29.9***	55.9	54.8***	4.58	4.21***	70.7	68.1^{ns}	59.1	57.9*	76.8	76.2^{ns}	87.2	86.8^{ns}
16	12_30400	A/**C**	31.0	30.1***	56.1	54.8***	4.62	4.26***	72.4	66.3***	59.3	57.9**	77.1	76.0*	87.4	86.6*
20	SCRI_RS_144448	C/**T**	31.1	30.2***	55.8	55.5^{ns}	4.4	4.63**	69.2	71.5^{ns}	57.9	60.1***	75.5	78.3***	86.2	88.3***

LI = Locus Index as described in Table 3-6. Minor allele is bolded and underlined; mean is by allele and across all treatments. Significance of minor allele using a type III ANOVA: *** indicates that the minor allele impacts the WC1 trait at the 0.0001 probability level, ** at the 0.001 level, * at the 0.05 level, ns is not significant.

For loci 1, 2, 15, and 16; the minor allele was associated with higher forage quality, but lower forage and grain yield, and earlier plant maturation. In contrast, when loci 12 and 20 are examined, the minor allele was associated with both higher quality and higher yield traits while plant maturity was delayed. Finally, locus 6 was the only locus were the minor allele was associated with decreased forage quality. This locus had no significant impact on forage or grain yield but was associated with earlier maturity dates. Thus, while there is some evidence of pleiotropic effects for these SNPs, the relationship between forage quality, forage or grain yield, and maturity traits appeared to vary by loci (Table 3-7).

Table 3-8: Percent of each WC1 subpopulation containing the SNP minor allele and significance of interaction between allele x subpopulation for major forage and agronomic traits. Minor allele is bolded and underlined.

Locus	SNP	Alleles	% of WC1 Subpopulation with minor allele						ADF (%)	NDF (%)	Forage Yield (ton/ac)	Grain Yield (bu/ac)	Heading Date	Soft-Dough Date	Maturity Date
			1	2	3	4	5	6							
1	SCRI_RS_236623	C/**T**	59	4	8	19	7	100	***	ns	ns	ns	*	**	***
2	BK_14	T/**C**	52	17	10	7	4	0	ns	ns	ns	ns	ns	ns	ns
6	SCRI_RS_159212	G/**A**	25	40	14	4	0	100	ns	**	***	***	***	***	***
12	11_11470	G/**A**	12	25	63	50	56	0	*	ns	ns	ns	ns	*	**
15	SCRI_RS_12723	T/**C**	45	13	16	35	0	93	*	ns	ns	ns	ns	*	*
16	12_30400	A/**C**	67	44	40	28	0	0	*	ns	***	***	***	***	***
20	SCRI_RS_144448	C/**T**	14	44	55	81	33	0	*	ns	ns	***	ns	ns	ns

Significance of minor allele using a type III ANOVA: *** indicates that the minor allele impacts the WC1 trait at the 0.0001 probability level, ** at the 0.001 level, * at the 0.05 level, ns is not significant.

While pleiotropic effects are important, it is also crucial that the impact of subpopulation is explored. Using a Type III ANOVA test, six of the seven forage quality SNPs exhibited evidence of a marker by subpopulation interaction on either ADF or NDF (Table 3-8). Three SNPs showed evidence that either forage or grain yield was impacted by a marker x subpopulation interaction, while five SNPs exhibited the same for at least one plant maturity trait (Table 3-8). The impact of the interaction between subpopulation and SCRI_RS_12723 (Locus 15) on soft-dough date (SOF) can be visualized in Figure 3-5a. While the "TT" allele was typically related to later soft-dough date, this relationship was inversed in WC1 subpopulation 2.

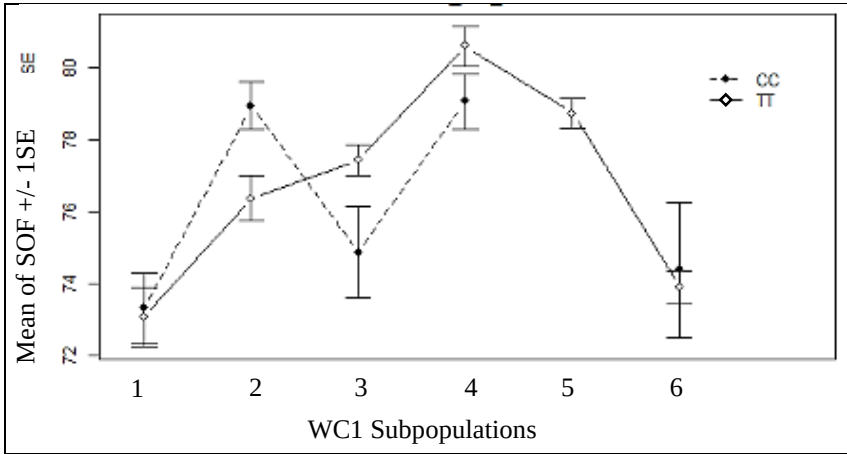


Figure 3-5a: The effect of the minor allele of SCRI_RS_12723 is dependent on subpopulation.

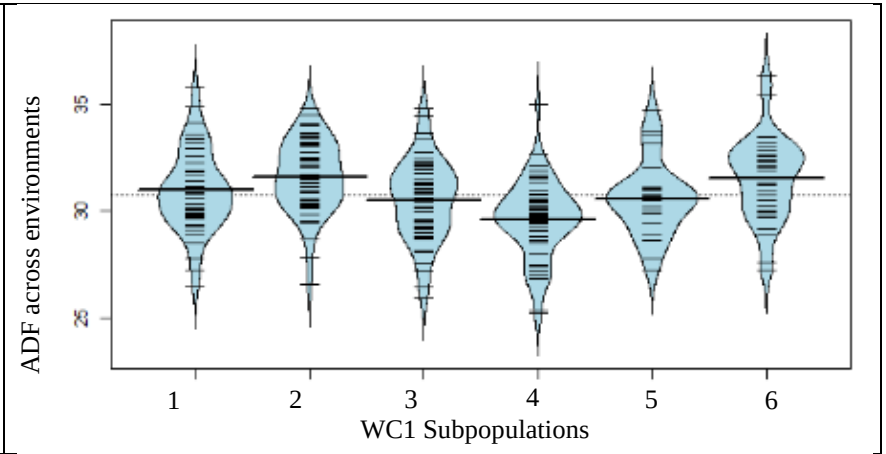


Figure 3-5b: All WC1 subpopulations exhibited equally large variation in ADF.

Additionally, it is important to note that the prevalence of the SNP minor allele for each locus varies between the subpopulations. From Table 3-8 we can see that the minor alleles were not spread equally among the subpopulations. Indeed, while subpopulation 6 exhibited a large variation for ADF (Figure 3-5b), the subpopulation was functionally monomorphic for all seven SNPs identified as impacting forage quality (Table 3-8). This is an indication that the current GWAS failed to identify the QTL impacting forage quality in subpopulation 6 and further work is needed to better understand the accessions which make up this subpopulation.

Discussion

Germplasm and Breeding

Germplasm repositories represent an invaluable genetic resource for plant breeders and researchers. To better understand the loci related to forage yield and quality in barley, and to identify barley accessions with superior forage traits, 260 spring, two-row barley lines were selected from the iCore population described by Munoz-Amatriain *et al.* (2014). The selected lines (WC1) were grown in an augmented field design under irrigated and dryland conditions across two field seasons and data were generated on each accession's forage and agronomic traits. From each accession, BLUP trait values were extracted within and across treatments. In the 260 lines within the WC1, 35% of the lines had higher biomass yield than the standard forage barley checks and ten superior accessions were selected based on digestible yield and grain yield (Table 3-3). Selecting for high grain yield as well as high forage yield ensures that released cultivars are compatible with forage seed production and facilitates cultivar market adoption.

The majority of the lines in this study with superior forage traits were cultivars from Europe and New Zealand. The highest-ranking line, PI 190200, out yielded the standard forage cultivars by more than a ton per acre and was a cultivar collected from an unknown country in Europe. The only landrace line amongst the highest-ranking accessions (Table 3-3) was PI 349896, a line collected in Serbia. The lines identified in this trial may provide a germplasm resource for breeders and researchers interested in forage yield and quality traits. From a breeder's perspective these lines do have two potential drawbacks: the identified lines are awned and, according to GRIN, have a lack of disease resistance.

The highest ranked line, PI 190200 (Table 3-3), was a cultivar collected in 1950 from an unknown country in Europe. While this line was identified as having positive forage traits in this trial, for breeders interested in using this line it is important to note that it is identified in the GRIN database as being resistant to scald but susceptible to net blotch, spot blotch, and stripe rust. Researchers with high disease pressure will need to select parents and offspring appropriately. Lines PI 498433, PI 498434, PI 414409, PI 467840, PI 485530 also provide net blotch resistance, however line PI 414409 was the only line in the top ten with even a moderate level of resistance to stripe rust and may be the best option for regions prone to these foliar diseases. While several of the best performing lines have some disease resistance to net blotch, resistance to other foliar diseases such as spot blotch and stripe rust was very limited. Forage barley is harvested much earlier in the season than grain barley, and thus avoids later season foliar diseases. However, environments with high disease pressure may still require some resistance. In

addition to considering disease pressure, it is important to note that the lines included in the WC1 are predominantly awned and, depending on the cultural practices of the area, may require introgression of desirable forage traits into hooded or awnless varieties.

Genome-Wide Association Mapping

Impact of environment and population: Across both the WC1 and WC2 populations, 118 marker-trait associations within twenty-three QTL were identified. Within those twenty-three loci, twenty-seven WC1-dryland, thirty WC1-irrigated, twenty-seven WC1-average, and thirty-four WC2 SNP-trait associations were detected. (Table 3-6)

Of the 84 WC1 marker-trait associations found across the three WC1 environments, nineteen (~23%) were found in more than one environment: 7 for plant height, 9 for heading/soft-dough/maturity, 2 for forage quality, and 1 for grain yield (Table 3-6). In all but one of these associations, the minor allele had the same directional effect; only the GYD SNP on 5H (Locus index #17, Table 3-6) exhibited a shift in the direction of the effect of the minor allele, with a positive effect in the dryland environment but a negative effect under the irrigated treatment. With the exception of this association, the size and direction of the impact of the minor alleles was exceedingly consistent between environments when an association was detected in multiple environments. (Table 3-6)

For the 65 WC1 marker-trait associations found in a single environment, those detected in the dryland and average conditions are of the most interest to Montana forage barley breeders. Since most of the cropland in Montana is under dryland farming

practices, the positive associations detected under dryland conditions should be positive for the largest number of growers. While associations detected across the treatments could be expected to be the most environmentally stable and would be expected to have consistent impacts in future environments regardless of whether the crop was grown under dryland or irrigated conditions.

Forage quality and plant maturity: Eight novel forage quality QTL were discovered. In the world core barley population, with its global genetic diversity, flowering time and plant maturity genes appeared to be important genetic influencers on forage quality and yield. Of the 8 genetic loci detected in connection to ADF or NDF across WC1 and WC2, all loci also included connections to heading date, soft-dough date or maturity date. Three of these loci (1, 2, 12, Table 3-6) have previously been linked to heading date in barley, one locus (15, Table 3-6) has only been previously linked to plant height, while the remaining four (6, 7, 16, 20, Table 3-6) are novel for all traits. Although the genetics underlying forage traits in barley are not well understood, the close relationship between plant maturity and forage quality found in this study has previously been shown in a genome-wise association study in perennial wildrye (Larson *et al.*, 2014).

Two loci related to forage quality contained known barley plant maturity genes. Locus 1 lies near a region on 1H containing the photoperiod response gene *Ppd-H2* (Laurie *et al.*, 1995), (Pauli *et al.*, 2014), while locus 2 spans a region of chromosome 2H that has previously been associated with heading date by Munoz-Amatriain *et al.* (2014) and Gordon *et al.* (2020) and which contains the long-day flowering time photoperiod

response regulator *Ppd-H1* (Turner, 2005). Across treatments the minor allele SNPs for these forage quality QTL were associated with earlier heading, soft-dough, and maturity dates; as well as improved forage quality and decreased biomass and grain yield (Table 3-7). The negative impact on biomass and grain yield may limit the usefulness of loci 1 and 2 for barley forage breeders, however the importance of maturity genes on forage traits is clear.

Of the remaining loci related to forage quality, the regions of loci 12 and 15 had previously been associated with barley heading date (Laurie *et al.*, 1995; Pauli *et al.*, 2014) and plant height respectively (Pauli *et al.*, 2014). Loci 6, 7, 16, and 20 had not been previously documented. When the pleiotropic effects of each loci were estimated (Table 3-7), loci 15 and 16 behaved similarly to loci 1 and 2. The minor alleles of the forage quality SNPs correlated to improved forage quality but reduced grain and biomass yield. Loci 6, 12, and 20 appear more promising for barley forage breeders. The minor forage quality alleles associated with these loci had positive impacts on forage quality while being positive or neutral for biomass and grain yield. Depending on the locus, the impact of plant maturity on forage yield or quality varied and additional investigation into the relationship between these complex traits would be beneficial.

Three additional loci (4, 9, and 21) also contained known genes. Locus 9, detected only in WC2, contains the *denso* gene (Wang *et al.*, 2012), which has been linked to both flowering time and plant height (Barua *et al.*, 1993; Bezant *et al.*, 1996; Laurie *et al.*, 1995; Pan *et al.*, 1994, Pauli *et al.*, 2014). Munoz-Amatriain *et al.* (2014) and Pauli *et al.* (2014) also previously identified a heading date QTL in the 519-525 Mbp region on

chromosome 2H, which corresponded to locus 4 of this study. Munoz-Amatriain *et al.* (2014) concluded that this QTL was tightly linked with an earliness per se flowering determinant mutation, *EPS2* (Comadran *et al*., 2012). Finally, locus 21 contains the *Vrn-H3* gene, a vernalization gene which is a homolog to the Flowering Locus T in Arabidopsis (Yan *et al.*, 2006).

