DISSERTATION

BREEDING HARD WINTER WHEAT (*Triticum aestivum* L.) FOR HIGH GRAIN YIELD AND HIGH GRAIN PROTEIN CONCENTRATION

Submitted by

Susan Patricia Latshaw

Department of Soil and Crop Sciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2021

Doctoral Committee:

Advisor: Scott Haley

Jesse Poland
Milt Thomas
Merle Vigil

ABSTRACT

BREEDING HARD WINTER WHEAT (*Triticum aestivum* L.) FOR HIGH GRAIN YIELD AND HIGH GRAIN PROTEIN CONCENTRATION

High grain yield (GY) is the primary selection target in commercial hard winter wheat (*Triticum aestivum* L.) breeding programs, with milling and bread-making quality as important secondary selection targets. Grain protein concentration (GPRO) is strongly correlated with important dough rheology and bread-making characteristics. Simultaneous improvement is difficult given the strong negative relationship of GY and GPRO in cereal crops. Nitrogen use efficiency (NUE), defined as the amount of grain produced per unit of N supply, promotes high GY through the component traits N uptake (NUpE) and N utilization (NUtE) efficiencies. Grain protein accumulation relies on N uptake from the soil and remobilization from plant tissue reserves. One study was conducted to characterize variation for NUE among a set of 20 breeding lines and varieties adapted to the west central Great Plains of the United States. Path analysis was applied to characterize the NUE component structure during the 2010-2011 growing season and then for two newly released varieties in the 2011-2012 growing season. Nitrogen use efficiency ranged from 39.9 kg kg^{-1} for 'RonL' to 46.7 kg kg^{-1} for 'Byrd'. By path analysis, we determined that variation in NUE depended on NUpE under N sufficiency and on NUtE under limiting N. Additionally, strategies for simultaneous improvement of GY and GPRO were explored. Analysis of standardized residuals of the linear regression of GPRO on GY, or 'grain protein deviation', identified one cultivar ('Brawl CL Plus') that had 6.7 g kg^{-1} higher GPRO than the average for all 20 genotypes. In a second study, selection strategies based on protein-yield

selection indices for a set of 775 breeding lines and varieties representing the Colorado State University hard winter wheat breeding program were evaluated based on field data obtained during the 2012-2015 growing seasons. Selection based on high values for a particular index delivered a characteristic emphasis on GY or GPRO. Correlation analysis between index values and GY or GPRO showed that each simultaneous selection strategy focused to differing extents on the primary traits. Genomic selection applied to index values in univariate models provided forward prediction accuracy ranging between r = .21 to .44 for the 2013 validation set, but approached zero for the 2014 validation set. Index values were also calculated from genomic estimated breeding values obtained in bivariate genomic selection models. Prediction accuracy for individual trait values was not substantially improved in the bivariate model. Protein-yield indices calculated from bivariate genomic estimated breeding values showed similar relationships to GY and GPRO as for the genomic estimated breeding values for indices calculated in the univariate models. A set of selection strategies generate sufficient predictive ability in phenotypic or genomic selection to be effective tools for simultaneous selection for GY and GPRO.

ACKNOWLEDGEMENTS

God gave me protection, courage and direction on the challenging path to completing my degree. I am indebted to the many people who loved and helped me, gave me grace, and invested time and money. I have unending appreciation for the seamless, skilled teamwork provided by Emily Hudson-Arns, Tori Anderson, Scott Seifert and John Stromberger during planting, management, and harvesting of the field trials for the Colorado State University Wheat Breeding Program. I am grateful to David J. Poss and Linda Hardesty (USDA-ARS) for field management at the USDA Research Station at Akron, CO, for grinding countless plots of wheat plants and for performing the C/N analyses. Assistance with statistical analysis of the N use study was provided by Dr. Philip Chapman (CSU). Tawney Campbell, James Hamilton, Donald (Collin) Dutro and Erin Krause assisted with sample processing, data collection and field support (and provided much merriment). I hold precious memories of my classmates Annie, Anna, Paul, Asma, Carolyn, Wahid, Jessica, Tori, Leon, Craig, Sarah, Hung, Steve, Melaku, Erena, Garrett, Rich, Salem and Mohammed while we shared the fun, struggle, debate and celebration that were part of life in the basement office.

My deepest appreciation to Martha L. Crouch, Brian Staskawicz, Robert Boswell, Donald R. McCarty, William G. Farmerie, and Robert E. Stall for providing training, knowledge and support to foster my development as a research scientist. I am indebted to Sal Edwards for coaching me through the endless last few miles. Her faithful determination kept my fires burning. Special appreciation to Judy Harrington for her keen editorial eye and sage advice during preparation of the dissertation.

I am indebted to Scott Haley for his steadfast faith in me and for the doors he opened through the training and financial support he provided. I am fortunate to know him and to have witnessed his passion for serving Colorado growers and for feeding the world by breeding better wheat. I treasure memories of following him through the headrows--seeing each one through his 'breeder's eye', all while he articulated an encyclopedia of wheat breeding history. I am grateful for the encouragement and guidance during the steps to my degree from Milt Thomas and Jesse Poland. Merle Vigil never failed to turn my eyes to the bright light of hope through his gifts of wisdom and encouragement. Byrd Curtis connected my present work to the paths that he and his generation of breeders trod into the wheat fields of the modern world. He always reminded me of the value of our work. I will miss his presence in this world and look forward to reunion in the one to come.

My mother and father implanted a desire to seek challenges, pursue knowledge, and to expect success. They taught me to climb the ladder of education to transform myself and our world. I dedicated this work to my son and I finished it to make good on the debt of time and attention owed him.

Funding was provided by the USDA-ARS Central Plains Resources Management Research Unit, Akron Colorado, under National Program 212, Soils and Global Climate change CRIS Project number 3010-12210-002-00D.

TABLE OF CONTENTS

CHAPTER 1

LITERATURE REVIEW

Introduction

Human global population is projected to reach 9.8 billion people by the year 2050, with most growth occurring in Asia and Africa (UNDESA, 2017). The global population is projected to be wealthier and thus demand more calories from animal products. With a basis of the 2010 agronomic practices and crop yields, that demand will create a food supply gap equivalent to 56% more crop calories, will require 593 million ha of additional agricultural land area, and will emit 11 Gt more greenhouse gases (Searchinger et al., 2019). Given these projections, the equivalent food demand of the 2050 human population will be that of 12.5 billion at current consumption levels (Baenziger et al., 2017). The three major cereal crops, rice (*Oryza sativa*), wheat (*Triticum aestivum* L.) and maize (*Zea mays* L.), provide 44.8% of the calories required for global populations (FAO, 2019).

The grand challenge for food security in the 20th century was the 'War on Hunger' which was fought with considerable success by developing and distributing improved cereal grain varieties to food insecure nations and by promoting modern agronomic practices to double or triple global food production during the 'Green Revolution' (Wharton, 1969). Concomitantly, there was a reduction of hunger from levels estimated at 50% of the global population in 1968 (Tillman, 1968) to 29% during 1979-1981 and to 18% during 1995-1997 (Donmez et al., 2001; FAO, 2004; Pingali, 2012). This achievement came through improved agronomic practices such

as nitrogen (N) fertilizer application and the distribution of new cereal grain varieties bred for N responsiveness for yield coupled with reduced height to prevent lodging (Donmez et al., 2001; Reitz & Salmon, 1968). High N input production systems push productivity of varieties with high yield potential (Graybosch et al., 2014; Lollato & Edwards, 2015), but this comes with risk of environmental degradation through N escape (Cameron et al., 2013). New priorities for the 21st century push food quality for human nutrition as a key part of food security (Baenziger et al., 2017).

Climate change mitigation efforts within the agricultural sector are needed as we strive to meet the caloric needs of a growing population. Nitrous oxide (NO_2) is a potent greenhouse gas with 8% of all US emissions attributed to fertilized agricultural fields (Millar et al., 2014). The future production goals for the wheat crop must be achieved without increasing emissions of greenhouse gases in the cropping system. While increased atmospheric carbon dioxide (CO_2) levels have been observed to increase cereal grain yield potential through increased photosynthetic rates and carbon translocation to the grain (Tester & Langridge, 2010), climate change also brings production risk through more frequent climate extremes which disrupt crop productivity (Liu et al., 2016; Reynolds et al., 2016). Global climate change is already changing production patterns of the top 10 food crops, with some regions showing marked declines, while others have increased productivity (Ray et al., 2019).

The grand challenge plant breeders face today is to develop new crop varieties which meet the 2050 production and food quality requirements, while targeting future production environments. Predictions of future production environments include factors such as current climate change mitigation efforts and continuing trends towards loss of arable crop land (Davidson et al., 2015; Reynolds et al., 2016; Tester & Langridge, 2010). Globally, wheat is

second only to rice for its contribution to daily calories (16% in the developing world and 26% in developed nations) and demand is projected to increase from 760 million tonnes in 2020 to 900 million tonnes in 2050 (Dixon et al., 2009). Wheat is a global trade commodity, with 150 million tonnes traded on an annual basis (Shewry & Hey, 2015). U.S. growers contributed 7.7% on average to global wheat production and 16.9% of the export market from 2010-2019 (USDA-ERS, 2019).

Nitrogen Use Efficiency

The 'Green Revolution' cereal grain production system coupled increased N fertilizer use with wide distribution of N responsive, lodging resistant short-statured varieties to meet food, feed, and fuel needs of the modern era. Nitrogen is an essential macronutrient for crop production, with a cost and revenue balance that encourages some producers in the developed world to apply it in excess, to insure against lost opportunities for optimal yields. On the other hand, in developing countries, or in other low-input production environments, N fertilizer may be the highest grower input cost, which may discourage optimal N fertilization for producing a quality crop (Tester & Langridge, 2010). Nitrogen losses from agroecosystems occur through N volatilization from the soil, biological denitrification, release of greenhouse gas forms of N oxides, nitrate leaching into water, and ammonia loss from plant leaves through photorespiration (Omara et al., 2019). Such losses reduce profitability and carry societal costs due to environmental degradation. Nitrogen use efficiency (NUE) is calculated as a ratio of either N harvested in the grain (GNY) or grain yield (GY) per unit of N supply (Ns) (Van Sanford & MacKown, 1986). Calculated values may account for soil residual N and other environmental inputs. When calculated as fertilizer recovery efficiency (GNY/Ns; (Hawkesford & Griffiths,

2019), where GNY is adjusted for environmental N sources and the Ns is applied N fertilizer, global NUE during cereal grain production was estimated to be 33% (Raun & Johnson, 1999). Production system changes which may improve NUE include optimizing crop rotations within an agro-environment, placement and timing of fertilizer application, choice of N fertilizer form, tillage methods, optimized timing and amount of irrigation, adoption of precision fertilizer application methods to account for within field N variability, and planting cultivars with superior NUE (Raun & Johnson, 1999).

Life cycle assessment in a wheat-to-bread supply chain was applied to assess the environmental impact of producing a loaf of bread in the U.K. during 2014 (Goucher et al., 2017). Surprisingly, 65% of the global warming potential of the loaf of bread was attributed to production of the wheat crop, with fertilizer accounting for 47% of the process load. In this highly productive environment, the NUE, reported as the ratio of harvested N to applied N, was estimated to be 71%. Commonly, the economic benefit to the grower of applying excess fertilizer to insure against lost yield potential is not offset by external pressures which might reduce application overages. The authors propose to incentivize responsible fertilizer use by integrating decisions across all stakeholders, including the consumers. An example of this sort of integrated policy was partially successful in resolving nitrate groundwater contamination from maize production in the Platte River Valley of Nebraska (Davidson et al., 2015; Ferguson, 2015). In this region, similar to most high production regions, rate of N fertilizer application increased linearly from 1955 to 1975. At the time that groundwater contamination was recognized in the early 1980s, nitrate levels were 30 to 40 mg NO_3-N L^{-1}. Mitigation policies were put in place and levels dropped to 10 mg NO_3-N L^{-1} by 2015, marking a large step forward towards reduction of N losses within this highly productive region. As an outcome of the integrated management

policy, Nebraska maize growers subsequently doubled grain yields with no significant increase in amount of applied N fertilizer through the adoption of grower education and incentive programs, improved N management practices and better hybrid varieties. Similarly, crop yield improvements with stable or decreasing N fertilizer levels were reported in a number of European countries after environmental policy changes were initiated in the 1980s (Lassaletta et al., 2014).

Recent re-analysis puts global NUE across all crops at 47% (Lassaletta et al., 2014), and 35% for cereal grain production (GNY/Ns) (Omara et al., 2019). In addition to an overall analysis, results on a country-by-country basis demonstrated trend categories for the 124 nations in the Lassaletta et al. (2014) study. Several sub-Saharan nations projected to have the greatest population gains in the next half-century show NUE trends indicative of N mining, where a low input cropping system results in low yields and depletion of soil fertility. In these N-limited systems, NUE approaches 100% as an effect of low N fertilization rates (Lassaletta et al., 2014). A group of highly productive regions, including the U.S. and Brazil, showed steady increases in N fertilizer use and yields until the 1980's, followed by a different trend line with increased yields per unit of N input and either stable or higher NUE. In the U.S., the trend change occurred at the time that fertilizer inputs were reduced and other agronomic changes continued to support yield improvements. Another set of highly productive regions, including China and India, show a steady rate of increased N input and yield response to N, followed by a flattening of the response, representing decreasing NUE. Here, the decreasing marginal yield benefit of additional units of N reflects excess fertilizer application.

A single solution to the NUE puzzle will not apply across the diversity of global agro-ecosystems. A method to find an optimum between conflicting factors, such as agricultural

productivity and environmental health, is to define the 'trade-off frontier'. Building consensus to correct agricultural N imbalances may start by parameterizing a N yield response curve. Outside of the frontier bounds, excess N causes environmental degradation or, within bounds, inadequate N fails to optimize crop production. The "sweet spot" of sustainable policy and practice may emerge within agroproduction systems through this methodology (Mueller et al., 2014).