Conclusion

Forage barley provides a valuable option for growers looking for a drought and salt tolerant crop that can produce in a single growing season. For plant breeders developing forage barley cultivars, improving forage quality as well as forage yield must be the goal. Genotyped global barley accessions from the iCore offer novel germplasm and the ability to discover the genetic regions underlying desirable traits. This study served to uncover ten global accessions with superior forage traits. A genome-wide association study identified twenty-three loci associated with key agronomic traits within and between irrigated and dryland environments. Eight loci across chromosomes 1-6 were identified associated with forage quality traits. All eight forage quality loci also included associations for plant maturity traits such as heading, soft-dough, and maturity date. Estimates of pleiotropic effects showed that the relationship between forage quality, forage yield, and maturity dates varied by loci and additional investigation was warranted. Several loci were identified which improved both forage quality and yield and such loci provide breeders with a potential resource for improvement of future barley cultivars.

CHAPTER FOUR: MODELING THE IMPACT OF AGRONOMIC TRAITS ON FORAGE YIELD AND QUALITY

Introduction

A forage is a crop grown for the vegetative biomass it produces as a source of animal roughage. For cattle and sheep producers in Montana, forages are foundational to their business model and the largest cost to most operations (USDA National Agricultural Statistics Service, 2020).

Annual cereal forage varieties are important to cattle and sheep production, especially in regions of the US with low rainfall and short growing seasons such as the Northern Great Plains and Rocky Mountain regions. Annual cereal forages provide rapid biomass production and have been found to out-yield established perennial forages under drought conditions in Montana (Cash *et al.*, 2006). Of the common annual cereal forages, barley has been found to have an excellent balance between drought tolerance, medium to high yield, and good forage quality. It even exhibits resistance to saline soils (Maas, 1993; Steppuhn & Raney, 2005), another concern in the region. Taken together, barley's characteristics indicate that it would be an ideal target for forage breeders trying to fill the increased need of Montana growers for an adaptable and sustainable roughage crop.

Although forage yield is very important, forage quality is also a significant factor when breeding for a forage cultivar. Forage quality is a broad term that encapsulates the many characteristics of a forage that determine how well a forage meets the nutritional requirements of the animal consuming it. When considering barley forage quality to meet

the nutritional needs of cattle in Montana, two related traits are of particular interest to breeders: intake and digestibility.

Intake is a measure of how much of a given forage an animal can consume before experiencing gut-fill and appetite suppression. This factor is primarily related to the concentration of plant biochemicals which are difficult (hemicellulose, cellulose) or essentially impossible (lignin) for even a ruminant animal to digest. If a forage is of poor quality and thus low intake, the animal may not be able to ingest a sufficient quantity of feed to meet its nutritional requirements. Forage intake values can be approximated with a laboratory-based test, Neutral Detergent Fiber (NDF). NDF estimates the percent of a forage sample that consists of hemicellulose, cellulose, and lignin.

Digestibility is a measure of how easily a forage is digested by an animal. Forage digestibility is primarily related to the percent of indigestible material, mostly lignin, present in a forage; it determines the nutritional value of the crop and can affect feeding efficiency. Like intake, the digestibility of a forage can be approximated with a laboratory-based test. Acid Detergent Fiber (ADF), estimates the percent of a forage sample that consists of lignin.

Even small improvements in forage quality can have a positive impact on growers. A conservative estimate is that a one percent increase in forage digestibility can lead to a three percent increase in the average daily gains of steers, a three percent increase in milk production for dairy cows, and an overall increase in animal production per unit area (Casler & Vogel, 1999; Mohammed, 1967).

While forage quality is important to both producers and plant breeders, breeding for forage quality is challenging. Forage quality is expensive and time-consuming to measure. In the early stages of a breeding pipeline a breeder must be able to evaluate tens of thousands of lines in the field and measuring forage quality, even using laboratory techniques, is not feasible. This means that selection for forage lines at the early stages of the breeding program is largely subjective, based on the physical appearance of the line rather than robust data on yield or quality. The lack of quantitative selection criteria at these crucial stages must compromise the breeder's ability to select the optimal lines to advance.

Unlike digestibility or biomass yield, flowering time and plant height are relatively easy to select for in early stages of the breeding program. If the relationship between these traits and forage yield or quality could be quantified in barley, they could be used to inform early generation selection. Such models have been developed in alfalfa. In pure stands of alfalfa, proxy-based modeling has been successful in predicting first cutting forage quality. Models for alfalfa have used growing degree-days, plant height, and stage of maturity to predict forage quality (Fick & Onstad, 1988; Hintz & Albrecht, 1991). These models were developed as tools to help growers know how to time the harvest of their crop, but the success of the models suggests a connection between the terms of the model and the quality traits that could also be leveraged by forage breeders.

The most successful predictive models for alfalfa are the Predictive Equations for Alfalfa Quality (PEAQ) models. The first of these models was developed in Wisconsin in 1991 by Hintz and Albrecht. They found that including two easily measured

morphological traits: stem height and stem maturity, produced multiple regression models that had R^2 values of 0.88 and 0.89 for ADF and NDF respectively.

The findings of Hintz and Albrecht have been validated for pure stands of alfalfa in additional environments in Wisconsin (Owens *et al.*, 1995; Sulc *et al.*, 1997), New York, Pennsylvania, Ohio, California (Sulc *et al.*, 1997), New Mexico, Northern Mexico (Santillano-Cázares *et al.*, 2014); Central Europe (Hakl *et al*., 2019); and Northern Europe (Andrzejewska *et al.*, 2014).

While PEAQ models were originally developed in pure alfalfa stands, more recently researchers have also shown that robust PEAQ models can also be developed for mixed alfalfa-grass stands (Parsons *et al.*, 2006; Parsons *et al*., 2013; Wood *et al.*, 2018; Wood *et al*., 2019). However, in alfalfa-grass mixtures, growing degree days and the grass contribution to total yield were also required in the PEAQ model.

The success of these predictive models suggests a connection between the terms of the model and the quality traits being measured. This connection could be leveraged by the forage breeder. If terms such as stem height and maturity stage are predictive of forage quality, these or similar terms could be promising as proxies for selection in a forage breeding program. To better understand the relationship between barley forage digestibility, yield, flowering time and plant height, a modeling study was conducted.

Materials and Methods

WC1 Population selection

The USDA-ARS National Plant Germplasm System (NSGC) is a global germplasm repository which contains more than more than 33,000 *Hordeum vulgare L.* accessions collected from more than 100 countries and categorized based on their level of genetic improvement. As part of the USDA-NIFA funded Triticeae Coordinated Agricultural Project (TCAP), Muñoz-Amatriaín *et al.* selected 1,860 unique barley accessions from the NSGC to represent the total diversity of the collection and genotyped them with a barley SNP iSelect platform with 7,842 markers (Munoz-Amatriain *et al.*, 2014). This population was designated the USDA Barley World informative Core Collection (iCore). Of the 1,860 lines in the iCore, 621 lines were of the spring, 2-row morphology, which is the form of barley most commonly grown in Montana.

The iCore was designed to capture the full genetic diversity present in the NPGS. From the 621 spring, 2-row lines in the iCore, a population of 260 genetically diverse lines was randomly selected (WC1).

WC1 Experimental Design and Data Collection

The selected 260 iCore lines, termed World Core population 1 (WC1), were grown under both irrigated and rainfed (dryland) treatment conditions in field trials at the Arthur H. Post Research Farm in Bozeman, MT (Latitude: 45.6729, Longitude: -111.1547, Elevation: 1455 m) on Amsterdam silt loam soil in both the 2016 and 2017 field seasons. Field plots in 2016 were planted as two rows plots, 8 ft in length and

planted on 9 ft centers. Plots in 2017 were planted as three row plots, 15 ft in length and planted on 18 ft centers. All plots were seeded at a rate of 1.1 g/ft^2 with a row spacing of 1 ft. The irrigated treatment received 5 inches of irrigation which was applied to the field in two, 2.5 inch increments immediately prior to the heading stage of development. Each WC1 field was planted in an augmented design with 10 blocks and 4 check plots per block. The check varieties, Lavina, Hays, Conlon, Stepford, were randomly assigned plots within each block and the 260 experimental lines were randomly assigned to the remaining plots. This design resulted in 300 plots per treatment, 260 plots for the genetically diverse lines and 40 plots for the checks, for a total of 600 plots per year and 1200 plots across both years.

Plots were monitored daily throughout the growing season and data was collected on plot heading date (Zadoks stage 51), soft-dough date (Zadoks stage 85), maturity date (Zadoks stage 92), and mature plant height. Zadok stages were defined as when 50% of the main tillers in a plot reach the designated stage. The Zadok stage 92 definition was modified to be assessed visually. Visually, maturity was defined as when no green color remained in 50% of the main tiller heads, including the awns and glumes. Plant height was measured on mature plants and was defined as the distance from ground level to the top of the head, excluding awns. Plant height was measured at two locations per plot by selecting a handful of stems from the middle of the plot and taking the height of the tallest stem excluding awns. The two measurements were then averaged.

When a plot reached the soft-dough stage of maturity, six 6-inch forage samples were collected from representative areas of the plot and bulked. Samples were cut one

inch from the soil surface. Collected forage samples were dried, weighed, milled, and analyzed. From each forage sample, samples were analyzed for ADF and NDF and an estimated dry ton/acre biomass yield was calculated. It is important to note that although digestibility is the forage quality variable of interest, the laboratory proxies ADF and NDF measure % non-digestibility – thus, as ADF or NDF values decrease, digestibility and intake respectively increase.

WC2 and BP Conformation Populations

Population development: To confirm the modeling findings from the WC1 2016-2017 field seasons, two additional populations were studied in the 2018 and 2019 field seasons. First, a new set of 260 genetically diverse spring 2-row lines was randomly selected from the remaining untested spring, 2-row lines in the iCore (WC2). Second, a set of 79 forage barley cultivars and lines from the advanced stages of the Montana State University Forage Barley Breeding Program (BP) were also screened.

The BP lines constitute the current culmination of forage breeding in the MSU barley program. Like many breeding programs, the lines in the most advanced stages of the selection process are largely the result of crosses including elite Montana cultivars. For example, Hays and Lavina are the two most recent barley forage lines to have been produced by the MSU forage barley breeding program. Hays and Lavina are sister lines, the result of a cross between the forage line Haybet and the feed variety Baronesse. Of the 79 lines in the BP, only 4 lines do not include Hays, Lavina, or Haybet as one of their parents. Consequently, the BP population represents a much less genetically diverse

population than either of the WC populations, but it is the population which is most like those that barley forage breeders might encounter during the breeding process.

Experimental design and data collection: The WC2 and the BP conformation populations were grown under dryland conditions in 2018 and 2019 at the Arthur H. Post farm on Amsterdam silt loam soil. All confirmation trials were monitored daily, and data were collected as previously described. Field plots in 2018 and 2019 were planted as three row plots, 15 ft in length and planted on 18 ft centers. All plots were seeded at a rate of 1.1 g/ft^2 with a row spacing of 1 ft.

The WC2 population was grown in both 2018 and 2019 in an augmented design as described for the WC1 populations. The BP trials were grown in two different experimental designs depending on year and stage of advancement in the breeding program. Preliminary Yield Trials (PYT) were grown in an augmented design. In 2018, the PYT had three blocks with two check varieties, Hays and Lavina, per block and 16 experimental entries: 22 plots total. The 2019 PYT had three blocks with two check varieties, Hays and Lavina, per block and 42 experimental entries: 48 plots total. Advanced Yield Trials (AYT) were grown in Randomized Complete Block design. Both the 2018 and 2019 AYTs had three blocks with two check varieties, Hays and Lavina, per block and 14 experimental entries: 48 plots total per year.

In general, experimental lines are progressed through the stages of the breeding program; with superior lines progressed from preliminary to advanced trials while inferior lines are eliminated. Thus, due to the nature of a breeding program, the tested experimental lines varied from year to year and the BP dataset was not balanced.

Forage Sample Preparation and Laboratory Analysis

All forage samples were dried in a forced-air oven at 40 °C for 96 hours, weighed, and then ground using a Wiley cutting mill to pass through a 2-mm screen. Two hundred fifty forage barley samples from the 2016 field season (WC1) were analyzed in technical triplicate for acid detergent fiber (ADF) and neutral detergent fiber (NDF) using an ANKOM 2000 Fiber Analyzer as per the manufacturer's instructions. The ADF on these samples was determined via Method 973.18 (Mertens, 1992). The NDF was determined using heat stable alpha-amylase and sodium sulfite via the methods described by Van Soest, Robertson, and Lewis (1991).

From these samples, a custom NIR calibration curve was developed using a Foss NIRSystems 6500 with proprietary WinISI software. Calibration equations were validated on 20 additional 2016 WC1 samples not included in the sample subset used in calibration creation. All subsequent forage samples were analyzed for ADF and NDF using this custom NIR calibration via Method 4.2 of the National Forage Testing Association ("Fiber (Acid Detergent) and Protein (Crude) in Animal Feed and Forages: Near-infrared Reflectance Spectroscopic Method (989.03)," 1990; Martin, Shenk, & Barton, 1989). This calibration was updated as needed with additional samples in order to ensure that it remained robust from year-to-year.

Year-to-Year Variation in Environmental Conditions

The 2016 and 2017 growing seasons had similar growing conditions, with the exception of higher temperatures and lower precipitation in July 2017 relative to July

2016 (Table 4-1). The 2017 July conditions, on average, caused a 7-day decrease in soft-dough date (Table 4-2).

Table 4-1: Summary of field season precipitation and temperatures at Arthur H. Post Farm, Bozeman, MT from 2016 to 2019. LTA = Long term average.

Year	May				June				July				August				Total Season Precip (in) (May-Aug)
	Ave high temp (F°)	Ave low temp (F°)	Total precip (in)	# Days above 90 (F°)	Ave high temp (F°)	Ave low temp (F°)	Total precip (in)	# Days above 90 (F°)	Ave high temp (F°)	Ave low temp (F°)	Total precip (in)	# Days above 90 (F°)	Ave high temp (F°)	Ave low temp (F°)	Total precip (in)	# Days above 90 (F°)	
LTA	66	38	2.9	0	74	45	2.8	1	84	51	1.4	8	84	49	1.3	7	8.4
2016	64	39	2.7	0	80	48	0.8	3	83	51	1.2	6	84	49	0.9	7	5.6
2017	67	39	2.6	0	76	46	2.2	2	88	55	0.1	15	83	51	0.6	1	5.5
2018	68	44	2.9	0	72	47	3.6	0	82	51	0.2	3	80	49	1.3	5	8.0
2019	60	39	1.7	0	73	45	2.1	0	80	50	2.7	1	82	50	0.7	0	7.2

This year-to-year variation indicated the need for including Year as a random effect in the WC1 models. For the WC2 and BP trials, the trial variable was unique to a given year and the inclusion of trial as a random effect in these models was sufficient to account to year-to-year variation.

Summary of Agronomic Data

As expected, the global populations behaved more similarly to each other than to the breeding program populations (Table 4-2). The WC populations are much more genetically diverse and exhibited a greater range in values than the BP population for all traits measured (Table 4-2). For example, the ADF of the WC1 and WC2 populations had ranges of 18-22 % depending on year and environment, while the BP population had a range of only 11-14% (Table 4-2).

On average, the BP population had higher grain yield, and biomass yield while the WC populations had lower % ADF and NDF values. Higher yields in the elite lines found in the BP was expected since these lines have been selected for these traits, however the superior average forage quality of the WC populations was surprising. (Table 4-2)

While the average number of days to heading was relatively consistent across populations and years, the number of days to soft-dough was more likely to be impacted by year to year variations in temperature and precipitation. These changes can more easily be seen in the data when the averages of the variable “days between heading and soft-dough” (“head_soft_dif”) are compared across years. Within a population, the average “head_soft_dif” varied by up to 7 days from year to year. (Table 4-2)

Table 4-2: Summary of variables by experimental population and environment

Trait	Year/Trial	N	Min.	Max.	Mean	St. Dev.	Range
Days to heading							
	2016 Dryland WC1	287	45	75	58.9	4.3	30.0
	2016 Irrigated WC1	285	46	74	58.3	4.2	28.0
	2017 Dryland WC1	300	49	68	59.1	3.7	19.0
	2017 Irrigated WC1	300	47	67	58.2	4.0	20.0
	2018 WC2	300	49	76	65.6	5.5	27.0
	2019 WC2	299	48	66	58.3	4.0	18.0
	2018 BP	69	55	61	58.0	1.8	6.0
	2019 BP	95	54.5	62.5	57.9	1.9	8.0
Number of days between heading and soft-dough							
	2016 Dryland WC1	287	14	34	20.6	2.9	20.0
	2016 Irrigated WC1	285	16	28	22.4	2.5	12.0
	2017 Dryland WC1	300	9	21	14.7	1.6	12.0
	2017 Irrigated WC1	300	11	25	14.9	2.1	14.0
	2018 WC2	300	11	27	15.6	2.4	16.0
	2019 WC2	299	14	28	19.2	2.2	14.0
	2018 BP	69	18.5	26.5	21.5	1.8	8.0
	2019 BP	95	20	28	23.3	1.7	8.0
Days to soft-dough							
	2016 Dryland WC1	287	65	96	79.5	5.7	31.0
	2016 Irrigated WC1	285	66	94	80.6	5.0	28.0
	2017 Dryland WC1	300	66	83	73.8	2.9	17.0
	2017 Irrigated WC1	300	65	81	73.1	3.5	16.0
	2018 WC2	300	70	89	81.1	4.1	19.0
	2019 WC2	299	70	84	77.5	3.0	14.0
	2018 BP	69	78	84	79.4	1.7	6.0
	2019 BP	95	81	82.5	81.3	0.6	1.5
Number of days between soft-dough and maturity							
	2016 Dryland WC1	287	2	21	8.9	3.7	19.0
	2016 Irrigated WC1	285	2	20	8.6	3.1	18.0
	2017 Dryland WC1	300	6	19	10.7	2.1	13.0
	2017 Irrigated WC1	300	8	20	13.4	2.1	12.0
	2018 WC2	300	12	31	18.7	3.1	19.0
	2019 WC2	299	10	27	18.0	2.7	17.0
	2018 BP	69	5	12.5	9.7	1.7	7.5
	2019 BP	95	11	19.5	15.1	2.0	8.5
Days to maturity							
	2016 Dryland WC1	287	78	101	88.4	3.2	23.0

	2016 Irrigated WC1	285	78	99	89.2	3.1	21.0
	2017 Dryland WC1	300	76	93	84.4	3.2	17.0
	2017 Irrigated WC1	300	78	92	86.5	2.7	14.0
	2018 WC2	300	88	109	99.8	3.6	21.0
	2019 WC2	299	85	105	95.5	3.5	20.0
	2018 BP	69	87	92	89.2	1.4	5.0
	2019 BP	95	92	100.5	96.4	2.1	8.5
Acid detergent fiber (%)							
	2016 Dryland WC1	287	19.1	37.2	27.7	3.3	18.1
	2016 Irrigated WC1	285	20.0	38.6	30.2	3.1	18.6
	2017 Dryland WC1	300	20.7	43.4	31.3	3.0	22.7
	2017 Irrigated WC1	300	24.1	46.4	33.8	3.1	22.3
	2018 WC2	300	21.3	41.2	34.5	2.9	19.9
	2019 WC2	299	26.5	46.5	37.3	3.2	20.1
	2018 BP	69	31.8	42.8	38.0	2.3	11.0
	2019 BP	95	27.5	41.9	35.7	3.2	14.4
Neutral detergent fiber (%)							
	2016 Dryland WC1	287	43.4	62.2	52.9	3.1	18.9
	2016 Irrigated WC1	285	47.0	62.6	54.1	2.8	15.6
	2017 Dryland WC1	300	47.9	67.8	56.6	3.4	19.9
	2017 Irrigated WC1	300	49.8	70.8	59.1	3.6	20.9
	2018 WC2	300	43.9	63.1	54.8	3.3	19.2
	2019 WC2	299	46.2	67.5	59.2	3.5	21.3
	2018 BP	69	55.1	65.1	60.8	2.3	10.0
	2019 BP	95	51.3	64.9	58.8	3.2	13.6
Forage yield (ton/ac)							
	2016 Dryland WC1	287	1.3	8.5	3.8	0.9	7.2
	2016 Irrigated WC1	285	2.1	9.8	5.5	1.2	7.7
	2017 Dryland WC1	300	2.5	6.5	4.2	0.6	4.0
	2017 Irrigated WC1	300	1.9	6.4	4.4	0.8	4.4
	2018 WC2	300	1.9	8.4	5.0	1.0	6.5
	2019 WC2	299	3.8	8.1	5.7	0.7	4.3
	2018 BP	69	3.2	6.8	4.8	0.9	3.7
	2019 BP	95	4.9	8.2	6.6	0.9	3.4
Grain yield (bu/ac)							
	2016 Dryland WC1	287	24.0	105.8	70.0	14.5	81.8
	2016 Irrigated WC1	285	43.9	182.9	91.7	20.5	139.0
	2017 Dryland WC1	300	28.5	93.7	62.9	11.0	65.2
	2017 Irrigated WC1	300	17.1	95.6	64.3	13.5	78.5
	2018 WC2	300	35.6	131.6	78.4	16.3	95.9
	2019 WC2	299	37.2	152.8	90.0	22.5	115.5

Plant height (in)	2018 BP	69	74.6	147.6	112.5	16.0	73.0
	2019 BP	95	89.7	148.7	119.3	13.0	58.9
	2016 Dryland WC1	287	13.0	33.0	23.3	3.4	20.0
	2016 Irrigated WC1	285	15.0	43.0	31.3	5.4	28.0
	2017 Dryland WC1	300	17.3	35.8	26.3	3.3	18.5
	2017 Irrigated WC1	300	14.6	39.4	27.4	3.8	24.8
	2018 WC2	300	14.2	46.1	32.9	5.1	31.9
	2019 WC2	299	18.5	44.1	33.6	3.8	25.6
	2018 BP	69	29.1	41.3	34.2	3.0	12.2
	2019 BP	95	32.1	40.5	35.4	1.9	8.4

Data Analysis and Modeling

All data processing and statistical modeling was performed in the open-source statistical platform R (R Core Team, 2020) through the graphical user interface RStudio (RStudio Team, 2020). All models were fit as linear mixed models with the lmer function in the 'lme4' package (Bates *et al.*, 2015), and model test statistics were assessed via the Anova function in the 'car' package (Fox & Weisberg, 2019). R^2 values (Nakagawa, Johnson, & Schielzeth, 2017) were calculated with the 'MuMIn' package (Barton, 2020). Model diagnostic plots were visually assessed using the 'car' and 'lattice' (Sarkar, 2008) packages. None of the models demonstrated evidence of serious violations of the linearity or normality assumptions. The 'effects' package (Fox & Weisberg, 2018) was used to generate plots to visualize the impact of each independent trait on the dependent trait of the model.

Table 4-3: Mixed effects model terms		
Dependent Terms		
Variable	Description	Units

ADF	Acid detergent fiber. A measure of the percent of a forage that is not digestible. An estimate of forage digestibility.	% indigestible
NDF	Neutral detergent fiber. A measure of the percent of a forage that is difficult to digest plus the percent that is not digestible. An estimate of forage potential.	% poor digestibility
dry_ton_acre	Quantitative variable estimating the potential yield of an accession in tons/acre based on the dry biomass weight of the harvested plot sample.	Tons/acre
Independent Terms		
Variable	Description	Units
Treatment (WC1 only)	2-category treatment variable: Dryland, which received no additional moisture and Irrigated, which prior to heading received 5 additional inches of water relative to the dryland environment	2-level factor variable
days_to_head	Number of days from planting required for an accession to reach heading stage – defined as 50% of main tillers reaching Zadok 51	Days
head_soft_dif	Number of days between heading and soft-dough for a given accession	Days
days_to_soft	Number of days from planting required for an accession to reach soft-dough stage– defined as 50% of main tillers reaching Zadok 85	Days
soft_mat_dif	Number of days between soft-dough and maturity for a given accession	Days
days_to_mat	Number of days from planting required for an accession to reach maturity – Defined visually no green color remained in 50% of the main tiller heads, including the awns and glumes	Days
plant_ht_in	Average plant height of accession at maturity, taken from ground level to top of head excluding awns if present.	Inch

Associated with each collected forage sample are digestibility values, the sample accession, the flowering date, soft-dough date, maturity date, and the estimated biomass yield (Table 4-3). From the heading, soft-dough, and maturity dates, the number of days between each of these key plant physiological points was also calculated (Table 4-3). All

fixed and random effects to be included in initial modeling have been summarized in Tables 4-3 through 4-5. Models were assessed for three dependent variables: forage dry matter yield, ADF, and NDF.

Data were collected on 1200 WC1 samples, 600 WC2 samples, and 164 BP samples. After removing observations with missing values, WC1 consisted of 1172 observations on 264 accessions, WC2 consisted of 599 observations on 264 accessions, and the BP trial consisted of 164 observations on 79 accessions.

Modeling Forage Traits with the WC1: WC1 models were fit with accession, year, and block as random effects with block nested within year (Table 4-4). Beginning with a full model including all independent terms (Table 4-3), a backward stepwise method was used to select the optimal models for each of the dependent variables ADF, NDF, and dry_ton_acre using an F-test based 0.05 p-value cutoff as the selection criteria. In order to ensure that the WC1 models could be compared to the validation models, the WC1 models were restricted to contain no interaction terms, as models containing interaction terms were too complex to be fit to the smaller BP data set. For each of the dependent terms, the model selected on the WC1 data was then assessed for fit on the validation data sets. The selected model terms are presented in Tables 4-6 and 4-7.

Table 4-4: Random effects to be included in the WC1 models

Variable	Description	Units
block	40-category variable with 10 levels nested within each treatment X year field environment	40-level factor variable
year	2-category field-season variable: 2016 and 2017	2-level factor variable
accession	Unique ID assigned to each genetically unique barley line. The same accessions were planted	264 level factor variable

	across all four treatment X year field environments	

Validating Modeling Results with the WC2 and BP: The WC2 and BP validation trials were modeled separately so that differences between the diverse, global population and the Montana-adapted lines could be assessed.