Genetics of Nitrogen Use Efficiency in Wheat

Nitrogen use efficiency in wheat is conditioned by physiological processes which integrate N absorption, translocation, assimilation, and remobilization processes with photosynthesis, carbon (C) assimilation and remobilization during wheat plant growth and development (Figure 1.1). It is measured as the ratio of GY produced per Ns, with further consideration of component traits (Hawkesford & Griffiths, 2019; Moll et al., 1982). Definitions for NUE-related traits are listed in Table 1.1. During the vegetative phase of growth, N uptake by the roots and translocation to the growing shoots and roots is followed by assimilation via nitrate reduction and conversion into amino acids. During the vegetative phase, amino acids are incorporated into proteins to build the plant architecture and photosynthetic complexes. These processes determine the N uptake efficiency (NUpE), defined as efficiency of the accumulation of N in the shoot biomass per unit of Ns. During the reproductive phase, photosynthetic and N uptake processes continue, while grain-filling requires remobilization of assimilates to the developing seeds and senescence of vegetative tissues. These processes determine the N utilization efficiency (NUtE) for GY production, measured as GY per unit of N accumulated in the shoot biomass. As ratios, the component traits have a product relationship with the resultant trait: NUE = NUpE * NUtE.

The relative genetic contributions of NUE component traits vary by crop species, genotype, agronomic management and other environmental conditions (Barraclough et al., 2010; Wang et al., 2011). In general, two methods are applied to characterize the relative importance of component trait contributions to variation for NUE. The degree of association is estimated by correlation and the degree of relationship by linear regression (Barraclough et al., 2010; de Oliveira Silva et al., 2020; Gaju et al., 2011; Guttieri et al., 2017; Kubota et al., 2018; Wang et al., 2011). An extension of correlation analysis was developed by Moll et al. (1982) and applies the product relationship of the component traits and their covariances to derive fractional contributions to variation for NUE (Le Gouis et al., 2000; Ortiz-Monasterio R. et al., 1997; Van Sanford & MacKown, 1986).

Variation for NUE may be determined by variation in NUpE (An et al., 2006; Brasier et al., 2020; Dhugga & Waines, 1989; Sadras & Lawson, 2013; Wang et al., 2011), or predominantly by variation in NUtE (Barraclough et al., 2010; Muurinen et al., 2006). Breeding progress for NUE within individual breeding programs reflects the history of selection pressure on the component traits. In a study of productivity trends for variety releases from 1958 to 2007, the rate of increase in N taken up by the crop (0.40 kg N ha^{-1} yr^{-1}) paralleled GY trends (18 kg ha^{-1} yr^{-1}), suggesting that breeding for GY applied indirect selection for N uptake (Sadras & Lawson, 2013). Similarly, for irrigated spring wheat grown in California and Australia, cultivars differed in NUpE at non-limiting N levels while genetic variation for NUpE explained most of the variance for NUE (Dhugga & Waines, 1989; Sadras & Lawson, 2013). Nitrogen utilization efficiency has a product relationship with component traits harvest index (HI) and biomass production efficiency (kg total dry weight (TDW) kg^{-1} N). In Finland, an environment with a breeding history of selection under high N, the genetic variance for NUtE, through its component trait HI, determined variation in NUE within an historic set of spring wheat cultivars (Muurinen et al., 2006). Moll et al. (1982) suggested that at

moderate N rates, selection for NUE would apply selection pressure for both NUpE and NUtE. Consistent with this proposal, a set of Mexican spring wheat genotypes that were selected under moderate N application rates had both increased GY under limiting Ns and greater responsiveness to N fertilizer application. The researchers reported both increased NUpE and NUtE during the breeding history (Ortiz-Monasterio R. et al., 1997). These studies illustrate that it is important to assess the physiological basis of NUE as it relates to agronomic conditions and the genetic diversity of a breeding population in advance of establishing a breeding strategy.

The extensive literature on component trait contributions in wheat encompasses a diversity of agroproduction systems, but typically restricts the study to a small set of genotypes relevant to a particular region or breeding program. Longitudinal studies explored genetic gain for NUE in Europe and the Great Plains of the US. A panel of 225 elite European winter wheat cultivars (1985-2010 release years) was grown in a multi-environment trial (MET) under optimal and limiting Ns, with mean GY 7.4 Mg ha^{-1} (Cormier et al., 2013). They found genetic gain for GY was not significantly different at limiting N levels (LN) or with optimal N levels (HN) and was estimated to be 0.45% yr^{-1}, or 33.2 kg ha^{-1} yr^{-1}. For NUE, they observed genetic gain of 0.37% yr^{-1} at LN and 0.30% yr^{-1} at HN. Both NUpE and NUtE were correlated with NUE, but they did not detect significant year effect for NUpE, possibly due to limitations of the measurement methodologies. A significant year effect was detected for NUtE (0.20% yr^{-1}), demonstrating breeding progress for this component trait. In a similar study of a panel of 299 landraces, breeding lines, and cultivars (release dates 1874-2014), but grown in the lower yielding environment (mean GY 4.7 Mg ha^{-1}) of the Great Plains of the United States. Genotype-by-year interactions were significant due to typically variant weather patterns between growing seasons, so data were analyzed by year (Gutierri et al, 2017). Separately estimated for 2012 and 2013, genetic gain during 1960-2014 for GY was 0.331 and

0.761% yr^{-1}, for NUpE was 0.076 and 0.165% yr^{-1}, and for NUtE was 0.115 and 0.367% yr^{-1}.

Nitrogen is remobilized from the canopy to the grain for assimilation into structural and storage proteins. Selection for N responsiveness for GY can have the unintended effect of selection for reduced GPRO (Acreche & Slafer, 2009; Simmonds, 1995). There are several theories proposed for this phenomenon. Since GPRO can be considered to be the ratio of C to N in the grain, differing dynamics of remobilization result in the 'N dilution effect' (Desai & Bhatia, 1978). A global meta-analysis of previously published NUtE data was performed on 524 observations from 54 publications, representing all major wheat growing regions (de Oliveira Silva et al., 2020). Data were centered to remove environmental effects to focus on main effects of N variables. The summary statistics showed normal distributions for all variables, with a wide range of values: 11 Mg ha^{-1} GY, 250 kg ha^{-1} shoot biomass N (BMN), 57 kg GY kg^{-1} NUtE, and 85 g kg^{-1} grain protein concentration (GPRO). There is a linear and negative relationship for BMN and NUtE, with substantial variation for NUtE at a given BMN level. These relationships predict: 1) as GY improves through NUtE, GPRO will decline, and 2) variation at each BMN level may enable selection for varieties which deviate from the negative relationship of GPRO with GY. These results differ from those reported by Cormier (2013) where, even as GY increased due to improved NUtE, GPRO did not decrease over the span of surveyed years due to increased NHI.

Through spike trimming experiments, a linear negative relationship between density of seeds and GN was demonstrated for a given N uptake level (Acreche & Slafer, 2009). The authors hypothesized that N accumulation in the grain is source limited, while carbohydrate accumulation is driven by sink strength. Sink strength refers to the yield capacity, through yield components such as spikes per area and seeds per spike. It has been under selection by breeding for GY and would result in N dilution when a finite amount of BMN is available for remobilization to ever-increasing

numbers of seeds. The end-point of N dilution is the minimum level of GPRO needed to produce viable grains and the constraints of the protein requirements of the wheat market class (Gaju et al., 2011). Selection for greater NHI among genotypes with high NUE is a potential strategy to counteract N dilution.

Grain yield drives profits for bread wheat production due to commodity pricing basis, with protein premiums sometimes part of the equation. Grain protein concentration is associated with bread-making quality traits in wheat. It has long been reported that for cereal grain crops, GY and GPRO hold a strong negative association when compared across genotypes in a population (Simmonds, 1995). Proposed mechanisms to account for the negative correlation are the 'N dilution effect' (Acreche & Slafer, 2009), the C cost for N assimilation and translocation (Munier-Jolain & Salon, 2005), or the 'self-destruct' hypothesis where C fixation and N assimilation and translocation processes are in physiological opposition (Barraclough et al., 2010; Sinclair & de Wit, 1975).

Exceptional genotypes have been reported that have higher GPRO than expected at a given GY level (Bogard et al., 2010; Ehdaie & Waines, 2001; Fortunato et al., 2019; Guttieri et al., 2015; Marinciu & Saulescu, 2009; Monaghan et al., 2001; Oury & Godin, 2007; Rapp et al., 2018; Thorwarth et al., 2018; Cristobal Uauy et al., 2006). Genotypes with high grain protein deviation (GPD) achieve GPRO that exceeds the predicted value for a given level of GY (Monaghan et al., 2001). Processes linked to this trait include anthesis date (Bogard et al., 2011), N partitioning (Gaju et al., 2014; Papakosta & Gagianas, 1991), N assimilation dynamics (Fortunato et al., 2019), plasticity of biomass production under N-limitation (Rahimi Eichi et al., 2019), total biomass N accumulation (de Oliveira Silva et al., 2020; Desai & Bhatia, 1978), reproductive N sink strength (Dhugga & Waines, 1989), post-anthesis N uptake (Bogard et al., 2010) and rates of canopy

senescence or plant N losses (Noulas et al., 2013). Grain protein deviation has been proposed as an effective breeding strategy to select genotypes with adequate GPRO for end-use quality, independently of GY (Monaghan et al., 2001).

Just as HI is a measure of remobilization efficiency of photosynthate to the grain, NHI is a measure of the proportion of BMN that is remobilized to the grain. These traits are strongly associated, but are under differential control by genetic and environmental factors (Desai & Bhatia, 1978). The component traits of NHI have a product relationship: HI_{GN} (g GN kg^{-1} TDW)* BPE (kg TDW g^{-1} BMN) = NHI (%). The chloroplast-localized photosynthetic enzyme, Rubisco (ribulose 1,5 biphosphate carboxylase oxygenase), accounts for about 50% of total plant protein and more than 25% of total plant N, and thus is central for N management in the plant (Hirel et al., 2007). Retaining N in photosynthetically active tissues promotes N utilization efficiency for GY through continued carbon assimilation and translocation to the grain (Barraclough et al., 2010). The 'self-destruction' hypothesis predicts that under N-limiting conditions, increased remobilization of N from proteins in vegetative tissues leads to declining photosynthesis, increased rate of senescence, and a shortened grain-filling period (Sinclair & de Wit, 1975). Accordingly, Barraclough et al. (2010) emphasize that to both increase GY and maintain GPRO, NHI must increase, while maintaining a functional photosynthetic system. These authors suggest that accumulation and subsequent transfer of non-photosynthetic sources of N might explain the observed variation among genotypes for NHI. However, in a subsequent study of N pools in a set of elite genotypes, N was remobilized efficiently from all plant tissues, suggesting that all remobilized N was in fact metabolic and not structural (Barraclough et al., 2014). Genetic variation in timing and rate of senescence has been reported, as reviewed in Cormier et al. (2016). They propose that a 'supply/demand framework for N dynamics' underlies the variation under N-limiting conditions.

Nitrogen sink demand within a grain is met by stored N pools in the stems and rachis through N remobilization after anthesis, but under N insufficiency, the leaf N pool is remobilized, with associated accelerated senescence. Modification of the timing of senescence relative to remobilization processes could enhance GN while maintaining GY.

Nitrogen remobilization from BMN accounts for 60-95% of GN (Papakosta & Gagianas, 1991; Van Sanford & MacKown, 1986). Post-anthesis N uptake (PANU) contributes to GNY (Bogard et al., 2010), although this is not significant under conditions of low soil moisture (Kubota et al., 2018). Relationships of GPRO with the physiological traits, N remobilization efficiency (NRE) and PANU, differ among genotypes and N supply levels and have been proposed as selection targets for maintaining GPRO under enhanced GY (Bahrani et al., 2013; Gaju et al., 2014; Monaghan et al., 2001). In some agronomic environments, PANU contributes to NUtE for GY (Gaju et al., 2011) and to increased GPRO (Bogard et al., 2010; Monaghan et al., 2001). In those environments, PANU would be sufficient to support N translocation for GN while maintaining BMN for continued photosynthesis (Sinclair & de Wit, 1975). Loss of N through volatilization of ammonia from the canopy occurs, is reported as a negative value for PANU, and is impacted by genotype and environment (Vikas Belamkar et al., 2018). Optimizing PANU is an important breeding objective in environments with low frequency occurrences of drought conditions during the grain filling period.

Recent decades have seen an unfolding of the research imperative to identify candidate genes controlling NUE in order to gain knowledge of its inheritance and genetic architecture and to identify genes that may be targeted for breeding (Hirel et al., 2007). Through consideration of the physiological processes which determine plant growth and development (Figure 1.1), candidate genes were proposed and evaluated for their contribution to variation in NUE (for review, Bharati &

Mandal, 2020). Candidate genes which are found to have significant positive effects may be introduced by breeding into elite germplasm, or may be utilized in marker-trait association analysis to further elucidate gene networks contributing to NUE (Nigro et al., 2019).

An example of successful implementation of this approach is found in rice. Knock-out mutants of an Arabidopsis (*Arabidopsis thaliana* L.) transcription factor known as a NIN-like protein (*AtNLP7*) cause a N-starved phenotype and impaired nitrate signaling, while overexpression stimulates C and N assimilation and total biomass accumulation (Wu et al., 2020). NIN-like proteins regulate nitrate-inducible gene expression (Konishi & Yanagisawa, 2014; Mu & Luo, 2019; Wang et al., 2018). Gene orthologs were identified in rice, including the NIN-like protein (*OsNLP4*). It was found to be a global regulator of N-responsive genes in rice. When overexpressed, it effects a 47% increase in NUE under moderate Ns. Additionally, when *OsNLP4* is overexpressed in Arabidopsis, it rescues the function of a knock-out null mutation of *AtNLP7*. This gene has a major effect on NUE and is a promising candidate to validate for use in breeding rice. To deploy this and other major NUE-related genes for breeding wheat, wheat orthologs have been identified for a number of candidate genes (Bajgain et al., 2018; Balyan et al., 2016; Good & Beatty, 2011; Nadolska-Orczyk et al., 2017; Wang et al., 2018).