Table 4-5: Random effects to be included in the WC2 models		
Variable	Description	Units
year	2-category field-season variable: 2017 and 2018	2-level factor variable
block	20-category variable with 10 levels nested within each trial	20-level factor variable
accession	Unique ID assigned to each genetically unique barley line. The same accessions were planted across both years	264-level factor variable
Random effects to be included in the BP models		
Variable	Description	Units
trial	4-category treatment variable: 2018-Preliminary Forage Trial, 2019-Preliminary Forage Trial, 2018-Advanced Forage Trial, 2019-Advanced Forage Trial	4-level factor variable
block	12-category variable with 3 levels nested within each trial	12-level factor variable
accession	Unique ID assigned to each genetically unique barley line. Accessions varied by trial and year as lines were advanced or dropped depending on performance within the breeding program.	79-level factor variable

Initial models were fit with accession, trial, and block as random effects with block nested within trial (Table 4-3, 4-4). After initial model assessment, block was found to explain no variation in the BP trials and was dropped as a random effect (Table 4-6, Table 4-7).

Table 4-6: Random effects by population and fixed effects determined by backward stepwise model selection on WC1 experimental population

Model Random Effects by Population	Dependent Variable Y	Model Fixed Effects selected on WC1 population
WC1 $Y_{ijk} = \mu_{ij} + accession_i + year_j + block_{jk} + \epsilon_{ijk}$	ADF	$\mu_{ij} = \beta_0 + \beta_1 days\_to\_head + \beta_2 head\_soft\_dif + \beta_3 soft\_mat\_dif + \beta_4 plant\_ht\_in + \beta_5 I_{treatment=irrigated}(WC1\ only)$
WC2 $Y_{ijk} = \mu_{ij} + accession_i + year_j + block_{jk} + \epsilon_{ij}$	NDF	$\mu_{ij} = \beta_0 + \beta_1 days\_to\_head + \beta_2 soft\_mat\_dif + \beta_3 plant\_ht\_in$
BP $Y_{ij} = \mu_{ij} + accession_i + trial_{ij} + \epsilon_{ij}$	Biomass	$\mu_{ij} = \beta_0 + \beta_1 days\_to\_head + \beta_2 soft\_mat\_dif + \beta_3 plant\_ht\_in + \beta_4 I_{treatment=irrigated}(WC1\ only)$

Model term definitions given for WC1 as an example:
$accession_i$: Accession random intercept (i^{th} level, i = 1, … , 267 = I), where $accession_i \sim N(0,\hat{\sigma}^2_{entryf})$;
$year_{ij}$: Year random intercept (j^{th} year within i^{th} accession), where $year_{ij} \sim N(0,\hat{\sigma}^2_{yearf})$;
$block_k$: Block random intercept (k^{th} block, k = 1, … , 40 = K), where $block_k \sim N(0,\hat{\sigma}^2_{blockf})$;
ϵ_{ij}: Random error at plot level (within j^{th} year within i^{th} accession, where $\epsilon_{ij} \sim N(0,\hat{\sigma}^2_{\epsilon})$;
μ_{ij}: within and between accession fixed effects

Table 4-7: Selected model components by dependent variable and experimental population

Dependent Variable	Data Set	Intercept	days_to_head	head_soft_dif	soft_mat_dif	plant_ht_in	Treatment=Irr	R^2m	R^2c
ADF									
	WC1	34.38	-0.0880**	-0.3585***	0.1090*	0.2192***	1.6769***	0.3457	0.4450
	WC2	30.86	-0.0426	-0.2824***	0.1924***	0.2693***	-	0.1229	0.6309
	BP	72.75	-0.4261*	-0.7204***	-0.2104**	0.2175*	-	0.2343	0.2443
Dependent Variable	**Data Set**	**Intercept**	**days_to_head**	**-**	**soft_mat_dif**	**plant_ht_in**	**-**	**R^2m**	**R^2c**
NDF									
	WC1	41.08	0.1549***	-	0.0816*	0.1700***	-	0.0575	0.5700
	WC2	37.60	0.1205***	-	0.2599***	0.2119***	-	0.0668	0.6808
	BP	60.94	0.0183	-	-0.3138***	0.0514	-	0.1015	0.1346
Dependent Variable	**Data Set**	**Intercept**	**days_to_head**	**head_soft_dif**	**-**	**plant_ht_in**	**Treatment=Irr**	**R^2m**	**R^2c**
Biomass Yield									
	WC1	-5.5826	0.0925***	0.1238***	-	0.0800***	0.5110***	0.5939	0.7812
	WC2	-4.4984	0.0938***	0.10341***	-	0.0683***	-	0.2360	0.6147
	BP	-14.33	0.1841**	0.2750***	-	0.0941*	-	0.1614	0.4731

Significance codes: p-value is less than 0.0001 '***', 0.001:'**', 0.01:'*', 0.05:'·'

Results

Results of Mixed Effects Modeling for ADF

Very strong evidence was found in all three populations for an impact of timing of plant development on ADF with the number of days between heading and soft-dough having the largest and most stable impact. For a one day increase in the period between heading and soft-dough date, ADF was estimated to decrease by 0.359, 0.282, and 0.720% in the WC1, WC2 and BP populations respectively. (Table 4-7)

The number of days between soft-dough date and maturity date was also a significant term in all three models, however the direction of the impact was not consistent. While increasing the number of days between soft-dough and maturity by one day was estimated to increase ADF by 0.109 to 0.192% in the genetically diverse WC populations, the same change was estimated to decrease ADF by 0.210% in the breeding program population. (Table 4-7)

In addition to plant developmental timing, strong evidence was also found for a stable impact of plant height on ADF. For a one-inch increase in plant height, ADF was estimated to increase in the WC1, WC2, and BP populations by 0.219, 0.269, or 0.218% respectively. (Table 4-7)

For the WC1, WC2, and BP populations, fixed effects alone explained 34.57, 12.29, and 23.42 % of the variance in the respective datasets. When accounting for fixed effects as well as random effects, a total of 44.50, 63.09, and 24.43 % of the variance was explained. (Table 4-7)

Results of Mixed Effects Modeling for NDF

In contrast to the ADF models, the variance explained by the fixed effects in the NDF models was much lower. For the WC1, WC2, and BP populations, fixed effects alone explained only 5.75, 6.68, and 10.15 % of the variance in the respective datasets. The total variance explained by each model was higher when accounting for both fixed and random effects. A total of 57.0, 68.08, and 13.46 % of the variance was then explained by each model. (Table 4-7)

Like ADF, evidence was found for an impact of plant developmental timing on NDF. Unlike ADF however, no developmental term exhibited an impact that was stable across all populations. The number of days between soft-dough date and maturity date was statistically significant in all three NDF models, but as was seen in the ADF models, the direction of the impact of the trait differed between the world core and breeding program populations. In the BP population, for a one day increase in the period between soft-dough and maturity date, NDF was estimated to decrease by 0.314%. In the global populations however, the same one-day increase was estimated to increase NDF by 0.081 to 0.260%. (Table 4-7)

Results of Mixed Effects Modeling for Biomass Yield

For biomass yield, fixed effects alone explained 59.39, 23.6, and 16.14 % of the variance in the WC1, WC2, and BP datasets. A total of 78.12, 61.47, and 47.31% of the variance was explained in each dataset when accounting for fixed effects as well as random effects. (Table 4-7)

Of the three dependent variables, biomass yield was the trait with the most stable independent variables across the three populations. Strong evidence was found for an impact of both days to heading and number of days between heading and soft-dough on biomass yield. For a one day increase in days to heading, biomass was estimated to increase by 0.093, 0.094, and 0.184 ton/acre for the WC1, WC2, and BP populations respectively. Very strong evidence was found for an impact of number of days between heading and soft-dough date on biomass. Similarly, for a one day increase in the period between heading and soft-dough date, biomass was estimated to increase by 0.124, 0.103, and 0.275 ton/acre for the three populations. (Table 4-7)

As expected, strong evidence was also found for an impact of plant height on biomass. For a one-inch increase in plant height, biomass was estimated to increase by 0.080, 0.068, and 0.094 ton/acre for the three populations. (Table 4-7)

Discussion

While the WC data sets were mostly balanced, the BP data set, due to the nature of breeding program trialing, was very imbalanced. Only 9 lines were present in both 2018 and 2019 and only the two check varieties, Lavina and Hays, were present in all trial x year environments. While this certainly limited the modeling power of the data set, nevertheless useful models for the dependent variables ADF and Biomass Yield were discovered. For NDF, only the BP experimental population was able to produce a model with a marginal R^2 value greater than 0.1, which limits the conclusions that can be drawn for this trait.

Modeling ADF

When modeling ADF there were three terms that had consistent impacts across all three models: days to heading, days between heading and soft-dough, and plant height (Table 4-7). Soft-dough to maturity duration was found to be important in all three models, however the sign of the effect was not consistent between populations. This would indicate that the term would not be a good candidate as a selection proxy for ADF.

Heading to soft-dough date: There was very strong evidence in all three models that as the period between heading and soft-dough date increased, ADF decreased (Table 4-7). Depending on the model, increasing the heading to soft-dough period by a single day led to a 0.28 to 0.72% reduction in ADF with the greatest impact seen in the BP population (Figure 4-1).

A link between plant maturity and forage quality has been well documents in crops like alfalfa (Andrzejewska *et a.*, 2014; Hakl *et al.*, 2019; Hintz & Albrecht, 1991; Owens *et al*., 1995; Santillano-Cázares *et a.*, 2014; Sulc *et al*., 1997). While most studies focus on a certain stage of maturity, our model indicates that for these barley populations the duration between heading and soft-dough is more important than either heading or soft-dough date alone. This leads to the conclusion that selecting for increased heading to soft-dough date in forage barley lines in future breeding trials could lead to lowering ADF.

Heading date: There was some evidence that delaying heading date was also linked to a reduction in ADF. This relationship was observed in both the WC1 and BP

populations. (Table 4-7) Taken together these two pieces of evidence seem to indicate that selecting for a delayed heading date could act as a proxy for selecting for lowered ADF in early stages of the breeding program, however the relationship is less consistent than that observed between ADF and heading to soft-dough duration.

Plant height: Finally, in all three ADF models a very consistent relationship was seen between ADF and plant height (Figure 4-1). In all three populations increasing plant height by one inch was linked to increasing ADF by about 0.24%. ADF increased by 0.23%, 0.27%, and 0.22% for the WC1, WC2, and BP populations respectively. (Table 4-7) This relationship has also been documented by numerous researchers in alfalfa (Andrzejewska *et al.*, 2014; Hakl *et al.*, 2019; Hintz & Albrecht, 1991; Owens *et al.*, 1995; Santillano-Cázares *et al.*, 2014; Sulc *et al.*, 1997).

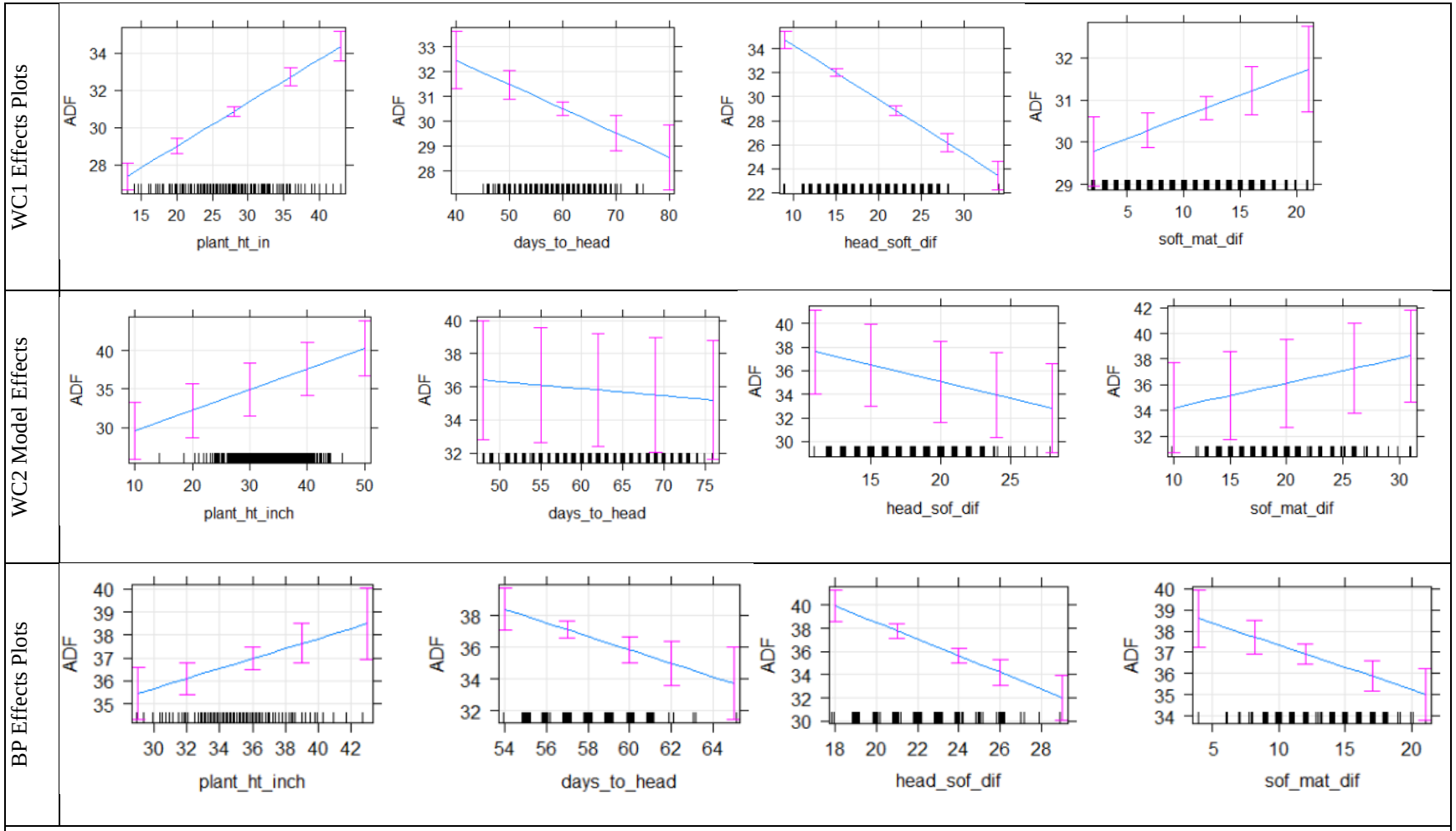


Figure 4-1: Effects plots demonstrate the impact of each of the independent terms on ADF. Bars represent 95% confidence intervals. The WC1 effect plot for treatment not depicted.

Modeling NDF

Low marginal R^2 values relative to conditional R^2 values indicate that the variance being explained by all three of the NDF models is largely the variance contained in the random effects, i.e. year-to-year or accession-to-accession variance (Table 4-7). Plant height and heading date had relatively consistent slopes across the models but based on their p-values there was very weak evidence for their impacts on NDF in the BP model ($F(1,90.75) = 0.247$, p-value = 0.62) and ($F(1,87.16) = 0.02$, p-value = 0.89) respectively.

Soft-dough to maturity date: The duration between soft-dough date and maturity date was the only term that met the F-test based 0.05 p-value cutoff across all three populations (Table 4-7). However, the direction of the effect varied between the WC and BP populations (Figure 4-2). While the relationship was positive in the WC populations (as the number of days between soft-dough and maturity increased, NDF increased), it was negative in the BP population. Interestingly this is the same trend that this trait exhibited in the models for ADF.

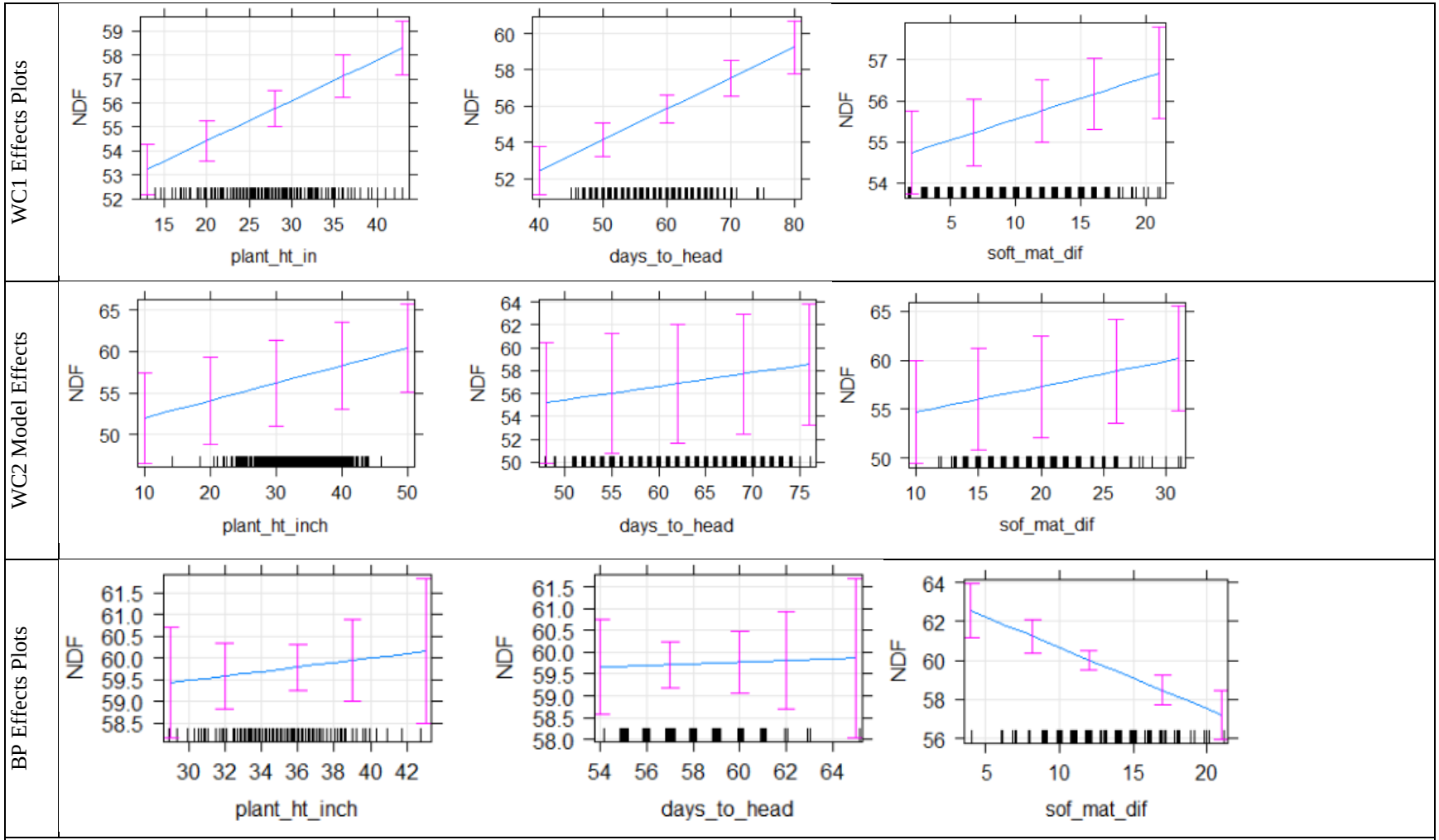


Figure 4-2: Effects plots demonstrate the impact of each of the independent terms on NDF. Bars represent 95% confidence intervals.

Although the NDF models have identified fixed effects that meet the F-test based 0.05 p-value cutoff, the high proportion of unexplained variation and the inconsistencies in the independent variables across the three models make it difficult to draw conclusions about the practical impact these variables may have on NDF from a breeding perspective.

Modeling Forage Yield Conclusions

Across the three modeling populations all three fixed effects terms, days to heading, days between heading and soft-dough, and plant height, had consistent impacts on Biomass Yield (Table 4-7, Figure 4-3). The variation captured by the models accounted for between 61.78 and 16.14% of the variation found in the populations (WC1 to BP respectively).

Heading to soft-dough date: The term with the largest impact on biomass (excluding irrigation in the WC1 population) was the number of days between heading and soft-dough date. Biomass was shown to increase by 0.103 to 0.275 tons/acre with an increase of one day in the heading to soft-dough period. Increases in days to heading also appeared to be linked to increases in biomass with a one day delay equating to a 0.09 to 0.18 increase in tons/acre (WC to BP populations respectively). (Table 4-7)

Plant height: Plant height also had a consistent impact on biomass with a one-inch increase in height resulting in a 0.07 to 0.09 increase in tons/acre (WC to BP populations respectively). (Table 4-7)

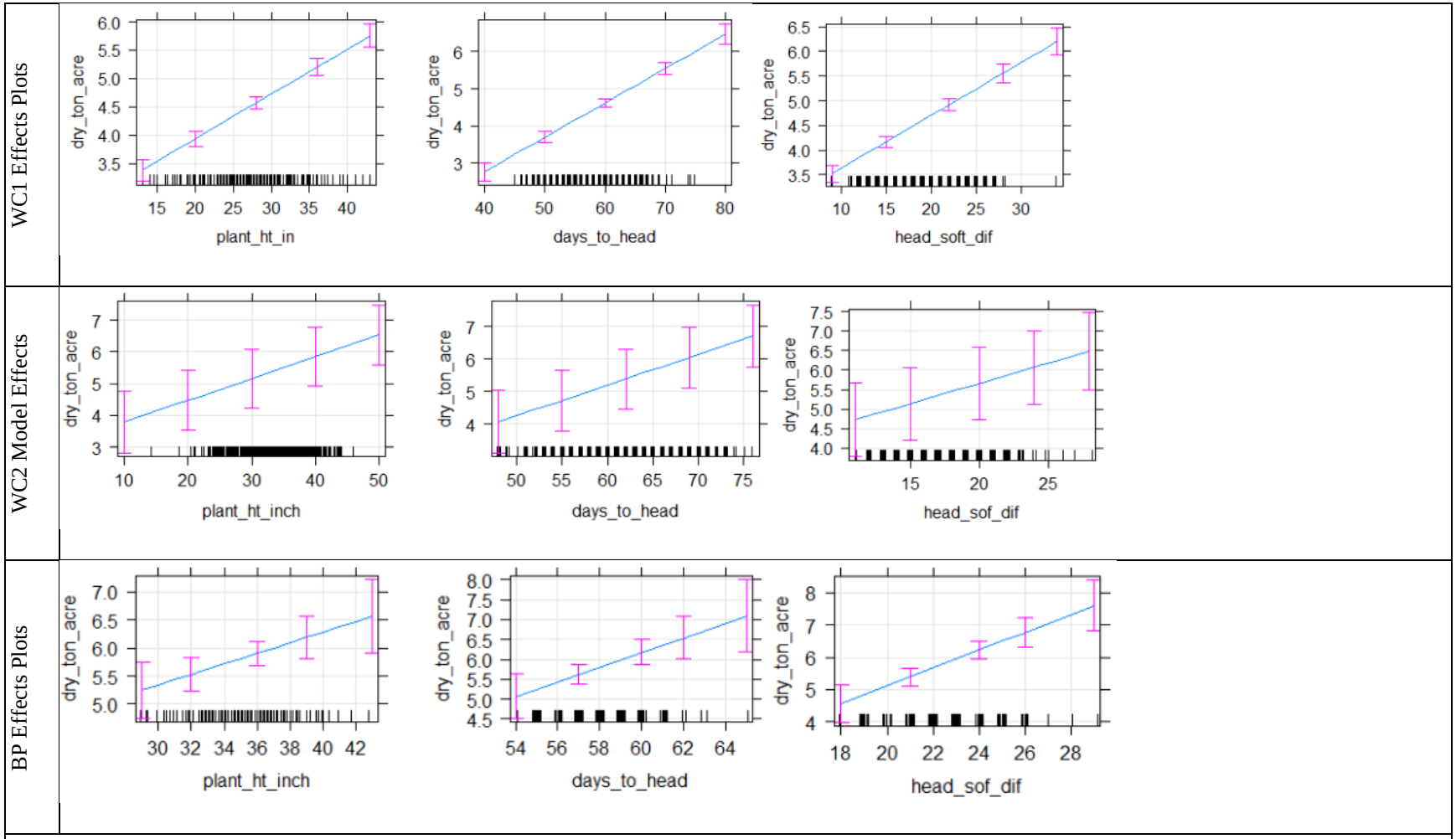


Figure 4-3: Effects plots demonstrate the impact of each of the independent terms on biomass yield. Bars represent 95% confidence intervals. The WC1 effect plot for treatment not depicted.

Overall Modeling Conclusions

Overall, the fixed effects terms with the clearest and most consistent impacts on forage traits were heading date, number of days between heading and soft-dough date, and plant height. Links between plant maturity, plant height, and forage traits has been well documented in intensively studied crops such as alfalfa (Andrzejewska *et al.*, 2014; Hakl *et al*., 2019; Hintz & Albrecht, 1991; Owens *et al*., 1995; Santillano-Cázares *et al*., 2014; Sulc *et al.*, 1997), but this represents the first time that these relationships have been modeled in forage barley.

Not all populations demonstrated the same relationships. The trait 'days between soft-dough and maturity', exhibited changes in the direction of the effect between the WC and BP populations. This could be due to the differences in the amount of genetic diversity in these populations. As previously stated, the BP lines are largely the result of crosses including elite Montana cultivars. The forage varieties Hays, Lavina, and Haybet are parents for all but 4 of the lines in the population. Since Hays and Lavina are sister lines, the result of a cross between the forage line Haybet and the feed variety Baronesse, this means that 95% of the lines in the BP are about 50% genetically identical.

For the highly genetically diverse WC populations, more of the genes impacting forage traits will be polymorphic and thus a greater number of genes will contribute to the variation in forage yield or quality seen in the population. Conversely, in a set of individuals such as BP population where many of the genes in the population are fixed, only the influence of the remaining genetic diversity can cause variation in the measured traits.

Based on modeling results, selecting for delayed heading date and a longer period between heading date and soft-dough date in the early stages of future breeding program populations could result in improved ADF and biomass yield. Selecting for an increase in plant height would have a more complicated impact on forage traits. Increasing plant height was linked to increasing both biomass yield (desirable) and increasing ADF (undesirable). However, when increasing plant height is combined with the aforementioned selection for delayed heading and longer periods between heading and soft-dough, the overall impact could still be net positive for both traits.

For example, looking at the BP models for ADF and Biomass Yield, when the population averages are entered into the models, the predicted forage traits for the average line are:

ADF:

$\hat{\mu}_{ij} = 72.75 + (-0.4261 * 57.94) + (-0.7204 * 22.48) + (-0.2104 * 12.95) + (0.2175 * 34.86)$ = 36.72 %ADF

Biomass:

$\hat{\mu}_{ij} = -14.33 + (0.1841 * 57.94) + (0.2750 * 22.48) + (0.0941 * 34.86)$ = 5.80 biomass tons/acre

If heading date was delayed one day, heading to soft-dough was extended one day, and plant height was increased by one inch, the models predict the following:

ADF:

$\hat{\mu}_{ij} = 72.75 + (-0.4261 * 58.94) + (-0.7204 * 23.48) + (-0.2104 * 12.95) + (0.2175 * 35.86)$ = 35.80 %ADF

Biomass:

$\hat{\mu}_{ij} = -14.33 + (0.1841 * 58.94) + (0.2750 * 23.48) + (0.0941 * 35.86)$ = 6.35 biomass tons/acre

The cumulative result predicted by the models is that biomass yield increases by 0.55 tons/acre while ADF decreases by 0.92%, an ideal scenario from a forage breeder's perspective.