In rice, a difference in NUE between japonica and indica rice was associated with single nucleotide polymorphisms (SNP) in the nitrate-transporter gene, NRT1.1B (*OsNPF6.5*) and the gene functional was confirmed by transgenic studies of near isogenic indica lines (Hu et al., 2015). Sequence similarity with Arabidopsis and other cereal N transporter genes identified wheat homoeologs located on the A,B, and D genomes (Bajgain et al., 2018). A subset of homoeologs showed differential expression in seedling roots and shoots. Further work to understand related

allelic variation and its contribution to efficient N uptake and N transport will support breeding efforts to optimize NUE.

Nitrogen use efficiency and increased N accumulation in the grain has been linked to a number of N metabolism pathway genes. Two durum wheat (*Triticum turgidum* L. ssp. *durum*) cultivars that differed in the level of N accumulation in the grain showed differing patterns of nitrate reductase activity. The cultivar that accumulated higher grain N showed N-inducible nitrate reductase activity in the roots and leaves, decreased ammonium ion concentration in the roots, and increased nitrate concentration in leaves (Fortunato et al., 2019). Improved NUE has been demonstrated in wheat through transgenic introduction of an ABRE-binding factor (*TabZIP60*). This ABF-like leucine zipper transcription factor mediates N uptake and GY via interaction with the binding site in the promoter of NADH-GOGAT (J. Yang et al., 2019). Root-specific gene expression led to cloning of an N-inducible NAC transcription factor, *TaNAC2-5A* (He et al., 2015). It binds to the promoters for N transporters and glutamine synthetase and its overexpression increased N uptake, GN, NHI, and GY under field conditions. The trimeric Nuclear Factor Y (NF-Y) binds to the CCAAT box, a universal element of the eukaryotic promoter. Its expression is up-regulated under N limitation via down-regulation of miR169 in Arabidopsis. Increased N uptake, root biomass, and GY under field conditions was observed in wheat overexpressing *TaNFYA-B1* (Qu et al., 2015).

Transgenic wheat containing the maize transcription factor *ZmDof1* provides an example of mixed outcomes for transgenic NUE genes (Peña et al., 2017). Transgenes were either constitutively expressed by a ubiquinone promoter, or were under tissue-specific light regulation in leaf mesophyll cells and leaf sheaths by the *rbc*S1 promoter. Constitutive overexpression down-regulated photosynthesis resulting in decreased plant height, biomass and GY. Tissue-specific

overexpression promoted increased biomass and GY, with no significant reduction in GPRO. Good et al. (2007) reported on a significant role for alanine transferase in promoting NUE in wheat. When expression was controlled by a root-specific inducible promotor, alanine unexpectedly accumulated in the shoot, measurable as increased BMN. This dramatic result contrasts with numerous other studies of NUE candidate genes that contributed no observable phenotype via overexpression. The authors suggest that design of a successful transgenic approach to improved NUE will require detailed knowledge of end-products, potential feed-back regulation by metabolic products, and correct tissue specificity and timing of gene expression.

Breeding For Nitrogen Use Efficiency

Genetic progress for NUE under N limiting conditions generally occurs through indirect selection under N sufficiency within the main breeding nurseries. Genetic correlation under high and low N conditions supports success of indirect selection for NUE. Heritability in low N conditions may be reduced relative to N sufficient conditions through low genetic variance and high environmental variance for GY (Cormier et al., 2016). Genetic variation exists within elite breeding populations for NUE and its component traits, NUpE and NUtE (as reviewed in Balyan et al., 2016; Cormier et al., 2016; Guttieri et al., 2017). Selection for GY or NUE under contrasting Ns may capture variation in efficiencies of N utilization and uptake, separately, thus enabling breeding crosses to combine superior alleles for both component traits (Dhugga & Waines, 1989; Ortiz-Monasterio R. et al., 1997; Wang et al., 2011). A compromise between the resource demand of screening a breeding nursery at two or more Ns levels for most accurate ranking of NUE components and the masking of genetic variation for N uptake at optimal Ns may be to use moderate Ns in selection nurseries (Cormier et al., 2016).

Breeding progress has been investigated through comparative studies of cultivars released over time within a target region for GY (Battenfield et al., 2013; Maeoka et al., 2020; Rife et al., 2019; Sadras & Lawson, 2013) and for NUE (Cormier et al., 2013; Guarda et al., 2004; Kubota et al., 2018; Muurinen et al., 2006; Ortiz-Monasterio R. et al., 1997). The rate of progress is measured in comparison to the included variety with the earliest release date (selected reports are summarized in Table 1.2). Rates show either linear or curvilinear relationships with date of cultivar release. A period of rapid change through the 1980s during the adoption of N-responsive, semi-dwarf varieties was followed by a shift to decreased gains in recent years in several regions. Despite the range of estimated gains of only 0.4 to 1.1% during the modern era, substantial variation for GY exists among a collection of elite entries in the Southern Regional Performance Nursery (SRPN) (Battenfield et al., 2016; Rife et al., 2019). The reported gains are less than the required 2% gain per year to meet 2050 projected food demands (Tester & Langridge, 2010). Understanding and optimizing the underlying genetic contributions to NUE will to provide breeders effective strategies for accelerating genetic gain to meet this imperative.

Nitrogen use efficiency is a quantitative trait with polygenic inheritance (Hirel et al., 2007). Quantitative genetics methods and candidate gene approaches are employed to detect chromosomal regions which contribute to variation for complex traits (Bernardo, 2010). Quantitative trait loci (QTL) mapping is a linkage-based statistical method applied to bi-parental populations to identify bi-allelic loci that are significantly associated with phenotypic values. Similarly, candidate gene analysis applies co-segregation analysis to identify statistical associations of sequence variants in genes known to have functions related to the trait of interest. Genome-wide association studies are in wide use for exploring the genetic architecture underlying quantitative traits and for fine-mapping of QTL. Genome-wide association studies (GWAS) assay all haplotypes present in a population for

significant association with trait values, providing both effects estimation and QTL discovery (Hamblin et al., 2011). These methods contribute to understanding the genetic architecture of a quantitative trait through estimation of numbers of loci controlling traits and the relative contributions of the QTL. Fine mapping of QTL is possible when GWAS is applied in the context of an extensive history of recombination events captured by diverse germplasm collections (Bernardo, 2016). These types of studies identify markers linked to QTL that may deployed for breeding through trait introgression and marker-assisted selection.

A caveat for to consider prior to deployment in breeding is that effects of a QTL may be specific to a population or environment. Validation studies in relevant germplasm are essential prior to effective deployment in a breeding program. Additionally, the ability to detect QTL depends on a number of variables, including: the frequency distribution of alleles at causal loci, magnitude of the effect at each locus, population size and relatedness structure, quality of phenotyping data, and, by extension, trait heritability (Bernardo, 2008; Rafalski, 2010). An example of the impact of these factors on QTL detection power was illustrated in a maize bi-parental population for detecting plant height and GY QTL (Bernardo, 2010). Mapping populations from maize inbreds Mo17 and B73, with differing numbers of recombinant inbred lines were developed by two research groups and were tested in differing environments. The numbers and effects distributions of QTL were similar between the studies, but the map locations and the relative contributions to trait variances differed. One QTL was mapped at the same location in both studies. When a population was subdivided from 400 members to 100 members, the number of QTL detected was reduced, with counts depending on the random subset. As such, there was an upward biasing of estimated effects attributed to the detected QTL in the subsetted populations. An additional limitation of QTL applications in breeding is due to the polygenic nature of these traits. To obtain the desired phenotypic effect under an additive effect model, a number of QTL would

need to be combined within each selection candidate line. Stacking more than a few genes requires prohibitively large populations to obtain the desired recombinant.

Quantitative trait loci for nitrogen use efficiency

QTL mapping has been applied to identify chromosomal locations linked to genes that underlie NUE-related traits (An et al., 2006; Balyan et al., 2016; Cormier et al., 2014; Guttieri et al., 2017). In a GWAS applied to a panel of 214 European winter wheat elite varieties (release dates 1985-2010) under two N levels, 15 SNP associated with NUE were detected, with average effect of 8.7% , consistent with the expected polygenic nature of the trait (Cormier et al., 2014). Under an additive model for gene action, predicted values for each variety were calculated by summing effects for each QTL and then were regressed against adjusted means. Together, the predicted effects explained 55.7% of the genetic variation. More recently released varieties contained a higher percentage of favorable alleles, reflecting a breeding history with selection pressure for increased GY through improved NUtE and NHI (Cormier et al., 2013). The 2014 study included a co-localization network analysis that linked 28 traits based on the percentage of QTL in common between traits. This analysis reveals expected patterns of pleiotropic effects resulting from selection on QTL that co-localize among networked traits. Of interest was the lack of co-localized QTL for NUpE, likely due to its low genetic variance in this panel where the selection target was GY.

A panel of 299 winter wheat cultivars and breeding lines from Great Plains breeding programs was used for an association analysis for QTL contributing to NUE-related traits (Guttieri et al., 2017). Two stable QTL were found on chromosomes 4B and 2D that map near chromosome locations of a cytosolic glutamine synthetase (GS1) gene and plastic glutamine synthetase 2 (GS2) genes. In a candidate gene study of GS2 variants among Chinese winter wheats, particular gene variants were significantly associated with improved N uptake and GY (Li et al., 2011). Eleven

studies reviewed in Balyan et al. (2016) reported between 11 and 380 significant QTL were detected in each study for NUE and its component traits, explaining as much as 39% of variation for NUE. Association analysis with enzyme activity of NUE candidate genes was undertaken which identified significant QTL for GS1 (Fontaine et al., 2009; Habash et al., 2007), GS2, glutamate synthase (GOGAT), glutamate dehydrogenase (GDH) (Bordes et al., 2013), and NADH-GOGAT (Nigro et al., 2019). These studies have provided helpful insights into the metabolic pathways which may respond to selective pressure for NUE components, but they also confirm its highly polygenic nature and the need for gene pyramiding (Cormier et al., 2016).

QTL mapping in a Chinese bi-parental population under a range of N levels identified five QTL for total above-ground N (TN) that controlled 14-21.9% of variation for TN and were stable across N treatments (An et al., 2006). Plant and seedling biomass related traits showed positive correlation with TN. Several QTL detected for TN co-localized with plant biomass related QTLs. Traits positively associated with the five stable QTL included tillering, root dry weight, kernel number, shoot dry weight and seedling vigor, supporting the hypothesis that vigorous early shoot and root growth are associated with higher N uptake.

Quantitative trait loci for grain protein deviation

Although both traits respond to N fertilizer application, GPRO is commonly negatively correlated with GY (Nuttall et al., 2017; Simmonds, 1995). Breeding could contribute to improved GPD by increasing the favorable allele frequencies for marker alleles and QTL that condition high GPRO independently of GY. A number of association and candidate gene studies have identified potentially useful variants and QTL. Through association mapping, a QTL on chromosome 6B was detected in a wild emmer accession [*Triticum turgidum* ssp. *dicoccoides* (Körn)] which contributes up to 66% of variation for GPRO independently of GY (Joppa &

Cantrell, 1990). The underlying gene (GPC-B1) is a NAC transcription factor (NAM-B1) with pleiotropic effects on GPRO, as well as on zinc and iron concentration in the grain. Its action is mediated through accelerated canopy senescence and increased rate of nutrient remobilization from leaves to grain (Avni et al., 2014; C. Uauy et al., 2006). Homoeologs of the gene identified in Argentinean, European and Australian wheats are also associated with variation in GPRO (Cormier et al., 2015; Tabbita et al., 2013; R. Yang et al., 2019). Marker-assisted gene pyramiding efforts are underway for Australian wheats (R. Yang et al., 2019).

In a population of recombinant inbred lines developed from a cross between modern Chinese winter wheat varieties, four QTL on chromosomes 2B, 4A, 7A, and 5B were identified that explained 23% of variation in GPRO, independently of GY (Wang et al., 2012). An association analysis of a multi-parent interconnected population of French breeding lines identified QTL on chromosomes 3A and 5D that controlled GPRO independently of GY (Bogard et al., 2013). In a GWAS on panel of winter wheat from the U.S. Great Plains, five SNP marker alleles were identified that were associated with GPD (Guttieri et al., 2017). The minor allele frequencies for the SNPs on chromosome 2B and 2D were the favorable alleles, while for the SNPs on 1D and 4B (2 loci) the minor alleles had adverse effects. A panel of 1,604 European wheat hybrids was uitilized for GWAS to identify the genetic architecture underlying GPD (Thorworth et al., 2018). They observed antagonistic gene action for most of the pleiotropic QTL, confirming a genetic basis for the difficulty of simultaneous improvement.

Nitrogen metabolism-related gene sequences identified in wheat and other organisms were used for *in silico* queries of the wheat genome to identify, map, and develop allele-specific molecular markers for wheat homoeologs and orthologs of N metabolism candidate genes (Nigro et al., 2019). It has been proposed that GY and GPRO may be under independent control by their

gene action. In a diverse set of seven subspecies of tetraploid wheat (*Triticum turgidum* L.) grown in southern Italy, candidate gene association analysis revealed eight N metabolism candidate genes that explained 34.2% of variance for GPD and that three QTL on chromosome 5B and one on 4A were significantly associated with GPD (Nigro et al., 2019). Increasing the frequency of positive alleles for these QTL in hard winter wheat breeding populations may lead to selection for higher GPD. Validation of the significant associations is particularly important, given that GPD may be more difficult to detect or may result in a higher rate of false positives as a consequence of its mathematical derivation (Nigro et al., 2019; Wang et al., 2012).