In conclusion, models developed using genetically diverse populations can provide insight into less diverse breeding populations. Traits such as heading date, number of days between heading and soft-dough date, and plant height have been shown to consistently impact forage traits such as yield and quality in the modeled populations. While direct selection for forage quality and yield is not possible in the early stages of a forage breeding program, selection for agronomic traits such as heading date and plant height is feasible. Our findings indicate that heading date, number of days between heading and soft-dough date, and plant height may be useful proxies for forage traits when direct selection is not practical. It should be noted that the modeling results obtained in this study can only be directly applied to the environments tested and different places and environmental conditions may produce very different results. Future research should test the predictive ability of these models across more years and environments.

CHAPTER FIVE: PROJECT RESULTS AND CONCLUSIONS

Breeding for Forage Barley in Montana

Value of forage breeding research

Forage crops are important to Montana. Forages are the foundation of the cattle industry, which generates 2 billion dollars of revenue annually for the state. While many of the calories needed by a herd of cattle in Montana will be obtained by grazing rangeland, supplemental winter hay must also be grown, harvested, and fed out, a resources and labor intensive process. Winter feed is the single largest expense for most Montana cattle producers (USDA National Agricultural Statistics Service, 2020). Forage breeders can help cut the costs of forage production by creating and releasing forage varieties with improved yield and quality.

Barley (*Hordeum vulgare spp. vulgare L.*) is considered one of the most drought tolerant of the annual cereals and spring barley has been shown to out yield established perennial forages under drought conditions in central Montana (Cash *et al.*, 2006). It is saline tolerant (Maas, 1993; Steppuhn & Raney, 2005) and adapted to the short growing season of the region. Just as important to the long-term success of the crop, it is considered to be adaptable to a greater range of climate than any other cereal (Britannica, 2020) and is a promising model organism for how grasses adapt to climate change (Dawson *et al.*, 2015). Taken together, barley's characteristics indicate that it would be an ideal target for forage breeders trying to fill the increased need of Montana growers for an adaptable and sustainable roughage crop.

Hurdles to barley forage breeding research

Successful plant breeding requires the ability to select for desirable traits and a population that is genetically diversity for the traits of interest. The larger the genetic variability in a population under selection pressure, the greater the potential genetic progress (Moreno-González & Cubero, 1993). For barley breeders this means that selection in the earliest rounds of inbreeding, when genetic diversity is the greatest, is the most important point in the selection process. However, early stages of the breeding program have very large populations and only single plants might be available for selection. Since forage yield and quality are very difficult to measure on single plants or large populations, selection opportunities at these crucial early stages are lost.

Accelerating barley forage breeding research

As stated previously, in order to accelerate the discovery and release of superior forage barley lines for Montana, the following research and breeding goals were identified:

1) Since stage of maturity is known to impact yield and quality (Casler & Vogel, 1999), implement daily monitoring of forage trials to ensure experimental plots are harvested at the same stage of maturity, eliminating that key source of variation in forage quality and thus making forage lines comparable across plots, environments, and year.

2) Develop near-infrared reflectance (NIR) technology to make screening for forage quality faster and cheaper – allowing more samples, and thus

more barley lines at earlier stages of the breeding program and across more environments, to be tested.

3) Identify exotic germplasm with superior digestibility and biomass yield for incorporation into the MSU barley breeding program.

4) Utilize a genome wide association analysis to find genetic regions related to key forage and agronomic traits.

5) Use statistical modeling to a) examine the relationship between difficult to measure forage traits such as quality and yield, and easy to measure agronomic traits such as flowering time and plant height, b) identify agronomic traits that can be used as proxies for yield and quality in the earliest stages of the breeding program when genetic and phenotypic variability are at their greatest.

At the beginning of this project, several key points were identified in the forage barley breeding program where advances were needed to make selection for superior forage lines more successful. Steps were taken to make current selection practices better. Stage of maturity is known to impact forage quality, if the forage quality of lines is to be compared the lines must be harvested at the same stage of maturity. This was the first issue addressed by the breeding program.

Next, for forage quality selection to be possible even in the later stages of the breeding program, when forage samples can be collected from full plot trials, NIR technology needed to be developed. NIR is the cheapest and fastest way to determine the forage quality of a sample. When hundreds of samples must be analyzed for forage

quality. To this end custom NIR calibration curves were developed for barley forage quality. Since their development, these curves have allowed more than 3000 cereal forage samples to be evaluated.

As previously stated, the larger the genetic variability in a population under selection pressure, the greater the potential genetic progress (Moreno-González & Cubero, 1993). Genetic variability is important in a breeding program because it allows for greater potential for improvement in traits of interest, but it also introduces the genetic diversity required to ensure that the breeding program can adapt to changing environmental factors. Immense resources for genetic diversity exist in the germplasm repositories such as the USDA-ARS National Plant Germplasm System. As part of the overall goal of accelerating the development of superior forage lines, exotic germplasm was screened for the beneficial traits. Since the start of the project, five hundred and twenty global accessions have been screened for forage yield and quality potential. Starting in the fall of 2017, global lines identified in the trials to have superior forage yield and quality characteristics were incorporated into the MSU barley forage breeding program. The offspring of these crosses will be in field yield trials for the first time in the 2020 field season.

Finally, one of the best ways to accelerate the development of forage varieties is to make selection possible at the earliest stages of the breeding program when genetic diversity is the highest. Measuring forage traits is very difficult in the early stages of the breeding program. In the first two rounds of field selection thousands of individuals must be evaluated for each cross and only single plants or a single row of plants are available

for evaluation. Finding agronomic traits that can be measured on thousands of single plants and that are related to forage traits could allow for some selection for forage quality or yield when traditional forms of direct measurement are not possible.

Five hundred and twenty global barley accessions and seventy-nine lines selected from crosses between Montana adapted forage varieties were screened for both forage and agronomic traits. Mixed effects models were then developed from these populations indicated that traits such as heading date, plant height, and the number of days between heading and soft-dough date, impact forage yield and quality and could provide proxies for forage trait selection. These models provide a promising resource for selection when direct selection is not feasible.

Conclusions on research goals

Accelerating the development of a superior forage variety for Montana means finding ways to make selection for forage traits faster, more effective, and feasible as early in the breeding process as possible. Genetically diverse populations taken from global germplasm repositories can enable several of these goals.

Screening highly diverse populations allows beneficial exotic germplasm to be discovered and incorporated into the breeding program. The high diversity of these populations also means that these populations are highly variable for the forage and agronomic traits measured. Thus, these lines lend themselves to creating sample sets for NIR calibration development and mixed modeling.

REFERENCES CITED

Åberg, E. (1938). *Hordeum Agriocrithon Nova Sp: A Wild Sixrowed Barley*: Almqvist & Wilksells Boktryckeri-AB.

Alforex Seeds. (2015). Hi-Gest Alfalfa: The first low lignin alfalfa on the market. In a. a. o. D. A. Alforex Seeds LLC (Ed.). Woodland, CA.

Allen, M., & Beck, J. (1996, 4-5 March 1996). *Relationship between spring harvest alfalfa quality and growing degree days.* Paper presented at the Twenty-sixth National Alfalfa Symposium, East Lansing, MI.

Allen, M. S., O'Neil, K. A., Dado, R. G., & Kohn, R. A. (1992, 20-21 May 1992). *Variation in fiber content and fiber digestibility of alfalfa and corn forages.* Paper presented at the Tri-State Dairy Nutrition Conference, Fort Wayne, IN.

Alqudah, A. M., Koppolu, R., Wolde, G. M., Graner, A., & Schnurbusch, T. (2016). The Genetic Architecture of Barley Plant Stature. *Front Genet, 7*, 117. doi:10.3389/fgene.2016.00117

Andrzejewska, J., Contreras-Govea, F. E., & Albrecht, K. A. (2014). Field prediction of alfalfa (Medicago sativa L.) fibre constituents in northern Europe. *Grass and Forage Science, 69*(2), 348-355. doi:10.1111/gfs.12069

Arnold, A. M., Albrecht, K. A., Barker, D. J., & Sulc, R. M. (2019). Predictive Equations for Alfalfa Quality (PEAQ) Can Be Used with Reduced-Lignin Alfalfa. *Crop, Forage & Turfgrass Management, 5*(1). doi:10.2134/cftm2019.01.0004

Arnold, A. M., Cassida, K. A., Albrecht, K. A., Hall, M. H., Min, D., Xu, X., Sulc, R. M. (2019). Multistate Evaluation of Reduced-Lignin Alfalfa Harvested at Different Intervals. *Crop Science, 59*(4), 1799-1807. doi:10.2135/cropsci2019.01.0023

Arojju, S. K., et al. (2016). "Markers associated with heading and aftermath heading in perennial ryegrass full-sib families." *BMC Plant Biol*, 16(1): 160

Backes, G., Graner, A., Foroughi-Wehr, B., Fishbeck, G., Wenzel, G., & Jahoor, A. (1995). Localization of quantitative trait loci (QTL) for agronomic important characters by the use of a RFLP map in barley (Hordeum vulgare L.). *Theoretical & Applied Genetics, 90*(294–302). doi:10.1007/BF00222217

Barton, K. (2020). MuMIn: Multi-Model Inference. R package version 1.43.17. Retrieved from https://CRAN.R-project.org/package=MuMIn

Barua, U. M., Chalmers, K. J., Thomas, W. T. B., Hackett, C. A., Lea, V., Jack, P., Powell, W. (1993). Molecular mapping of genes determining height, time to

heading, and growth habit in barley (Hordeum vulgare). *Genome, 36*(1080–1087). doi:10.1139/g93-143

Bates, D., Maechler, M., Ben Bolker, S., & Walker, K. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:doi:10.18637/jss.v067.i01

Beier, S., et al. (2017). "Construction of a map-based reference genome sequence for barley, Hordeum vulgare L." *Sci Data,* 4: 170044.

Bekele, E. (1983). A differential rate of regional distribution of barley flavonoid patterns in Ethiopia, and a view on the center of origin of barley. *Hereditas, 98*(2), 269-280.

Ben-Ghedalia, D., Kabala, A., & Miron, J. (1995). Composition and in-vitro digestibility of carbohydrates of wheat plants harvested at bloom and soft-dough stages. *Journal of the Science of Food and Agriculture, 68*(1), 111-116. doi:10.1002/jsfa.2740680118

Bezant, J., Laurie, D., Pratchett, N., Chojecki, J., & Kearsey, M. (1996). Marker regression mapping of QTL controlling flowering time and plant height in a spring barley (Hordeum vulgare L.) cross. *Heredity, 77*, 64–73. doi:10.1038/hdy.1996.109

Biazzi, E., et al. (2017). "Genome-Wide Association Mapping and Genomic Selection for Alfalfa (Medicago sativa) Forage Quality Traits." *PLoS One,* 12(1): e0169234.

Bittman, S. (1999). *Advanced Forage Management - a production guide for coastal British Columbia and the Pacific Northwest*. Agassiz, BS, Canada: Pacific Field Corn Association.

Brenner, E. A., Zein, I., Chen, Y., Andersen, J. R., Wenzel, G., Ouzunova, M., . . . Lübberstedt, T. (2010). Polymorphisms in O-methyltransferase genes are associated with stover cell wall digestibility in European maize (Zea mays L.). *BMC plant biology, 10*, 27-27. doi:10.1186/1471-2229-10-27

Britannica, T. E. o. E. (2020). Barley. In P. B. Adam Augustyn, Brian Duignan, Alison Eldridge, Erik Gregersen, Amy McKenna, Melissa Petruzzello, John P. Rafferty, Michael Ray, Kara Rogers, Amy Tikkanen, Jeff Wallenfeldt, Adam Zeidan, and Alicja Zelazko (Ed.), *Encyclopædia Britannica*: Encyclopædia Britannica, inc.

Calinski, T., & Harabasz, J. (1974). A Dendrite Method for Cluster Analysis: Communications in Statistics. *Theory and Methods, 3*, 1-27. Retrieved from http://dx.doi.org/10.1080/03610927408827101

Cash, D. (2006). Montana State University Animal & Range Sciences Extension Service. *Extension Forage Program*. Bozeman, MT. https://animalrangeextension.montana.edu/forage/. Accessed: Oct 23, 2016.

Cash, D., Surber, L. M. M., & Wichman, D. M. (2006). *Annual forage production in drought.* Paper presented at the Idaho Alfalfa and Forage Conference, Twin Falls, ID.

Casler, M., & Vogel, K. (1999). Accomplishments and impact from breeding for increased forage nutritional value. *Crop Science, 39*(1), 12-20. doi:DOI 10.2135/cropsci1999.0011183X003900010003x

Casler, M. D. (2013). Breeding for Increased Forage Quality. In S. B. a. D. Hunt (Ed.), *Cool forages - Advanced management of temperate forages* (pp. 60-63). Agassiz, BC, Canada: Pacific Field Corn Association.

Casler, M. D., Buxton, D. R., & Vogel, K. P. (2002). Genetic modification of lignin concentration affects fitness of perennial herbaceous plants. *Theoretical and Applied Genetics, 104*(1), 127-131. doi:10.1007/s001220200015

Casler, M. D., & Vogel, K. P. (1999). Accomplishments and impact from breeding for increased forage nutritional value. *Crop Science, 39*(1), 12-20. doi:DOI 10.2135/cropsci1999.0011183X003900010003x

Catling, P. M., McElroy, A. R., & Spicer, K. W. (1994). Potential forage value of some eastern Canadian sedges (Cyperaceae: Carex). *Journal of Range Management, 47*(3), 220-230.