Genomic selection for nitrogen use efficiency

A method to apply genome-wide markers for breeding value prediction without requiring significant marker-trait associations was developed for predicting breeding values of bulls (Meuwissen et al., 2001). Genomic selection (GS) combines phenotypic data and polymorphic marker genotypes from a training population to build a predictive model for performance of untested, but genotyped, individuals (Bernardo, 2016). Under the infinitesimal model, predictions are based on summation of genome-wide marker effects, thus capturing not only major QTL, but also unknown and minor effect QTL (Bernardo, 2014). The method aims to improve the mean performance of a population, without requiring gene discovery or knowledge of trait mechanisms. Its base assumption is that with marker density adequate to capture all linkage disequilibrium intervals, all marker effects can be estimated and their sum will be the additive genetic value for an individual. Genome-wide enrichment of favorable alleles among selection candidates would then be achieved through directional selection of the best individuals during inbreeding based on predicted genotypic values, with no requirement to identify significant associations with the underlying genes (Jannink et al., 2010). When these individuals

are cycled back into the breeding population as crossing parents, the frequencies of favorable alleles are enriched.

The first publication of the GS applied in plants was a simulation study of maize testcross performance (Bernardo & Yu, 2007). This work generated enthusiasm in the plant breeding community by demonstrating substantial improvements in response to selection. The first application within a wheat breeding program was for recurrent selection of quantitative stem rust resistance and the correlated trait pseudo-black chaff (Rutkoski et al., 2011; Rutkoski et al., 2014). Consistent with simulated schemes, realized genetic gain per unit time did not differ significantly between GS and phenotypic selection. Without increasing cycle time, while maintaining the same rate of genetic improvement, GS would reduce costs by enabling early generation selection prior to phenotyping. Optimization of GS is monitored by measuring prediction accuracy and cycle time relative to a benchmark method (Heffner et al., 2009; Larkin et al., 2019; Norman et al., 2018). Prediction accuracy, defined as the correlation between genomic estimated breeding values (GEBV) and phenotypic values, is impacted by factors that are characteristic of the breeding population and the targeted traits. As reviewed in Larkin et al. (2019), simulations and empirical studies have evaluated impacts of training population design and size, marker density, population structure, relatedness of training and validation sets, trait heritability, and choice of statistical model. For traits with one to three major QTL, with each contributing 10% or more to genetic variance, genomic prediction accuracy is improved when they are included as fixed effects (Arruda et al., 2016; Bernardo, 2014; Sarinelli et al., 2019). For complex traits, such as Fusarium head blight resistance, the selection differential obtained

with genomic selection is higher than for marker-assisted selection (simulation models included one to five QTL) under the same selection intensity (Arruda et al., 2016).

For the first time in wheat, a very large panel of lines provided an opportunity to explore the experimental space beyond which more data do not contribute to higher prediction accuracy (Norman et al., 2018). In this study, 10,375 lines in an association panel were genotyped with 18,101 markers. These data were applied via cross-validation analysis to examine impacts of the design factors on prediction accuracy for four traits with differing genetic architectures. Prediction accuracy follows a curvilinear relationship with training set size for all traits, with the curve flattening above 2,000 individuals. This response is independent of the genetic complexity of the predicted trait. Accuracy is improved with higher levels of relatedness between the training and validation sets and with increased diversity in the training set. There is an interaction between marker density, training set diversity, and relatedness wherein response to increased marker density is greatest when predicting from a diverse training set to a less related validation set. This work and similar studies provide breeders with parameters for designing an effective genomic selection program (V. Belamkar et al., 2018; Dawson et al., 2013; He et al., 2016; Michel et al., 2017).

Plant breeders usually target improvement of multiple traits to increase the economic value of plants. Phenotypic multi-trait selection strategies have included tandem selection, independent culling, and index selection (Bernardo, 2010). While GS typically has targeted single traits, multi-trait GS includes correlated traits and can produce higher prediction accuracies for those traits with unbalanced data or low heritability (Schulthess et al., 2016). Additionally, a selection index may be treated as a single trait in a univariate GS model to obtain multi-trait improvement (Schulthess et al., 2016). Application of genome wide molecular

markers for simultaneous improvement of GY and GPRO has been reported for durum wheat (Rapp et al., 2018) and bread wheat (Michel et al., 2016; Michel et al., 2019c).

Research Objectives

The objectives of this work are to detect component trait contributions to NUE, observe variation for NUE-related traits, and to develop phenotypic and genomic selection methods for simultaneous improvement of GY and GPRO within the winter wheat breeding population at Colorado State University.

Table 1.1. List and definitions for N use efficiency-related traits in wheat (*Triticum aestivum* L.)

Abbreviation	Trait	Description	Unit	Calculation or method
BMN	Shoot biomass N concentration	N concentration on a dry weight basis in the stem and leaves	g kg^{-1}	AACCI Method 46-30.01
BMY	Biomass yield	Total dry weight of leaves, stems, and chaff per unit area	Mg ha^{-1}	
BMNY	Biomass N yield	N accumulated in the above ground biomass per unit area	kg ha^{-1}	BMNY=0.001*BMN*BMY
GN	Grain N concentration	N concentration on a dry weight basis in the grain	g kg^{-1}	AACCI Method 46-30.01
GNY	Grain N yield	N accumulated in the harvested grain per unit area	kg ha^{-1}	GNY=0.001*GN*GY
GPRO	Grain protein concentration	Proportionate dry weight basis for protein in the grain	g kg^{-1}	GN * 5.7†
GPD	Grain protein deviation	Residuals of the linear regression of GPRO on GY	g kg^{-1}	Linear regression
GY	Grain yield	Grain weight adjusted to a defined moisture basis (eg 12%) per unit area	Mg ha^{-1}	[Grain dry weight * $(1\text{-}0.12)^{-1}$] * harvested $area^{-1}$
HI	Harvest index	Proportion of total biomass harvested as grain	Mg Mg^{-1}	GY * TDW^{-1}
NHI	N harvest index	Proportion of BMN translocated to the grain	g g^{-1}	GN * BMN^{-1}
NRE	N remobilization efficiency	Proportion of the BMNY that is not recovered in the grain	kg ha^{-1}	(BMNY – GNY) * $BMNY^{-1}$
N_s	N supply	Measurable N available to the crop per unit area	kg ha^{-1}	Example: residual N + applied N
NUE	N use efficiency	Grain production per unit of N supply	kg kg^{-1}	GY * N_s^{-1}
NUpE	N uptake efficiency	Efficiency of accumulation of BMNY per unit of N_s	g kg^{-1}	BMNY * N_s^{-1}
NUtE	N utilization efficiency	Efficiency of GY production per unit of BMN per unit area (BMNY)	kg kg^{-1}	GY * $BMNY^{-1}$
PANU	Post-anthesis N uptake	N translocated to the grain after flowering on a dry weight basis in the grain	g kg^{-1}	GN - BMN at anthesis
TDW	Total dry weight	Total above ground plant dry weight per unit area	Mg ha^{-1}	GY + BMY
TN	Total above ground N	N accumulated in all above ground plant parts per unit area	g ha^{-1}	(GN + BMN) * harvested $area^{-1}$

† Sosulski and Imafidon, 1990

Table 1.2. Breeding progress for grain productivity in wheat (*Triticum aestivum* L.), selected study summaries.

Target environment	Moisture source, type	N level for selection	Release dates	Trait †	Genetic gain per year	Contributing trait(s)	Citation
Canada	Rainfed, diverse	optimal	1910-2009	NUE	0.34%	N utilization	Kubota, Iqbal et al., 2018
Mexico	Irrigated	moderate	1950-1985	NUE	1.0%	N uptake & N utilization	Ortiz-Monasterio, Sayre et al., 1997;
Finland	Rainfed, replete	moderate	1901-2000	NUE	0.05 kg kg^{-1} N ha^{-1}	N uptake	Muurinen, Slafer et al., 2006;
France	Rainfed, replete	diverse	1969-2010	NUE	0.33%	N utilization	Cormier, Faure et al., 2013;
Northern Italy	Rainfed, diverse	optimal	1900-1994	AE	0.11 kg kg^{-1} N ha^{-1}	N utilization	Guarda, Padovan et al., 2004
US Southern Great Plains	Rainfed or irrigated, diverse	optimal	1971-2008	GY	0.40%	not specified	Battenfield, Klatt, et al., 2013
US Southern Great Plains	Rainfed, diverse	optimal	1992-2014	GY	1.1%	not specified	Rife, Graybosch, et al, 2020
Kansas	Rainfed, diverse	optimal	1920-2016	GY	17/62/8 kg ha^{-1} (by time period)	N utilization	Maeoka, Sadras, et al, 2020
Australia	Rainfed, diverse	moderate	1958-2007	GY	18 kg ha^{-1}	N uptake	Sadras and Lawson, 2013;

† NUE, N use efficiency, AE, agronomic efficiency, GY, grain yield

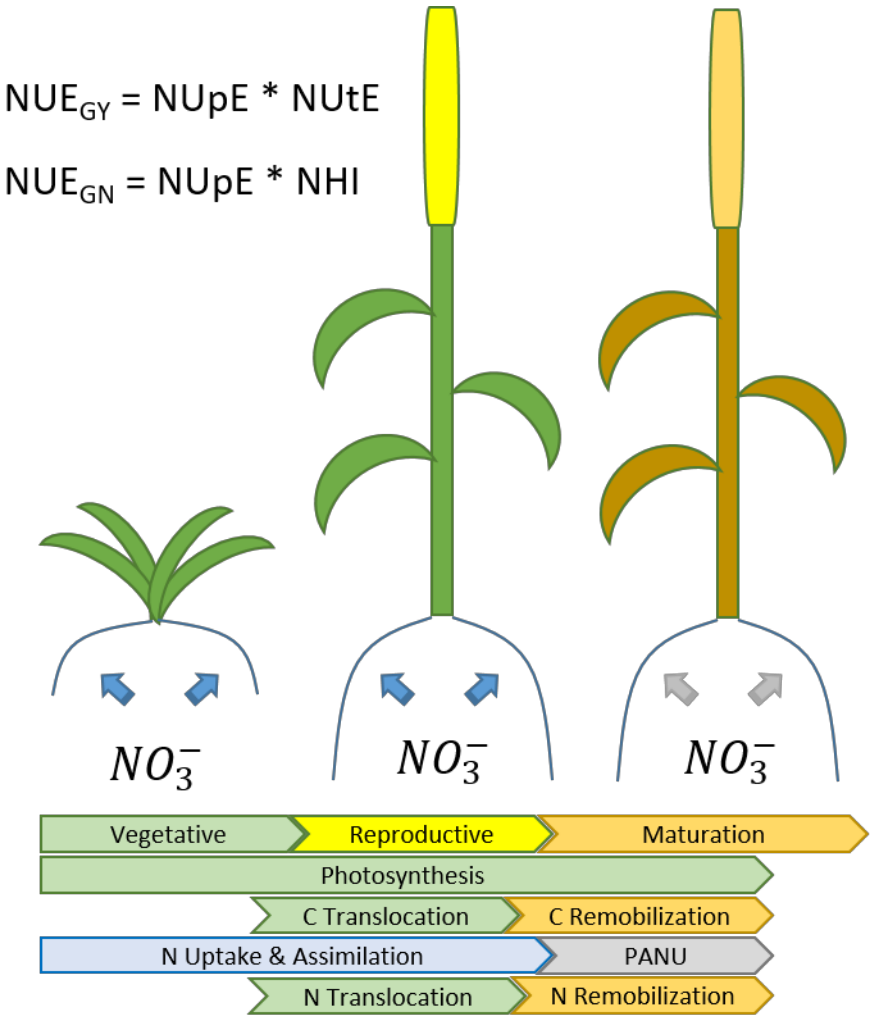


Figure 1.1. Physiological processes which determine plant growth and development. NUE_{GY}, N use efficiency for grain yield production; NUE_{GN}, N use efficiency for grain N yield; NUpE, N uptake efficiency; NUtE, N utilization efficiency; NHI, N harvest index; PANU, post-anthesis N uptake.

References

Acreche, M. M., & Slafer, G. A. (2009). Variation of grain nitrogen content in relation with grain yield in old and modern Spanish wheats grown under a wide range of agronomic conditions in a Mediterranean region. *Journal of Agricultural Science, 147*(6), 657.

An, D., Su, J., Liu, Q., Zhu, Y., Tong, Y., Li, J., . . . Li, Z. (2006). Mapping QTLs for nitrogen uptake in relation to the early growth of wheat (*Triticum aestivum* L.). *Plant and Soil, 284*(1-2), 73-73-84. doi:10.1007/s11104-006-0030-3

Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C., Brown-Guedira, G., Dong, Y., Foresman, B. J., & Kolb, F. L. (2016). Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). *Molecular Breeding, 36*(7), 84. doi:10.1007/s11032-016-0508-5

Avni, R., Zhao, R., Pearce, S., Jun, Y., Uauy, C., Tabbita, F., Fahima, T., Slade, A., Dubcovsky, J., & Distelfeld, A. (2014). Functional characterization of GPC-1 genes in hexaploid wheat. *Planta, 239*(2), 313-324. doi:10.1007/s00425-013-1977-y

Baenziger, P. S., Mumm, R. H., Bernardo, R., Brummer, E. C., Langridge, P., Simon, P., & Smith, S. (2017). *Plant breeding and genetics: a paper in the series on The Need for Agricultural Innovation to Sustainably Feed the World by 2050*. Retrieved from Ames: https://www.cast-science.org/publication/plant-breeding-and-genetics/

Bahrani, A., Abad, H. H. S., & Aynehband, A. (2013). Nitrogen remobilization in wheat as influenced by nitrogen application and post-anthesis water deficit during grain filling. *Afr. J. Biotechnol., 10*(52), 10585-10594.

Bajgain, P., Russell, B., & Mohammadi, M. (2018). Phylogenetic analyses and in-seedling expression of ammonium and nitrate transporters in wheat. *Scientific reports, 8*(1), 1-13.

Balyan, H. S., Gahlaut, V., Kumar, A., Jaiswal, V., Dhariwal, R., Tyagi, S., Agarwal, P., Kumari, S., & Gupta, P. K. (2016). Nitrogen and phosphorus use efficiencies in wheat: physiology, phenotyping, genetics, and breeding. *Plant Breed Rev, 40*, 167-214.

Barraclough, P. B., Howarth, J. R., Jones, J., Lopez-Bellido, R., Parmar, S., Shepherd, C. E., & Hawkesford, M. J. (2010). Nitrogen efficiency of wheat: genotypic and environmental variation and prospects for improvement. *European Journal of Agronomy, 33*(1), 1-11.

Barraclough, P. B., Lopez-Bellido, R., & Hawkesford, M. J. (2014). Genotypic variation in the uptake, partitioning and remobilisation of nitrogen during grain-filling in wheat. *Field Crops Research, 156*, 242-248.

Battenfield, S. D., Guzman, C., Gaynor, R. C., Singh, R. P., Pena, R. J., Dreisigacker, S., Fritz,

A. K., & Poland, J. A. (2016). Genomic selection for processing and end-use quality traits in the CIMMYT spring bread wheat breeding program. *Plant Genome, 9*(2). doi:10.3835/plantgenome2016.01.0005

Battenfield, S. D., Klatt, A. R., & Raun, W. R. (2013). Genetic yield potential improvement of semidwarf winter wheat in the Great Plains. *Crop Science, 53*(3). doi:10.2135/cropsci2012.03.0158

Belamkar, V., Guttieri, M. J., Hussain, W., Jarquín, D., El-basyoni, I., Poland, J., Lorenz, A. J., & Baenziger, P. S. (2018). Genomic selection in preliminary yield trials in a winter wheat breeding program. *G3: Genes|Genomes|Genetics, 8*(8), 2735. doi:10.1534/g3.118.200415

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: Learning from the last 20 years. *Crop Science, 48*(5), 1649-1664. doi:10.2135/cropsci2008.03.0131

Bernardo, R. (2010). *Breeding for Quantitative Traits in Plants* (2nd ed.): Stemma Press.

Bernardo, R. (2014). Genomewide selection when major genes are known. *Crop Science, 54*(1), 68-75. doi:10.2135/cropsci2013.05.0315

Bernardo, R. (2016). Bandwagons I, too, have known. *Theoretical and Applied Genetics, 129*(12), 2323-2332.

Bernardo, R., & Yu, J. (2007). Prospects for genomewide selection for quantitative traits in maize. *Crop Science, 47*(3), 1082-1090.

Bharati, A., & Mandal, P. K. (2020). Strategies for identification of genes toward enhancing nitrogen utilization efficiency in cereals. *Nutrient Dynamics for Sustainable Crop Production* (pp. 157-187): Springer.

Bogard, M., Allard, V., Brancourt-Hulmel, M., Heumez, E., Machet, J. M., Jeuffroy, M. H., Gate, P., Martre, P., & Le Gouis, J. (2010). Deviation from the grain protein concentration–grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. *Journal of Experimental Botany, 61*(15), 4303-4312.

Bogard, M., Allard, V., Martre, P., Heumez, E., Snape, J., Orford, S., Griffiths, S., Gaju, O., Foulkes, J., & Le Gouis, J. (2013). Identifying wheat genomic regions for improving grain protein concentration independently of grain yield using multiple inter-related populations. *Molecular Breeding, 31*(3), 587-599. doi:10.1007/s11032-012-9817-5

Bogard, M., Jourdan, M., Allard, V., Martre, P., Perretant, M. R., Ravel, C., Heumez, E., Orford, S., Snape, J., & Griffiths, S. (2011). Anthesis date mainly explained correlations between post-anthesis leaf senescence, grain yield, and grain protein concentration in a winter wheat population segregating for flowering time QTLs. *Journal of Experimental Botany, 62*(10), 3621-3636.

Bordes, J., Ravel, C., Jaubertie, J. P., Duperrier, B., Gardet, O., Heumez, E., Pissavy, A. L., Charmet, G., Le Gouis, J., & Balfourier, F. (2013). Genomic regions associated with the nitrogen limitation response revealed in a global wheat core collection. *Theoretical and Applied Genetics, 126*(3), 805-822.

Brasier, K., Ward, B., Smith, J., Seago, J., Oakes, J., Balota, M., Davis, P., Fountain, M., Brown-Guedira, G., & Sneller, C. (2020). Identification of quantitative trait loci associated with nitrogen use efficiency in winter wheat. *PloS one, 15*(2), e0228775.

Cameron, K. C., Di, H. J., & Moir, J. L. (2013). Nitrogen losses from the soil/plant system: a review. *Annals of Applied Biology, 162*(2), 145-173. doi:10.1111/aab.12014

Cormier, F., Faure, S., Dubreuil, P., Heumez, E., Beauchêne, K., Lafarge, S., Praud, S., & Le Gouis, J. (2013). A multi-environmental study of recent breeding progress on nitrogen use efficiency in wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics, 126*(12), 3035-3048. doi:10.1007/s00122-013-2191-9

Cormier, F., Foulkes, J., Hirel, B., Gouache, D., Moënne-Loccoz, Y., & Le Gouis, J. (2016). Breeding for increased nitrogen-use efficiency: a review for wheat (*T. aestivum* L.). *Plant Breeding, 135*(3), 255-278. doi:10.1111/pbr.12371

Cormier, F., Le Gouis, J., Dubreuil, P., Lafarge, S., & Praud, S. (2014). A genome-wide identification of chromosomal regions determining nitrogen use efficiency components in wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 1-15.

Cormier, F., Throude, M., Ravel, C., Le Gouis, J., Leveugle, M., Lafarge, S., Exbrayat, F., Duranton, N., & Praud, S. (2015). Detection of NAM-A1 natural variants in bread wheat reveals differences in haplotype distribution between a worldwide core collection and European elite germplasm. *Agronomy, 5*(2), 143-151.

Davidson, E. A., Suddick, E. C., Rice, C. W., & Prokopy, L. S. (2015). More Food, Low Pollution (Mo Fo Lo Po): A grand challenge for the 21st century. *Journal of Environmental Quality, 44*(2), 305-311. doi:10.2134/jeq2015.02.0078

Dawson, J. C., Endelman, J. B., Heslot, N., Crossa, J., Poland, J., Dreisigacker, S., Manès, Y., Sorrells, M. E., & Jannink, J.-L.. (2013). The use of unbalanced historical data for genomic selection in an international wheat breeding program. *Field Crops Research, 154*, 12-22.

de Oliveira Silva, A., Ciampitti, I. A., Slafer, G. A., & Lollato, R. P. (2020). Nitrogen utilization efficiency in wheat: A global perspective. *European Journal of Agronomy, 114*, 126008. doi:https://doi.org/10.1016/j.eja.2020.126008

Desai, R. M., & Bhatia, C. R. (1978). Nitrogen uptake and nitrogen harvest index in durum wheat cultivars varying in their grain protein concentration. *Euphytica, 27*(2), 561-566.

doi:10.1007/bf00043182

Dhugga, K., & Waines, J. (1989). Analysis of nitrogen accumulation and use in bread and durum wheat. *Crop Science, 29*(5), 1232-1239.

Dixon, J., Braun, H.-J., Kosina, P., & Crouch, J. (2009). *Wheat facts and futures 2009*. Retrieved from Mexico, D.F.:

Donmez, E., Sears, R. G., Shroyer, J. P., & Paulsen, G. M. (2001). Genetic gain in yield attributes of winter wheat in the Great Plains. Contribution no. 00-440-J from the Kansas Agric. Exp. Stn. *Crop Science, 41*(5), 1412-1419. doi:10.2135/cropsci2001.4151412x

Ehdaie, B., & Waines, J. (2001). Sowing date and nitrogen rate effects on dry matter and nitrogen partitioning in bread and durum wheat. *Field Crops Research, 73*(1), 47-61.

FAO. (2004) FAO statistical yearbook - 2004. *Vol. 1. Number of undernourished and proportion in total population*. Rome: Food and Agriculture Organization of the United Nations.

FAO. (2019) World Food and Agriculture – Statistical pocketbook. Rome.

Ferguson, R. B. (2015). Groundwater quality and nitrogen use efficiency in Nebraska's Central Platte River Valley. *Journal of Environmental Quality, 44*(2), 449-459.

Fontaine, J. X., Ravel, C., Pageau, K., Heumez, E., Dubois, F., Hirel, B., & Le Gouis, J. (2009). A quantitative genetic study for elucidating the contribution of glutamine synthetase, glutamate dehydrogenase and other nitrogen-related physiological traits to the agronomic performance of common wheat. *Theoretical and Applied Genetics, 119*(4), 645-662.

Fortunato, S., Nigro, D., Paradiso, A., Cucci, G., Lacolla, G., Trani, R., Agrimi, G., Blanco, A., de Pinto, M. C., & Gadaleta, A. (2019). Nitrogen metabolism at tillering stage differently affects the grain yield and grain protein content in two durum wheat cultivars. *Diversity, 11*(10), 186. doi:10.3390/d11100186

Gaju, O., Allard, V., Martre, P., Le Gouis, J., Moreau, D., Bogard, M., Hubbart, S., & Foulkes, M. J. (2014). Nitrogen partitioning and remobilization in relation to leaf senescence, grain yield and grain nitrogen concentration in wheat cultivars. *Field Crops Research, 155*, 213-223.

Gaju, O., Allard, V., Martre, P., Snape, J. W., Heumez, E., Le Gouis, J., Moreau, D., Bogard, M., Griffiths, S., Orford, S., Hubbart, S., & Foulkes, M. J. (2011). Identification of traits to improve the nitrogen-use efficiency of wheat genotypes. *Field Crops Research, 123*(2), 139-152. doi:10.1016/j.fcr.2011.05.010

Good, A. G., & Beatty, P. H. (2011). Biotechnological approaches to improving nitrogen use efficiency in plants: Alanine aminotransferase as a case study. *The Molecular and Physiological Basis of Nutrient Use Efficiency in Crops* (pp. 165-191): Wiley-Blackwell.

Good, A. G., Johnson, S. J., De Pauw, M., Carroll, R. T., Savidov, N., Vidmar, J., Lu, Z., Taylor, G., & Stroeher, V. (2007). Engineering nitrogen use efficiency with alanine aminotransferase. *Canadian Journal of Botany, 85*(3), 252-262.

Goucher, L., Bruce, R., Cameron, D. D., Lenny Koh, S. C., & Horton, P. (2017). The environmental impact of fertilizer embodied in a wheat-to-bread supply chain. *Nature Plants, 3*, 17012. doi:10.1038/nplants.2017.12.

Graybosch, R., Bockelman, H. E., Garland-Campbell, K. A., Garvin, D. F., & Regassa, T. (2014). Wheat. In J. Specht & B. Carver (Eds.), *Yield Gains in Major U.S. Field Crops* (pp. 459-488): American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc.

Guarda, G., Padovan, S., & Delogu, G. (2004). Grain yield, nitrogen-use efficiency and baking quality of old and modern Italian bread-wheat cultivars grown at different nitrogen levels. *European Journal of Agronomy, 21*(2), 181-192.

Guttieri, M. J., Baenziger, P. S., Frels, K., Carver, B., Arnall, B., & Waters, B. M. (2015). Variation for grain mineral concentration in a diversity panel of current and historical Great Plains hard winter wheat germplasm. *Crop Science, 55*(3), 1035-1052. doi:10.2135/cropsci2014.07.0506

Guttieri, M. J., Frels, K., Regassa, T., Waters, B. M., & Baenziger, P. S. (2017). Variation for nitrogen use efficiency traits in current and historical great plains hard winter wheat. *Euphytica, 213*(4), 87. doi:10.1007/s10681-017-1869-5

Habash, D. Z., Bernard, S., Schondelmaier, J., Weyen, J., & Quarrie, S. A. (2007). The genetics of nitrogen use in hexaploid wheat: N utilisation, development and yield. *Theoretical and Applied Genetics, 114*(3), 403-419. doi:10.1007/s00122-006-0429-5

Hamblin, M. T., Buckler, E. S., & Jannink, J.-L. (2011). Population genetics of genomics-based crop improvement methods. *Trends in Genetics, 27*(3), 98-106. doi:10.1016/j.tig.2010.12.003

Hawkesford, M. J., & Griffiths, S. (2019). Exploiting genetic variation in nitrogen use efficiency for cereal crop improvement. *Current Opinion in Plant Biology, 49*, 35-42.

He, S., Schulthess, A. W., Mirdita, V., Zhao, Y., Korzun, V., Bothe, R., Ebmeyer, E., Reif, J. C., & Jiang, Y. (2016). Genomic selection in a commercial winter wheat population. *Theoretical and Applied Genetics, 129*(3), 641-651. doi:10.1007/s00122-015-2655-1

He, X., Qu, B., Li, W., Zhao, X., Teng, W., Ma, W., Ren, Y., Li, B., Li, Z., & Tong, Y. (2015). The nitrate-inducible NAC transcription factor TaNAC2-5A controls nitrate response and increases wheat yield. *Plant Physiology, 169*(3), 1991-2005. doi:10.1104/pp.15.00568

Heffner, E. L., Sorrells, M. E., & Jannink, J. L. (2009). Genomic selection for crop improvement. *Crop Science 49*, 1-12.