Cherney, J. H. (1995, 27-28 February 1995). *Spring alfalfa harvest in relation to growing degree days.* Paper presented at the Twenty-fifth National Alfalfa Symposium, Syracuse, NY.

Cherney, J. H., & Marten, G. C. (1982a). Small Grain Crop Forage Potential: I. Biological and Chemical Determinants of Quality, and Yield1. *Crop Science, 22*(2), cropsci1982.0011183X002200020007x. doi:10.2135/cropsci1982.0011183X002200020007x

Cherney, J. H., & Marten, G. C. (1982b). Small Grain Crop Forage Potential: II. Interrelationships Among Biological, Chemical, Morphological, and Anatomical Determinants of Quality1. *Crop Science, 22*(2), cropsci1982.0011183X002200020010x. doi:10.2135/cropsci1982.0011183X002200020010x

Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., Waugh, R. (2012). Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed

to spring growth habit and environmental adaptation in cultivated barley. *Nature Genetics, 44*(12), 1388-1392.

Crovetto, G. M., Galassi, G., Rapetti, L., Sandrucci, A., & Tamburini, A. (1998). Effect of the stage of maturity on the nutritive value of whole crop wheat silage. *Livestock Production Science, 55*(1), 21-32. doi:https://doi.org/10.1016/S0301-6226(98)00125-0

Dawson, I. K., Russell, J., Powell, W., Steffenson, B., Thomas, W. T. B., & Waugh, R. (2015). Barley: a translational model for adaptation to climate change. *New Phytologist, 206*(3), 913-931. doi:10.1111/nph.13266

Digel, B., Tavakol, E., Verderio, G., Tondelli, A., Xu, X., Cattivelli, L., von Korff, M. (2016). Photoperiod-H1 (Ppd-H1) Controls Leaf Size. *Plant Physiol, 172*(1), 405-415. doi:10.1104/pp.16.00977

Ellis, R. H., Roberts, E. H., Summerfield, R. J., & Cooper, J. P. (1988). Environmental Control of Flowering in Barley (Hordeum vulgare L.). II. Rate of Developement as a Function of Temperature and Photoperiod and its Modification by Low-temperature Vernalization. *Annals of Botany, 62*(2), 145-158. doi:10.1093/oxfordjournals.aob.a087645

Ellis, R. H., Summerfield, R. J., Roberts, E. H., & Cooper, J. P. (1989). Environmental Control of Flowering in Barley (Hordeum vulgare). III. Analysis of Potential Vernalization Responses, and Methods of Screening Germplasm for Sensitivity to Photoperiod and Temperature. *Annals of Botany, 63*(6), 687-704. doi:10.1093/oxfordjournals.aob.a087797

Endelman, J. B. (2011). Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome, 4*(3), 250-255. doi:10.3835/plantgenome2011.08.0024

Everitt, B., & Hothorn, T. (2011). *An Introduction to Applied Multivariate Analysis with R*: Springer Science + Business Media, LLC.

Fang, Y., & Wang, J. (2012). Selection of the number of clusters via the bootstrap method. *Computational Statistics & Data Analysis, 56*(3), 468-477.

Federer, W. T. (1956). "Augmented (or Hoonuiaku) Designs." *Biometrics Unit Technical Reports Number BU-74-M.*

Federer, W. T. (1961). "Augmented Designs with One-Way Elimination of Heterogeneity." *Biometrics,* 17(3): 447-473

Fiber (Acid Detergent) and Protein (Crude) in Animal Feed and Forages: Near-infrared Reflectance Spectroscopic Method (989.03). (1990). In *Official Methods of Analysis* (15th Edition ed.).

Fick, G. W., & Onstad, D. W. (1988). Statistical Models for Predicting Alfalfa Herbage Quality from Morphological or Weather Data. *Journal of Production Agriculture, 1*(2), 160-166. doi:10.2134/jpa1988.0160

Fox, J., & Weisberg, S. (2018). Visualizing Fit and Lack of Fit in Complex Regression Models with Predictor Effect Plots and Partial Residuals. *Journal of Statistical Software, 87*(9), 1-27. Retrieved from https://www.jstatsoft.org/v087/i09

Fox, J., & Weisberg, S. (2019). An {R} Companion to Applied Regression. Thousand Oaks, CA. Retrieved from https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Fuller, D. Q., Willcox, G., & Allaby, R. G. (2011). Cultivation and domestication had multiple origins: arguments against the core area hypothesis for the origins of agriculture in the Near East. *World Archaeology, 43*(4), 628-652.

Garnsworthy, P. C., & Stokes, D. T. (1993). The nutritive value of wheat and oat silages ensiled on three cutting dates. *The Journal of Agricultural Science, 121*(2), 233-240. doi:10.1017/S0021859600077108

González, F. G., Slafer, G. A., & Miralles, D. J. (2002). Vernalization and photoperiod responses in wheat pre-flowering reproductive phases. *Field Crops Research, 74*(2), 183-195. doi:https://doi.org/10.1016/S0378-4290(01)00210-6

González, F. G., Slafer, G. A., & Miralles, D. J. (2003). Floret development and spike growth as affected by photoperiod during stem elongation in wheat. *Field Crops Research, 81*(1), 29-38. doi:https://doi.org/10.1016/S0378-4290(02)00196-X

Gordon, T., Wang, R., Bowman, B., Klassen, N., Wheeler, J., Bonman, J. M., Chen, J. (2020). Agronomic and genetic assessment of terminal drought tolerance in two-row spring barley. *Crop Science, 60*(3), 1415-1427. doi:10.1002/csc2.20040

Gorthy, S., Mayandi, K., Faldu, D., & Dalal, M. (2013). Molecular characterization of allelic variation in spontaneous brown midrib mutants of sorghum (Sorghum bicolor (L.) Moench). *Molecular Breeding, 31*(4), 795-803. doi:10.1007/s11032-012-9834-4

Gowri, G., Bugos, R. C., Campbell, W. H., Maxwell, C. A., & Dixon, R. A. (1991). Stress Responses in Alfalfa (Medicago sativa L.): X. Molecular Cloning and Expression of S-Adenosyl-1-Methionine:Caffeic Acid 3-O-Methyltransferase, a Key Enzyme of Lignin Biosynthesis, *Plant Physiol, 97*(1), 7-14. doi:10.1104/pp.97.1.7

Guo, D., Chen, F., Wheeler, J., Winder, J., Selman, S., Peterson, M., & Dixon, R. A. (2001). Improvement of in-rumen digestibility of alfalfa forage by genetic manipulation of lignin O-methyltransferases. *Transgenic Research, 10*(5), 457-464. doi:10.1023/A:1012278106147

Hakl, J., Mofidian, S. M. A., Kozova, Z., Fuksa, P., & Jaromir, S. (2019). Estimation of lucerne yield stability for enabling effective cultivar selection under rainfed conditions. *Grass and Forage Science, 74*(4), 687-695. doi:10.1111/gfs.12456

Helsel, Z. R., & Thomas, J. W. (1987). Small Grains for Forage1. *Journal of Dairy Science, 70*(11), 2330-2338. doi:https://doi.org/10.3168/jds.S0022-0302(87)80293-X

Hennig, C. (2015). fpc: Flexible Procedures for Clustering. R package version 2.1-10. Retrieved from https://CRAN.R-project.org/package=fpc

Hintz, R. W., & Albrecht, K. A. (1991). Prediction of Alfalfa Chemical Composition from Maturity and Plant Morphology. *Crop Science, 31*(6), 1561-1565. doi:10.2135/cropsci1991.0011183X003100060036x

Hopper, J. A., Peterson, H. H., & Burton, R. O. (2004). Alfalfa Hay Quality and Alternative Pricing Systems. *Journal of Agricultural and Applied Economics, 36*(3), 675-690. doi:10.1017/S1074070800026948

Huang, M., Liu, X., Zhou, Y., Summers, R. M., & Zhang, Z. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience, 8*(2). doi:10.1093/gigascience/giy154

Jia, C., et al. (2017). "Identification of genetic loci associated with crude protein and mineral concentrations in alfalfa (Medicago sativa) using association mapping." *BMC Plant Biol,* 17(1): 97.

Kennelly, J. J., & Weinberg, Z. G. (2015). Small Grain Silage. In D. R. Buxton, R. E. Muck, & J. H. Harrison (Eds.), *Silage Science and Technology* (Vol. 42, pp. 749-779).

Klevenhusen, F., Emsenhuber, C., Grausgruber, H., Petri, R. M., & Zebeli, Q. (2019). Effects of the orange lemma mutant line of barley cv. ‘Optic’ compared with its wild-type on the ruminal microbiome and fermentation tested with the rumen simulation technique. *Crop and Pasture Science, 70*(9), 789-799, 711. Retrieved from https://doi.org/10.1071/CP18288

Larson, S. R., Jensen, K. B., Robins, J. G., & Waldron, B. L. (2014). Genes and Quantitative Trait Loci Controlling Biomass Yield and Forage Quality Traits in Perennial Wildrye. *Crop Science, 54*(1), 111-126. doi:10.2135/cropsci2013.04.0261

Larson, S. R., Kishii, M., Tsujimoto, H., Qi, L., Chen, P., Lazo, G. R., Wang, R. R. C. (2012). Leymus EST linkage maps identify 4NsL–5NsL reciprocal translocation, wheat-Leymus chromosome introgressions, and functionally important gene loci. *Theoretical and Applied Genetics, 124*(1), 189-206. doi:10.1007/s00122-011-1698-1

Laurie, D. A., Pratchett, N., Snape, J. W., & Bezant, J. H. (1995). RFLP Mapping of Five Major Genes and Eight Quantitative Trait Loci Controlling Flowering Time in a Winter × Spring Barley (Hordeum Vulgare L.) Cross. *Genome = Génome, 38*(3), 575-585.

Li, X., Chen, W., Zhao, Y., Xiang, Y., Jiang, H., Zhu, S., & Cheng, B. (2013). Downregulation of caffeoyl-CoA O-methyltransferase (CCoAOMT) by RNA interference leads to reduced lignin production in maize straw. *Genetics and Molecular Biology, 36*, 540-546. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3873186/pdf/gmb-36-540.pdf

Li, J., et al. (2018). "Genome-Wide Association Studies for Five Forage Quality-Related Traits in Sorghum (Sorghum bicolor L.)." *Front Plant Sci*, 9: 1146.

Lin, S., et al. (2020). "Identification of genetic loci associated with forage quality in response to water deficit in autotetraploid alfalfa (Medicago sativa L.)." *BMC Plant Biol*, 20(1): 303.

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics, 28*(18), 2397-2399. doi:10.1093/bioinformatics/bts444

Liu, X. P. and L. X. Yu (2017). "Genome-Wide Association Mapping of Loci Associated with Plant Growth and Forage Production under Salt Stress in Alfalfa (Medicago sativa L.)." *Front Plant Sci*, 8: 853.

Lopez-Malvar, A., et al. (2019). "Genome-wide association analysis for maize stem Cell Wall-bound Hydroxycinnamates." *BMC Plant Biol*, 19(1): 519.

Lubberstedt, T. (2007). Application of genomics to forage crop breeding for quality traits. In R. K. Varshney & R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement: Vol 2: Genomics Applications in Crops* (pp. 281–306): Springer Science & Business Media.

Lundqvist, U. (2009). Eighty years of Scandinavian barley mutation genetics and breeding. In S. QY (Ed.), *Induced mutations in the genomics era* (pp. 39-43). Rome: FAO.

Maas, E. V. (1993, 1-4-1993). *Testing Crops for Salinity Tolerance.* Paper presented at the Workshop on Adaptation of Plants to Soil Stresses, Lincoln, NE.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2019). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.0.

Marita, J. M., Ralph, J., Hatfield, R. D., Guo, D., Chen, F., & Dixon, R. A. (2003). Structural and compositional modifications in lignin of transgenic alfalfa down-regulated in caffeic acid 3-O-methyltransferase and caffeoyl coenzyme A 3-O-methyltransferase. *Phytochemistry, 62*(1), 53-65. doi:https://doi.org/10.1016/S0031-9422(02)00434-X

Marquez-Cedillo, L. A., Hayes, P., Kleinhofs, A., Legge, W. G., Rossnagel, B. G., Sato, K., Wesennberg, D. M. (2001). QTL analysis of agronomic traits in barley based on the double haploid progeny of two elite North American varieties representing different germplasm groups. *Theoretical & Applied Genetics, 103*, 625–637. doi:10.1007/PL00002919

Martin, G. C., Shenk, J. S., & Barton, F. E. I. (1989). *Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality*.

Mertens, D. R. (1992). Critical conditions in determining detergent fiber. In *Proceedings of NFTA Forage Analysis Workshop* (15th ed ed., pp. C1-C8). Denver, CO.

Mertens, D. R., & McCaslin, M. (2008). Evaluation of alfalfa hays with down-regulated lignin biosynthesis. *Journal of Dairy Science, 91*.

Mesfin, A., Smith, K. P., Dill-Macky, R., Evans, C. K., Waugh, R., Gustus, C. D., & Muehlbauer, G. J. (2003). Quantitative trait loci for Fusarium head blight resistance in barley detected in a two-rowed by six-rowed population. *Crop Science, 43*, 307–318. doi:10.2135/cropsci2003.0307

Miedes, E., Vanholme, R., Boerjan, W., & Molina, A. (2014). The role of the secondary cell wall in plant resistance to pathogens. *Frontiers in Plant Science, 5*, 358-358. doi:10.3389/fpls.2014.00358

Mohammed, A. S., J. M. Anderson, L. M. Safky, and K. M. Barth. (1967) The Importance of Good Quality in Alfalfa Hay. In*: Vol. 61. Tennessee Farm and Home Science Progress Report* (pp. 10-13). Knoxville: Agricultural Experiment Station, University of Tennessee: University of Tennessee, Knoxville., & University of Tennessee.