Hirel, B., Le Gouis, J., Ney, B., & Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *Journal of Experimental Botany, 58*(9), 2369-2387. doi:10.1093/jxb/erm097

Hu, B., Wang, W., Ou, S., Tang, J., Li, H., Che, R., Zhang, Z., Chai, X., Wang, H., & Wang, Y.. (2015). Variation in NRT1. 1B contributes to nitrate-use divergence between rice subspecies. *Nature Genetics, 47*(7), 834.

Jannink, J. L., Lorenz, A. J., & Iwata, H. (2010). Genomic selection in plant breeding: from theory to practice. *Briefings in functional genomics, 9*(2), 166-177. doi:10.1093/bfgp/elq001

Joppa, L., & Cantrell, R. (1990). Chromosomal location of genes for grain protein content of wild tetraploid wheat. *Crop Science, 30*(5), 1059-1064.

Konishi, M., & Yanagisawa, S. (2014). Emergence of a new step towards understanding the molecular mechanisms underlying nitrate-regulated gene expression. *Journal of Experimental Botany, 65*(19), 5589-5600. doi:10.1093/jxb/eru267

Kubota, H., Iqbal, M., Dyck, M., Quideau, S., Yang, R.-C., & Spaner, D. (2018). Investigating genetic progress and variation for nitrogen use efficiency in spring wheat. *Crop Science, 58*(4), 1542-1557. doi:10.2135/cropsci2017.10.0598

Larkin, D. L., Lozada, D. N., & Mason, R. E. (2019). Genomic selection—considerations for successful implementation in wheat breeding programs. *Agronomy, 9*(9), 479.

Lassaletta, L., Billen, G., Grizzetti, B., Anglade, J., & Garnier, J. (2014). 50 year trends in nitrogen use efficiency of world cropping systems: the relationship between yield and nitrogen input to cropland. *Environmental Research Letters, 9*(10), 105011.

Le Gouis, J., Béghin, D., Heumez, E., & Pluchard, P. (2000). Genetic differences for nitrogen uptake and nitrogen utilisation efficiencies in winter wheat. *European Journal of Agronomy, 12*(3-4), 163-173. doi:10.1016/s1161-0301(00)00045-9

Li, X. P., Zhao, X. Q., He, X., Zhao, G. Y., Li, B., Liu, D. C., Zhang, A. M., Zhang, X. Y., Tong, Y. P., & Li, Z. S. (2011). Haplotype analysis of the genes encoding glutamine synthetase plastic isoforms and their association with nitrogen-use-and yield-related traits in bread wheat. *New Phytologist, 189*(2), 449-458.

Liu, B., Asseng, S., Müller, C., Ewert, F., Elliott, J., Lobell, D. B., Martre, P., Ruane, A. C., Wallach, D., & Jones, J. W. (2016). Similar estimates of temperature impacts on global wheat yield by three independent methods. *Nature Climate Change, 6*(12), 1130-1136.

Lollato, R. P., & Edwards, J. T. (2015). Maximum attainable wheat yield and resource-use efficiency in the southern Great Plains. *Crop Science, 55*(6), 2863-2876. doi:10.2135/cropsci2015.04.0215

Maeoka, R. E., Sadras, V. O., Ciampitti, I. A., Diaz, D. R., Fritz, A. K., & Lollato, R. P. (2020). Changes in the phenotype of winter wheat varieties released between 1920 and 2016 in response to in-furrow fertilizer: Biomass allocation, yield, and grain protein concentration. *Frontiers in Plant Science, 10*, 1786.

Marinciu, C., & Saulescu, N. (2009). Grain yield and protein concentration in winter wheat cultivars tested with and without nitrogen fertilizer. *Romanian Agricultural Research, 26*, 13-19.

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics, 157*(4), 1819-1829.

Michel, S., Ametz, C., Gungor, H., Akgöl, B., Epure, D., Grausgruber, H., Löschenberger, F., & Buerstmayr, H. (2017). Genomic assisted selection for enhancing line breeding: merging genomic and phenotypic selection in winter wheat breeding programs with preliminary yield trials. *Theoretical and Applied Genetics, 130*(2), 363-376.

Michel, S., Ametz, C., Gungor, H., Epure, D., Grausgruber, H., Löschenberger, F., & Buerstmayr, H. (2016). Genomic selection across multiple breeding cycles in applied bread wheat breeding. *Theoretical and Applied Genetics*, 1-11. doi:10.1007/s00122-016-2694-2

Michel, S., Löschenberger, F., Ametz, C., Pachler, B., Sparry, E., & Bürstmayr, H. (2019). Simultaneous selection for grain yield and protein content in genomics-assisted wheat breeding. *Theoretical and Applied Genetics*. doi:10.1007/s00122-019-03312-5

Millar, N., Doll, J. E., & Robertson, G. P. (2014) Management of nitrogen fertilizer to reduce nitrous oxide (N_2O) emissions from field crops. *Climate Change and Agriculture Fact Sheet Series*: Michigan State University Extension.

Moll, R. H., Kamprath, E. J., & Jackson, W. A. (1982). Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization. *Agronomy Journal, 74*(May-June), 562-564.

Monaghan, J. M., Snape, J. W., Chojecki, A. J. S., & Kettlewell, P. S. (2001). The use of grain protein deviation for identifying wheat cultivars with high grain protein concentration and yield. *Euphytica, 122*(2), 309-317.

Mu, X., & Luo, J. (2019). Evolutionary analyses of NIN-like proteins in plants and their roles in nitrate signaling. *Cellular and Molecular Life Sciences, 76*(19), 3753-3764. doi:10.1007/s00018-019-03164-8

Mueller, N. D., West, P. C., Gerber, J. S., MacDonald, G. K., Polasky, S., & Foley, J. A. (2014). A tradeoff frontier for global nitrogen use and cereal production. *Environmental Research Letters, 9*(5), 054002.

Munier-Jolain, N. G., & Salon, C. (2005). Are the carbon costs of seed production related to the quantitative and qualitative performance? An appraisal for legumes and other crops. *Plant, Cell & Environment, 28*(11), 1388-1395.

Muurinen, S., Slafer, G., & Peltonen-Sainio, P. (2006). Breeding effects on nitrogen use efficiency of spring cereals under northern conditions. *Crop Science, 46*(2), 561-568.

Nadolska-Orczyk, A., Rajchel, I. K., Orczyk, W., & Gasparis, S. (2017). Major genes determining yield-related traits in wheat and barley. *Theoretical and Applied Genetics, 130*(6), 1081-1098. doi:10.1007/s00122-017-2880-x

Nigro, D., Gadaleta, A., Mangini, G., Colasuonno, P., Marcotuli, I., Giancaspro, A., Giove, S. L., Simeone, R., & Blanco, A. (2019). Candidate genes and genome-wide association study of grain protein content and protein deviation in durum wheat. *Planta, 249*(4), 1157-1175. doi:10.1007/s00425-018-03075-1

Norman, A., Taylor, J., Edwards, J., & Kuchel, H. (2018). Optimising genomic selection in wheat: Effect of marker density, population size and population structure on prediction accuracy. *G3: Genes|Genomes|Genetics, 8*(9), 2889. doi:10.1534/g3.118.200311

Noulas, C., Alexiou, I., Herrera, J. M., & Stamp, P. (2013). Course of dry matter and nitrogen accumulation of spring wheat genotypes known to vary in parameters of nitrogen use efficiency. *Journal of Plant Nutrition, 36*(8), 1201-1218. doi:10.1080/01904167.2013.779706

Nuttall, J. G., O'Leary, G. J., Panozzo, J. F., Walker, C. K., Barlow, K. M., & Fitzgerald, G. J. (2017). Models of grain quality in wheat—A review. *Field Crops Research, 202*, 136-145. doi:https://doi.org/10.1016/j.fcr.2015.12.011

Omara, P., Aula, L., Oyebiyi, F., & Raun, W. R. (2019). World cereal nitrogen use efficiency trends: Review and current knowledge. *Agrosystems, Geosciences & Environment, 2*(1), 1800045.

Ortiz-Monasterio R., J. I., Sayre, K. D., Rajaram, S., & McMahon, M. (1997). Genetic progress in wheat yield and nitrogen use efficiency under four nitrogen rates. *Crop Science, 37*(3), 898-904. doi:10.2135/cropsci1997.0011183X003700030033x

Oury, F.-X., & Godin, C. (2007). Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favourable genotypes? *Euphytica, 157*(1-2), 45-57. doi:10.1007/s10681-007-9395-5

Papakosta, D. K., & Gagianas, A. (1991). Nitrogen and dry matter accumulation, remobilization, and losses for Mediterranean wheat during grain filling. *Agronomy Journal, 83*(5), 864-870.

Peña, P. A., Quach, T., Sato, S., Ge, Z., Nersesian, N., Changa, T., Dweikat, I., Soundararajan, M., & Clemente, T. E. (2017). Expression of the maize Dof1 transcription factor in wheat and sorghum. *Frontiers in Plant Science, 8*, 434. doi:10.3389/fpls.2017.00434

Pingali, P. L. (2012). Green revolution: impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences, 109*(31), 12302-12308.

Qu, B., He, X., Wang, J., Zhao, Y., Teng, W., Shao, A., Zhao, X., Ma, W., Wang, J., Li, B., Li, Z., & Tong, Y. (2015). A wheat CCAAT box-binding transcription factor increases the grain yield of wheat with less fertilizer input. *Plant Physiology, 167*(2), 411-423. doi:10.1104/pp.114.246959

Rafalski, J. A. (2010). Association genetics in crop improvement. *Current Opinion in Plant Biology, 13*(2), 174-180.

Rahimi Eichi, V., Okamato, M., Haefele, S. M., Jewell, N., Brien, C., Garnett, T., & Langridge, P. (2019). Understanding the interactions between biomass, grain production and grain protein content in high and low protein wheat genotypes under controlled environments. *Agronomy, 9*(11), 706.

Rapp, M., Lein, V., Lacoudre, F., Lafferty, J., Müller, E., Vida, G., Bozhanova, V., Ibraliu, A., Thorwarth, P., Piepho, H. P., Leiser, W. L., Würschum, T., & Longin, C. F. H. (2018). Simultaneous improvement of grain yield and protein content in durum wheat by different phenotypic indices and genomic selection. *Theoretical and Applied Genetics, 131*(6), 1315-1329. doi:10.1007/s00122-018-3080-z

Raun, W. R., & Johnson, G. V. (1999). Improving nitrogen use efficiency for cereal production. *Agronomy Journal, 91*(3), 357-363.

Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., & Chatterjee, S. (2019). Climate change has likely already affected global food production. *PLoS One, 14*(5), e0217148.

Reitz, L. P., & Salmon, S. C. (1968). Origin, history, and use of Norin 10 wheat. *Crop Science, 8*(6), 686-689.

Reynolds, M. P., Quilligan, E., Aggarwal, P. K., Bansal, K. C., Cavalieri, A. J., Chapman, S. C., Chapotin, S. M., Datta, S. K., Duveiller, E., Gill, K. S., Jagadish, K. S. V., Joshi, A. K., Koehler, A.-K., Kosina, P., Krishnan, S., Lafitte, R., Mahala, R. S., Muthurajan, R., Paterson, A. H., …Yadav, O. P. (2016). An integrated approach to maintaining cereal productivity under climate change. *Global Food Security, 8*, 9-18. doi:https://doi.org/10.1016/j.gfs.2016.02.002

Rife, T. W., Graybosch, R. A., & Poland, J. A. (2019). A field-based analysis of genetic improvement for grain yield in winter wheat cultivars developed in the US Central Plains from 1992 to 2014. *Crop Science, 59*(3), 905-910.

Rutkoski, J., Heffner, E., & Sorrells, M. (2011). Genomic selection for durable stem rust resistance in wheat. *Euphytica, 179*(1), 161-173. doi:10.1007/s10681-010-0301-1

Rutkoski, J. E., Poland, J. A., Singh, R. P., Huerta-Espino, J., Bhavani, S., Barbier, H., Rouse, M. N., Jannink, J.-L., & Sorrells, M. E. (2014). Genomic selection for quantitative adult plant stem rust resistance in wheat. *The Plant Genome, 7*(3), 1-10. doi:doi: 10.3835/plantgenome2014.02.0006

Sadras, V., & Lawson, C. (2013). Nitrogen and water-use efficiency of Australian wheat varieties released between 1958 and 2007. *European Journal of Agronomy, 46*, 34-41.

Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., Mason, R. E., Babar, A., Harrison, S., Sutton, R., Griffey, C. A., & Brown-Guedira, G.. (2019). Training population selection and use of fixed effects to optimize genomic predictions in a historical USA winter wheat panel. *Theoretical and Applied Genetics, 132*(4), 1247-1261. doi:10.1007/s00122-019-03276-6

Schulthess, A. W., Wang, Y., Miedaner, T., Wilde, P., Reif, J. C., & Zhao, Y. (2016). Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. *Theoretical and Applied Genetics, 129*(2), 273-287. doi:10.1007/s00122-015-2626-6

Searchinger, T., Waite, R., Hanson, C., Ranganaathan, J., & Dumas, P. (2019). *World Resources Report: Creating a Sustainable Food Future*. Retrieved from Washington, DC: https://wrr-food.wri.org/

Shewry, P. R., & Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food and Energy Security, 4*(3), 178-202.

Simmonds, N. W. (1995). The relation between yield and protein in cereal grain. *Journal of the Science of Food and Agriculture, 67*(3), 309-315. doi:10.1002/jsfa.2740670306

Sinclair, T. R., & de Wit, C. T. (1975). Photosynthate and nitrogen requirements for seed production by various crops. *Science, 189*(4202), 565-567. doi:10.2307/1740987

Tabbita, F., Lewis, S., Vouilloz, J. P., Ortega, M. A., Kade, M., Abbate, P. E., & Barneix, A. J. (2013). Effects of the Gpc-B1 locus on high grain protein content introgressed into Argentinean wheat germplasm. *Plant Breeding, 132*(1), 48-52. doi:10.1111/pbr.12011

Tester, M., & Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science, 327*(5967), 818-822. doi:10.1126/science.1183700

Thorwarth, P., Piepho, H. P., Zhao, Y., Ebmeyer, E., Schacht, J., Schachschneider, R., Kazman, E., Reif, J. C., Würschum, T., & Longin, C. F. H. (2018). Higher grain yield and higher grain protein deviation underline the potential of hybrid wheat for a sustainable agriculture. *Plant Breeding, 137*(3), 326-337. doi:10.1111/pbr.12588

Tillman, A. D. (1968). The world food problem. *Naval War College Review, 21*(2), 66-79.

Uauy, C., Brevis, J. C., & Dubcovsky, J. (2006). The high grain protein content gene Gpc-B1 accelerates senescence and has pleiotropic effects on protein content in wheat. *Journal of Experimental Botany, 57*(11), 2785-2794.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., & Dubcovsky, J. (2006). A NAC Gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science, 314*(5803), 1298-1301. doi:10.1126/science.1133649

UNDESA. (2017). World Population Prospects 2017 – Data Booklet (ST/ESA/SER.A/401 ed.): United Nations, Department of Economic and Social Affairs, Population Division.

USDA-ERS. (2019). *World and U.S. wheat production, exports, and ending stocks*. Wheat Data-Recent. Retrieved from: https://www.ers.usda.gov/data-products/wheat-data/

Van Sanford, D. A., & MacKown, C. T. (1986). Variation in nitrogen use efficiency among soft red winter wheat genotypes. *Theoretical and Applied Genetics, 72*(2), 158-163. doi:10.1007/bf00266987

Wang, L., Cui, F. A., Wang, J., Jun, L. I., Ding, A., Zhao, C., Li, X., Feng, D., Gao, J., & Wang, H. (2012). Conditional QTL mapping of protein content in wheat with respect to grain yield and its components. *Journal of Genetics, 91*(3), 303-312. doi:10.1007/s12041-012-0190-2

Wang, R. F., An, D. G., Hu, C. S., Li, L. H., Zhang, Y. M., Jia, Y. G., & Tong, Y. P. (2011). Relationship between nitrogen uptake and use efficiency of winter wheat grown in the North China Plain. *Crop and Pasture Science, 62*(6), 504-514.

Wang, Y.-Y., Cheng, Y.-H., Chen, K.-E., & Tsay, Y.-F. (2018). Nitrate transport, signaling, and use efficiency. *Annual Review of Plant Biology, 69*(1), 85-122. doi:10.1146/annurev-arplant-042817-040056

Wharton, C. R. (1969). The Green Revolution: cornucopia or Pandora's box? *Foreign Affairs, 47*(3), 464-476.

Wu, J., Zhang, Z.-S., Xia, J.-Q., Alfatih, A., Song, Y., Huang, Y.-J., Wan, G.-Y., Sun, L.-Q., Tang, H., & Liu, Y. (2020). Rice NIN-LIKE PROTEIN 4 is a master regulator of nitrogen use efficiency. *bioRxiv*. doi:DOI: 10.1101/2020.01.16.908558

Yang, J., Wang, M., Li, W., He, X., Teng, W., Ma, W., Zhao, X., Hu, M., Li, H., & Zhang, Y.

(2019). Reducing expression of a nitrate-responsive bZIP transcription factor increases grain yield and N use in wheat. *Plant Biotechnology Journal, 17*(9), 1823-1833.

Yang, R., Juhasz, A., Zhang, Y., Chen, X., Zhang, Y., She, M., Zhang, J., Maddern, R., Edwards, I., & Diepeveen, D. (2019). Molecular characterisation of the NAM-1 genes in bread wheat in Australia. *Crop and Pasture Science, 69*(12), 1173-1181.

CHAPTER 2

GENOTYPIC DIFFERENCES FOR NITROGEN USE EFFICIENCY AND GRAIN PROTEIN DEVIATION IN HARD WINTER WHEAT[1]

Summary

Breeding superior bread wheat (*Triticum aestivum* L.) genotypes requires sufficient genetic variation to obtain high grain yield and adequate protein concentration. This study was conducted to determine variation for nitrogen (N) use efficiency (NUE) and grain protein deviation among 20 hard winter wheat genotypes in one season and for two recently released cultivars ('Snowmass' and 'Byrd') in a second season, under five N application rates (0, 28, 56, 84, 112 kg ha^{-1}). Among these genotypes, the proportionate contributions of component traits to total variance for NUE ranged widely: N uptake efficiency (57-89 kg kg^{-1}) and N utilization efficiency (11-43 kg kg^{-1}). Across all genotypes, N utilization efficiency contributed the most to variance for NUE under moderate to high N supply while N uptake efficiency contributed more under N-limiting conditions. Increased NUE promotes high grain yields, but may result in decreased grain N concentration through the commonly observed negative correlation of these traits. Analysis of residuals from regression of grain protein concentration on grain yield, or 'grain protein deviation', identified one cultivar ('Brawl CL Plus') that had 6.7 g kg^{-1} higher grain protein concentration than the average for all 20 genotypes. These results for a

[1] A version of this chapter was published in Agronomy Journal 108:2201–2213 (2016) doi:10.2134/agronj2016.02.0070.

representative sample of a breeding population suggest that sufficient variation is available to improve NUE and grain protein deviation through breeding.

Introduction

Nitrogen use efficiency for cereal grain yield is estimated to average 33% worldwide due to limitations imposed by fertilizer N losses from the crop and soil system (Raun & Johnson, 1999) and to inefficiencies of N uptake and utilization by the crop (Hawkesford, 2014). These limitations constrain cereal grain production and pose risks for N escape and consequent environmental damage. Nitrogen use efficiency in cereals is defined as the amount of grain yield produced per unit of N supply (Moll et al., 1982). The major physiological processes contributing to NUE are: 1) N uptake from the soil and its assimilation into plant parts and, 2) utilization of the assimilated N for grain production (Moll et al., 1982). Bread wheat quality requires adequate grain protein concentration, but high-yielding wheat genotypes commonly have low grain protein concentration.

Breeding to improve NUE will improve both component traits (Moll et al., 1982). Efficient N uptake requires root systems that effectively explore the soil, transport N into root cells, and translocate N into the shoot for assimilation into plant parts (Hawkesford, 2011). Utilization of assimilated N for grain production is most efficient when photosynthesis and the rate of starch production are unhampered and occur at a high rate. These conditions are met when the environment provides sufficient N and moisture throughout the growing season and when the rate of canopy senescence allows continued nutrient remobilization during grain filling (Hawkesford, 2011).

The relative contributions of NUE component traits vary by crop species, genotype, agronomic management and other environmental conditions, as reviewed in Wang et al. (2011). In Finland, the genetic variance for N utilization efficiency was the sole contributor to variation in NUE within a set of spring wheat cultivars released over the past century when selection was done under high N (Muurinen et al., 2006). In contrast, for irrigated spring wheat grown in California and Australia, genetic variation for N uptake efficiency explained most of the variance for NUE (Dhugga & Waines, 1989; Sadras & Lawson, 2013). Moll et al. (1982) suggested that at moderate N rates, selection should improve NUE by selection for both N uptake efficiency and N utilization efficiency. A set of Mexican spring wheat genotypes that were selected under moderate N application rates had both increased yield at low N supply and responsiveness to N fertilizer application. The researchers reported both increased N uptake and utilization efficiencies (Ortiz-Monasterio R. et al., 1997). Clearly, it is important to assess the physiological basis of NUE as it relates to agronomic conditions and the genetic diversity of a breeding population in advance of establishing a breeding strategy.

Two general methods are applied to characterize the relative importance of component trait contributions to NUE. The degree of association is estimated by correlation and the degree of relationship by linear regression (Acreche & Slafer, 2009; Barraclough et al., 2010; Dhugga & Waines, 1989; Gaju et al., 2011; Muurinen et al., 2006; Wang et al., 2011). In cases where the relationship is established, path analysis can then quantify the net contribution of each component trait to variation for the resultant trait (Bogard et al., 2013; Dhugga & Waines, 1989; Le Gouis et al., 2000; Moll et al., 1982; Ortiz-Monasterio R. et al., 1997). Path analysis quantifies component trait contributions through standardized regression coefficients, relating them both directly and indirectly through the other components (Moll et al., 1982).

Selection for N responsiveness has had the unintended effect of selection for reduced grain protein concentration (Acreche & Slafer, 2009). This phenomenon is the 'N dilution effect' that describes an allometric relationship between crop biomass and biomass N concentration (Justes et al., 1994). Nitrogen is remobilized from the canopy to the grain for assimilation into structural and storage proteins. The N harvest index is defined as the fraction of the N in the canopy that is harvested in the grain. The 'self-destruction' hypothesis predicts that under N-limiting conditions, increased translocation of N from proteins in vegetative plant parts leads to declining photosynthesis, increased rate of senescence, and a shortened grain-filling period (Sinclair & de Wit, 1975). Accordingly, Barraclough et al. (2010) emphasize that to both increase grain yield and maintain grain protein concentration, N harvest index must increase, while maintaining a functional photosynthetic system. The authors suggest that accumulation and subsequent transfer of non-photosynthetic sources of N might explain the observed variation among genotypes for N harvest index. However, in a subsequent study of N pools in a set of elite genotypes, N was remobilized efficiently from all plant tissues, suggesting that all remobilized N was in fact metabolic and not structural (Barraclough et al., 2014).

Retaining N in the canopy in photosynthetically active tissues promotes N utilization efficiency for grain yield through continued C assimilation and translocation to the grain, but N remobilization to the grain may be constrained (Barraclough et al., 2010). In some agronomic environments, N absorbed after anthesis contributes to NUE for grain yield or grain protein concentration (Van Sanford & MacKown, 1986). Relationships of grain protein concentration with the physiological traits, N remobilization efficiency and post-anthesis N uptake, differ among genotypes and N supply (Bahrani et al., 2013; Monaghan et al., 2001; Gaju et al., 2014). The N harvest index is the proportion of the N in the crop canopy harvested with the grain. High

values result from efficient N remobilization from the canopy and translocation to the grain (Desai & Bhatia, 1978).

$$N\ harvest\ index = \frac{grain\ N\ yield}{total\ N\ uptake} = \frac{post\ anthesis\ N\ uptake}{total\ N\ uptake} * \frac{grain\ N\ yield}{post\ anthesis\ N\ uptake}$$

Variation in N harvest index is impacted by the combined variation in N remobilization and post-anthesis N uptake, the main contributors to grain N yield (Moll et al., 1982).

Exceptional genotypes that deviate from the negative relationship between grain yield and grain protein concentration have been reported (Bogard et al., 2010; Ehdaie & Waines, 2001; Marinciu & Saulescu, 2009; Monaghan et al., 2001; Oury & Godin, 2007; Cristobal Uauy et al., 2006). This deviation is known as 'grain protein deviation', and genotypes with high grain protein deviation achieve grain N concentration that exceeds the predicted value for a given level of grain yield (Monaghan et al., 2001). Genes have been identified that may underlie differences for this trait. A quantitative trait locus (QTL) that contributes to grain N concentration independently of grain yield (Joppa & Cantrell, 1990) was identified in a wild emmer accession (*Triticum turgidum* ssp. *dicoccoides* [Körn]). The underlying gene (Gpc-B1) is a NAC transcription factor with pleiotropic effects on grain N concentration through accelerated senescence and increased rate of nutrient remobilization from leaves to grain (Uauy et al., 2006b). Additionally, important contributions by post-anthesis N uptake have been reported (Bogard et al., 2010). In a study of a population of recombinant inbred lines developed from a cross between modern Chinese winter wheat varieties, two QTL were identified that condition high grain protein concentration independently of grain yield (Wang et al., 2012). A QTL analysis of a multi-parent interconnected population of French breeding lines identified QTL on chromosomes 3A and 5D that controlled grain N concentration independently of grain yield

(Bogard et al., 2013). These examples of genetic variants that produce high values for grain protein deviation encourage evaluation of diversity among breeding lines.

No prior assessments have been reported for variation in NUE among adapted US hard winter wheat germplasm in the west-central Great Plains. The objectives of this study were (i) to determine genetic variability for NUE among a set of 20 U.S. hard winter wheat genotypes, (ii) to identify physiological traits associated with high N use efficient genotypes, and (iii) to characterize genetic variation for grain protein deviation. We hypothesize that sufficient variation exists in this breeding population to identify new genotypes with both high grain yield and sufficient grain protein concentration to meet market standards for bread wheat.

Materials and Methods

Plant material

Twenty hard winter wheat genotypes were assembled that included six hard white winter and fourteen hard red winter wheat genotypes (Table 2.1). Genotypes were chosen based on factors that included: current and past popular Colorado cultivars, advanced lines from the Colorado State University (CSU) wheat breeding program, parents from molecular marker mapping populations (El-Feki et al., 2015), and several cultivars from neighboring-state breeding programs. Two genotypes were chosen to include in a second year of the trial, based on their importance to Colorado growers. The hard red winter wheat cultivar, 'Byrd' (Haley et al., 2012a) was released in 2011, having superior grain yield and good milling and baking quality. Hard white winter wheat cultivar, 'Snowmass' (Haley et al., 2011) was released in 2009, having exceptional milling and baking quality with good grain yield.

Growing conditions

Research was conducted during two field seasons at the USDA-ARS Central Great Plains Research Station (40°15′ N, 103°15′ W, 1383 m elevation above sea level), near Akron, CO. The location has a semiarid climate with 418 mm average annual precipitation. Cumulative values for precipitation and growing degree days (GDD) were determined for each growing season, and were compared to the 103 year averages (1910-2012, Figure 2.1). Climatic variables were tracked by 'day of year' (DOY) to simplify analysis for winter wheat growing seasons that extended across two calendar years. The first day of January is DOY = 1. In Fig. 1, earlier dates in a growing season are in the negative DOY range, while later dates are in the positive range. Growing degree days were calculated as $GDD = [(T_{MAX} + T_{MIN})/2] - T_{MIN}$. T_{MAX} and T_{MIN} are daily maximum and minimum temperatures at 2 m above the soil surface. We used a floor of 0 °C and a ceiling of 25 °C for T_{MIN} and T_{MAX} (McMaster & Wilhelm, 1997).