Molina-Cano, J., Fra-Mon, P., Salcedo, G., Aragoncillo, C., De Togores, F. R., & García-Olmedo, F. (1987). Morocco as a possible domestication center for barley: biochemical and agromorphological evidence. *Theoretical and Applied Genetics, 73*(4), 531-536.

Molina-Cano, J., Igartua, E., Casas, A., & Moralejo, M. (2002). New views on the origin of cultivated barley. *Barley science: Recent advances from molecular biology to agronomy of yield and quality*, 15-29.

Montana Salinity Control Association. (2017). *Montana Saline Seep Reclamaiton Project*. Natural Resources Conservation Sevice Retrieved from https://www.nrcs.usda.gov/wps/portal/nrcs/mt/programs/farmbill/rcpp/9aa43aef-a460-4a07-ab1f-ab6ff135be69/

Moreno-González, J., & Cubero, J. I. (1993). Selection strategies and choice of breeding methods. In M. D. Hayward, N. O. Bosemark, I. Romagosa, & M. Cerezo (Eds.), *Plant Breeding: Principles and prospects* (pp. 281-313). Dordrecht: Springer Netherlands.

Munoz-Amatriain, M., Cuesta-Marcos, A., Endelman, J. B., Comadran, J., Bonman, J. M., Bockelman, H. E., Muehlbauer, G. J. (2014). The USDA barley core collection: genetic diversity, population structure, and potential for genome-wide association studies. *PLoS One, 9*(4), e94688. doi:10.1371/journal.pone.0094688

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R(2) and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J R Soc Interface, 14*(134). doi:10.1098/rsif.2017.0213

Owens, V. N., Albrecht, K. A., & Hintz, R. W. (1995). A Rapid Method for Predicting Alfalfa Quality in the Field. *Journal of Production Agriculture, 8*(4), 491-495. doi:10.2134/jpa1995.0491

Pan, A., Hayes, P. M., Chen, F., Chen, T. H. H., Blake, T., Wright, S., Bedo, Z. (1994). Genetic analysis of the components of winterhardiness in barley (Hordeum vulgare L.). *Theoretical & Applied Genetics, 89*(900–910).

Parsons, D., Cherney, J. H., & Gauch, H. G. (2006). Alfalfa Fiber Estimation in Mixed Stands and Its Relationship to Plant Morphology. *Crop Science, 46*(6), 2446-2452. doi:10.2135/cropsci2006.03.0144

Parsons, D., Peterson, P. R., & Cherney, J. H. (2013). Estimation of Nutritive Value of Spring Alfalfa–Grass Mixtures using In-Field Measurements and Growing Degree Data. *Forage and Grazinglands, 11*(1). doi:10.1094/fg-2012-0162-rs

Pauli, D., Muehlbauer, G. J., Smith, K. P., Cooper, B., Hole, D., Obert, D. E., Blake, T. K. (2014). Association Mapping of Agronomic QTLs in US Spring Barley Breeding Germplasm. *Plant Genome, 7*(3). doi:10.3835/plantgenome2013.11.0037

Pillen, K., Zacharias, A., & Leon, J. (2003). Advanced backcross QTL analysis in barley (Hordeum vulgare L.). *Theoretical & Applied Genetics, 107*(2), 340–352. doi:10.1007/s00122-003-1253-9

R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from https://www.R-project.org/

Roberts, E. H., Summerfield, R. J., Cooper, J. P., & Ellis, R. H. (1988). Environmental Control of Flowering in Barley (Hordeum vulgare L.). I. Photoperiod Limits to Long-day Responses, Photoperiod-insensitive Phases and Effects of Low-temperature and Short-day Vernalization. *Annals of Botany, 62*(2), 127-144. doi:10.1093/oxfordjournals.aob.a087644

Rollins, J. A., Drosse, B., Mulki, M. A., Grando, S., Baum, M., Singh, M., von Korff, M. (2013). Variation at the vernalisation genes Vrn-H1 and Vrn-H2 determines growth and yield stability in barley (Hordeum vulgare) grown under dryland conditions in Syria. *Theor Appl Genet, 126*(11), 2803-2824. doi:10.1007/s00122-013-2173-y

RStudio Team. (2020). RStudio: Integrated Development for R. Boston, MA: RStudio, Inc. Retrieved from http://www.rstudio.com/

Rubin, E. M. (2008). Genomics of cellulosic biofuels. *Nature, 454*(7206), 841-845. doi:10.1038/nature07190

Sakiroglu, M. and E. C. Brummer (2017). "Identification of loci controlling forage yield and nutritive value in diploid alfalfa using GBS-GWAS." *Theor Appl Genet*, 130(2): 261-268

Sanderson, M. A. (1992). Predictors of Alfalfa Forage Quality: Validation with Field Data. *Crop Science, 32*(1), cropsci1992.0011183X003200010049x. doi:10.2135/cropsci1992.0011183X003200010049x

Santillano-Cázares, J., Carmona-Victoria, M., Contreras-Govea, F. E., Idowu, O. J., & Albrecht, K. A. (2014). Applicability of Predictive Equations for Alfalfa Quality to Southwestern United States and Northern Mexico. *Crop Science, 54*(6), 2880-2886. doi:10.2135/cropsci2014.01.0075

Sarkar, D. (2008). Lattice: Multivariate Data Visualization with R. In. New York: Springer.

Sattler, S. E., Funnell-Harris, D. L., & Pedersen, J. F. (2010). Brown midrib mutations and their importance to the utilization of maize, sorghum, and pearl millet lignocellulosic tissues. *Plant Science, 178*(3), 229-238. doi:10.1016/j.plantsci.2010.01.001

Sattler, S. E., Palmer, N. A., Saballos, A., Greene, A. M., Xin, Z., Sarath, G., Pedersen, J. F. (2012). Identification and Characterization of Four Missense Mutations in Brown midrib 12 (Bmr12), the Caffeic O-Methyltranferase (COMT) of Sorghum. *BioEnergy Research, 5*(4), 855-865. doi:10.1007/s12155-012-9197-z

Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences, 106*(37), 15594. doi:10.1073/pnas.0906865106

Siahsar, B. A., Taleei, A. R., Peighambari, S. A., Naghavi, M. R., & Ghanbari, A. (2009). QTL analysis of forage quantity and quality-related traits of Steptoe x Morex doubled haploid lines of barley. *New Biotechnology, 25,* S303-S304. doi:10.1016/j.nbt.2009.06.760

Slafer, G. A., & Rawson, H. M. (1994). Sensitivity of wheat phasic development to major environmental factors: a re-examination of some assumptions made by physiologists and modellers. *Australian Journal of Plant Physiology, 21*, 393-426. Retrieved from http://hdl.handle.net/102.100.100/239426?index=1

Spaner, D., Rossnagel, B. G., Legge, W. G., Scoles, G. J., Eckstein, P. E., Penner, G. A., Mather, D. (1999). Verification of a quantitative trait locus affecting agronomic traits in two-row barley. *Crop Science, 39*, 248–252. doi:10.2135/cropsci1999.0011183X003900010038x

Steinfeld, H., & Otte, J. (2003). World agriculture: towards 2015/2030. An FAO perspective. In H. d. H. Jelle Bruinsma, and Paul Harrison (Ed.), *World agriculture: towards 2015/2030. An FAO perspective* (pp. 158-176). Retrieved from http://www.fao.org/3/y4252e/y4252e05.pdf

Stephens, J., & Halpin, C. (2008). *Barley 'orange lemma' is a mutant in the CAD gene*.

Steppuhn, H., & Raney, J. P. (2005). Emergence, height, and yield of canola and barley grown in saline root zones. *Canadian Journal of Plant Science, 85*(4), 815-827. doi:10.4141/P04-199

Sulc, R. M., Albrecht, K. A., Cherney, J. H., Hall, M. H., Mueller, S. C., & Orloff, S. B. (1997). Field Testing a Rapid Method for Estimating Alfalfa Quality. *Agronomy Journal, 89*(6), 952-957. doi:10.2134/agronj1997.00021962008900060017x

Surber, L., Abdel-Haleem, H., Martin, J., Hensleigh, P., Cash, D., Bowman, J., & Blake, T. (2011). Mapping quantitative trait loci controlling variation in forage quality traits in barley. *Molecular Breeding, 28*(2), 189-200. doi:10.1007/s11032-010-9473-6

Taleei, A., Siahsar, B. A., & Peighambari, S. A. (2009). QTL Mapping for Forage Quality-Related Traits in Barley. *Bio-Science and Bio-Technology, 57*, 53-+. Retrieved from <Go to ISI>://WOS:000279574000008

Tinker, N. A., Mather, D. E., Blake, T. K., Briggs, K. G., Choo, T. M., Dahleen, L., . . . Kasha. (1995). Regions of the genome that affect agronomic performance in two-row barley. *Crop Science, 36*, 1053–1062. doi:10.2135/cropsci1996.0011183X003600040040x

Tsuberi, A., Nashef, K., Reddy, B., Ben-Meir, Y., J., M., Chen, Y., . . . Ben-David, R. (2016). Comparative analysis of winter cereals for animal feed in Mediterranean environments. *Nir VaTelem, 67*, 16-21.

Turner, A. (2005). The Pseudo-Response Regulator Ppd-H1 Provides Adaptation to Photoperiod in Barley. *Science, 310*(5750), 1031-1034. doi:10.1126/science.1117619

USDA National Agricultural Statistics Service. (2020). *NASS - Quick Stats*. https://quickstats.nass.usda.gov/. Retrieved from: https://data.nal.usda.gov/dataset/nass-quick-stats

Van Soest, P. J. (1996). *Environment and forage quality*. Paper presented at the 58th Cornell Nutrition Conference, Ithaca, NY.

Van Soest, P. J., Robertson, J. B., & Lewis, B. A. (1991). Methods for dietary fiber, neutral detergent fiber and on-starch polysaccharides in relation to animal nutrition. *Journal of Dairy Science, 74*, 3583-3597.

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science, 91*, 4414–4423.

Vignols, F. (1995). The brown midrib3 (bm3) Mutation in Maize Occurs in the Gene Encoding Caffeic Acid O-Methyltransferase. *THE PLANT CELL ONLINE, 7*(4), 407-416. doi:10.1105/tpc.7.4.407

von Bothmer, R., Jacobsen, v., N, Baden, C., Jorgensen, R., & Linde-Laursen, I. (1995). *An ecogeographical study of the genus Hordeum*.

Von Bothmer, R., & Komatsuda, T. (2011). Barley origin and related species. *Barley: Production, Improvement and Uses*, 14-62.

Walesiak, M., & Dudek, A. (2017). clusterSim: Searching for Optimal

Clustering Procedure for a Data Set. R package version 0.45-2,. Retrieved from http://CRAN.R-project.org/package=clusterSim

Wang, H., Smith, K. P., Combs, E., Blake, T., Horsley, R. D., & Muehlbauer, G. J. (2012). Effect of population size and unbalanced data sets on QTL detection using genome-wide association mapping in barley breeding germplasm. *Theoretical and Applied Genetics, 124*(1), 111-124. doi:10.1007/s00122-011-1691-8

Wang, Y., Ren, X., Sun, D., & Sun, G. (2015). Origin of worldwide cultivated barley revealed by NAM-1 gene and grain protein content. *Front Plant Sci, 6*, 803. doi:10.3389/fpls.2015.00803

Wang, H., et al. (2016). "Genome-wide association analysis of forage quality in maize mature stalk." *BMC Plant Biol*, 16(1): 227.

Ward, C. E. (2000). *Marketing and Pricing Alfalfa Hay*. Oklahoma State University

Whitechurch, E. M., Slafer, G. A., & Miralles, D. J. (2007). Variability in the Duration of Stem Elongation in Wheat and Barley Genotypes. *Journal of Agronomy and Crop Science, 193*(2), 138-145. doi:10.1111/j.1439-037X.2007.00260.x

Wood, S., Seguin, P., Tremblay, G. F., Bélanger, G., Lajeunesse, J., Martel, H., . . . Claessens, A. (2018). Validation of Predictive Equations of Pre-Harvest Forage Nutritive Value for Alfalfa-Grass Mixtures. *Agronomy Journal, 110*(3), 950-960. doi:10.2134/agronj2017.09.0542

Wood, S., Seguin, P., Tremblay, G. F., Bélanger, G., Lajeunesse, J., Martel, H., . . . Claessens, A. (2019). Predicting Pre-harvest Forage Nutritive Value of Spring and Summer Growth of Alfalfa–Grass Mixtures. *Agronomy Journal, 111*(6), 3172-3181. doi:10.2134/agronj2019.03.0199

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., . . . Dubcovsky, J. (2006). The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proceedings of the National Academy of Sciences, 103*(51), 19581-19586. doi:10.1073/pnas.0607142103

Yu, L. X. (2017). "Identification of Single-Nucleotide Polymorphic Loci Associated with Biomass Yield under Water Deficit in Alfalfa (Medicago sativa L.) Using Genome-Wide Sequencing and Association Mapping." *Front Plant Sci,* 8: 1152.

Zhao, Q., & Dixon, R. A. (2011). Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends in Plant Science, 16*(4), 227-233. doi:https://doi.org/10.1016/j.tplants.2010.12.005

Zhao, J., et al. (2016). "Genome-Wide Association Study for Nine Plant Architecture Traits in Sorghum." *Plant Genome,* 9(2).