Individual experimental units (plots) were direct seeded with a double disc no-till drill into proso millet (*Panicum miliaceum* L.) stubble in both growing seasons. Phosphorus as ammonium polyphosphate ($[NH_4PO_3]_n$) and N as ammonium was applied as a starter with the seed as a 10-34-0 liquid formulation at an application rate of 7.3 kg ha^{-1} of inorganic P and 4.9 kg ha^{-1} ammoniacal N. At or just before the booting stage [Feekes stages 9-10, Large (1954)], 1.2 m alleys were sprayed with glyphosate. Weed management was conducted according to the standard agronomic practice for wheat production. No disease or insect controls were necessary.

Experimental design

2010-2011 Growing Season

The field was managed prior to this experiment in a wheat-proso millet-summer fallow rotation, with no-till management for the season prior to planting. The soil type is Rago silt loam (fine montmorillonitic mesic Pachic Argiustolls [Soil Survey Staff, 1999]). As sampled immediately before planting, soil residual N averaged 57.2 kg NO_3-N ha^{-1} in the 60 cm soil profile. The planting date was 6 October 2010 at a seeding rate of 222 seeds m^{-2}. Individual experimental units (plots) were direct seeded with a double disc no-till drill into proso millet stubble both years (2010 and 2011). The trial was conducted as a factorial design with a split plot randomization in three replicates (rep). Plot size was 9.8 m long x 1.52 m wide in six rows on 0.23 m spacing, with a harvested area of 13 m^2. Nitrogen rate (N rate) was the main plot, with sub-plots of 20 genotypes randomized in five incomplete blocks of four genotypes within each N rate by rep combination. Nitrogen fertilizer (urea 46-0-0) was surface broadcast at rates of 0, 28, 56, 84, and 112 kg N ha^{-1} before planting. Due to equipment failure, the 10-34-0 NPK starter fertilizer was not applied in the first replication.

2011-2012 Growing Season

The field had been planted to proso millet and then fallowed in the prior two growing seasons. The soil type is a Weld silt loam (Fine, smectitic, mesic Aridic Argiustolls). As sampled immediately before planting, soil residual N averaged 43.7 kg NO_3-N ha^{-1} in the 60 cm soil profile. The planting date was 3 October 2011 at a seeding rate of 173 seeds m^{-2}. The trial was conducted with a split plot randomization with four reps. Plot size was 4.88 m long and 1.52 m wide in six rows on 0.23 m spacing, with a harvested area of 5.6 m^2. Nitrogen rate was

the main plot, with sub-plots of two genotypes (Snowmass and Byrd) completely randomized within each N rate. Nitrogen fertilizer (urea 46-0-0) was surface broadcast at rates of 0, 28, 56, 84, 112, and 140 kg N ha^{-1} shortly after planting.

Data sampling

For both growing seasons, residual soil N in the form of NO_3-N was measured at planting and after harvest (Table 2.2). Sampling dates for the first season were 14 Sept 2010 and 11 Aug 2011 and for the second season were 28 Sept 2011 and 5 July 2012. Two soil cores per plot were combined, with depth increments of 0-15 cm, 15-30 cm, and 30-60 cm. Before sowing, the odd numbered plots were sampled in the first season and all plots were sampled in the second season. All plots were sampled after harvest in both seasons. Soil inorganic N was determined colorimetrically by cadmium reduction of KCl-extracted soil samples (Keeney & Nelson, 1982) in a Quickchem FIA 8000 series autoanalyzer (Lachat Instruments, Loveland, Colorado).

Anthesis date was recorded when 50% of the heads in a plot had begun to extrude anthers. Heading date was scored for each plot as the DOY on which 50% of the heads were fully emerged above the flag leaf. Physiological maturity was scored when 50% of the peduncles in a plot were golden in color (Hanft and Wych, 1982).

Total aboveground biomass samples were cut at the soil surface on dates near to the trial averages for anthesis date (TDWa). These samples were used to determine biomass dry weight and N concentration. Sampling at anthesis was done over a three-day period from 6-8 June 2011. Unpatterned variability in plant stands both across the field and within each plot resulted from dry conditions at sowing and other field effects. To best sample across the range of plant density within a plot, an entire outside row was cut. For the 2011-2012 trial, 1 m samples were

cut on 23 May 2012 from an internal row within each plot, leaving at least 15 cm adjacent to the alleyways. The entire sample from each plot was shredded in a chipper shredder and then thoroughly mixed by hand before sub-sampling (150-250 g). The sub-samples were oven dried (60 °C) to constant weight. At crop physiological maturity, 1 m samples were cut from the middle rows within each plot in both seasons (12 July 2011 and 6 June 2012). Stems and leaves were separated from the grain and chaff to record straw and grain biomass dry weight. The chaff was discarded and total aboveground biomass at maturity (TDWm) was approximated by the summed dry weights of grain and straw. Grain weight and grain moisture concentration were recorded for each plot by the on-combine weighing system during mechanical harvesting.

Nitrogen concentration (NC) on a dry weight basis of vegetative parts at anthesis and maturity and of the grain was determined by pyrolysis of finely ground samples and thermo-conductivity detection of N (AACCI Method 46-30.01) on a TruSpec Micro CHNS analyzer (LECO Corp. St Joseph, MI).

Treatments and statistical analysis

The 2010-2011 trial was conducted as a factorial design with a split plot randomization in three replicates (rep). Nitrogen at rates of 0, 28,56,84,and 112 kg N ha^{-1} (N rate) was the main plot, with subplots of 20 genotypes randomized in five incomplete blocks of four genotypes within each N rate by rep combination.

The 2011-2012 trial was conducted with a split plot randomization in four reps. Nitrogen rate was the main plot, at rates of 0, 28, 56, 84, 112, and 140 kg N ha^{-1}, with subplots of two genotypes (Snowmass and Byrd) completely randomized within each N rate.

The Box-Cox procedure was applied to evaluate whether data transformation was needed so that residuals of the regression model approximated the normal distribution (Box & Cox, 1964). In all cases where transformation was needed, taking the natural logarithm of the data was effective for obtaining a normal distribution. To obtain adjusted means for all traits, the LSMEANS statement was applied to calculate least squares means in the MIXED procedure of the SAS statistical software (SAS Institute Inc, 2010). Satterthwaite-type denominator degrees of freedom were calculated by selecting the Kenward-Rogers option in PROC MIXED. Year, N rate, genotype and their interactions were considered as fixed effects. Random effects included reps, their interaction with N rate, and, in the first growing season, incomplete blocks nested within rep x N rate. Spatial covariance structures were included in the REPEATED statement to account for patterns of field variability for some of the traits (Littell et al., 2006). Tested structures included row-column, power, anisotropic power, and Matérn (SAS Institute Inc, 2010). Stepwise model selection was guided by the Akaike Information Criterion (AIC, Akaike 1981). Adjusted values were examined graphically and evaluated against observed field variation to support choice of the best model for each trait.

To account for a pattern of reduced plant stand observed in three ranges of the 2010-2011 trial, a categorical covariate for range within the arrangement of plots was included when its fixed effect was significant in the analysis of variance. The three southernmost ranges had many gaps in the plant stands and were assigned a value of '1' and the remaining two ranges were assigned a value of '2'. The categorical covariate was assigned a value of '3' for those main plots with N rate level of 112 kg ha^{-1} since the pattern of response to reduced stand differed for this N rate. The categorical covariate was included in the models for grain yield, N yield of the

aboveground biomass at anthesis, post-anthesis N uptake, N uptake efficiency, N utilization efficiency, and N harvest index.

Analysis of variance within the MIXED procedure determined significance ($p < .05$) of fixed effects for genotype, applied N rate, and their interactions for the 2010-2011 trial. A combined analysis of variance was performed for the genotypes Byrd and Snowmass over both years, with year as a fixed effect. Among significant fixed effects, adjusted means were separated by the least significant difference test (LSD, $\alpha = .05$). Group means were compared in the TTEST procedure of the SAS statistical software (SAS Institute Inc, 2010).

Between pairs of traits, Pearson Product Moment (r, for linear relationships) and Spearman Rank Order (r_s, for comparing rankings of traits with non-linear associations) correlation coefficients were determined in JMP® Pro 11.0.0 (SAS Institute Inc., 2013).

Calculations

Yield (Y) was calculated per unit area for grain, aboveground biomass at anthesis and maturity, and straw. Grain yield was calculated for each plot from grain weight adjusted to a 12% moisture basis. Harvest index was calculated as units of grain yield per unit of total aboveground biomass at maturity. Grain protein concentration was calculated by multiplying grain N concentration by 5.7 (Sosulski & Imafidon, 1990).

Nitrogen supply (N_s) was defined as the sum of the applied N and the trial average for soil residual nitrate (NO_3-N) in the top 60 cm of the soil profile. Nitrogen removal for each plant tissue type was calculated as the total N yield (NY):

$$NY = 0.01 * NC * Y$$

where NC is the N concentration and Y is the yield per unit area. The total N uptake in the crop was estimated by the NY of the total aboveground biomass at anthesis (Achreche & Slafer, 2009). Total N accumulated at maturity was calculated as the sum of NY of the straw and grain. Nitrogen remobilized to the grain was estimated as the difference between the NY in the total aboveground biomass at anthesis and NY of the straw. Post-anthesis N uptake was estimated as the difference between the NY total aboveground biomass at maturity and NY in the total aboveground biomass at anthesis (Cox et al., 1985; Monaghan et al., 2001). The fraction of the N in the plant that is harvested in the grain, or nitrogen harvest index (NHI), was calculated as the ratio of grain NY and NY of the total aboveground biomass at anthesis.

Nitrogen use efficiency was calculated as the ratio of units of grain yield per unit of N supply. Nitrogen uptake efficiency was calculated as NY of the total aboveground biomass at anthesis divided by N supply. Nitrogen utilization efficiency was calculated as units of grain yield produced per unit NY for the total aboveground biomass at anthesis.

On a dry weight basis in common units among traits, NUE is the product of N uptake efficiency and N utilization efficiency:

$$N\ use\ efficiency = N\ uptake\ efficiency * \ N\ utilization\ efficiency$$

$$\frac{grain\ yield}{nitrogen\ supply} = \frac{total\ assimilated\ \ N}{nitrogen\ supply} * \frac{grain\ yield}{total\ assimilated\ \ N}$$

Path analysis of the physiological basis for NUE was done according the method of Moll (1982), where the product relationships of component traits to the resultant trait are linearized by taking the logarithms of each factor. The log of the resultant trait (Y_k) for experimental unit *k* is then the sum of the logs of the component traits (X_{ik}). When these are expressed as deviations from

the mean, the sum over all k treatments of the cross-products of the deviations for the ith component trait (x_{ik}) and Y_k (y_k) equals the sum over all k sums of squares for Y_k:

$$\Sigma_k y_k^2 = \Sigma_k (\Sigma_i y_k x_{ik})$$

With rearrangement of terms, the sum of the cross products of each term divided by the sums of squares for Y_k gives the relative contributions of each component trait, and includes indirect effects through covariance with the other component trait:

$$1 = \frac{\sum_k y_k x_{ik}}{\sum_k y_k^2} + \frac{\sum_k y_k x_{i\prime k}}{\sum_k y_k^2} \; for \; i \neq i'$$

Grain yield adjusted means (x_i) were regressed on grain protein concentration adjusted means (y_i) for all genotypes (i) with the function lm in the base R package (R Core Team, 2012). The linear regression model that represented typical genotypes was determined by stepwise exclusion of genotypes with standardized residuals (z_i) that exceeded a threshold of 2.5% of the standard normal distribution ($|z_i| > 1.96$). For each regression model, the standardized residual for each genotype was calculated as:

$$z_i = \frac{residuals_i}{standard\ deviation\ of\ residuals \times \sqrt{\left(1 - \frac{predicted\ y_i}{y_i}\right)}}$$

Genotypes that exceeded the 2.5% threshold were removed and then the regression equation was re-calculated. The Pearson's correlation was determined for grain yield and grain protein concentration for the retained genotypes at each step. The final regression model was chosen when no standardized residual exceeded the threshold. The final regression equation was applied to all of the data to calculate the predicted values for grain protein concentration. According to

methods developed by Oury and Godin (2007), grain protein deviation (GPD) was determined by first calculating residuals for genotype i as the difference between the adjusted mean and the calculated predicted value. The standardized residuals were then calculated for each genotype, as described above. Genotypes with grain protein deviation that exceed the 5% threshold of the standard normal distribution (|GPD|>1.64) are considered to have high values (Oury & Godin, 2007).

RESULTS AND DISCUSSION

Climatic conditions and phenology

Climate conditions for each growing season are compared to the 103 year averages for cumulative precipitation and temperatures in Figure 2.1. Cumulative precipitation during the 2010-2011 growing season was below the 103 year average until 19 May 2011 (DOY 139, Figure 2.1a). After that, significant rain events occurred, resulting in above average cumulative precipitation for the growing season. A record-breaking drought occurred in 2011-2012 (Figure 2.1a), producing the driest and warmest conditions measured in 107 years of weather data collection at the USDA-ARS Central Great Plains Research Station. Cumulative temperatures were higher throughout both seasons than the 103-year average (Figure 2.1b). In 2011-2012, the 140 kg ha^{-1} N level was included in the design of the trial in case of a growing season with above average precipitation. Not unexpectedly, given the 2012 drought conditions, no N response was observed for the 140 kg ha^{-1} N rate, so it was excluded from the analysis (data not shown).

Adjusted means for phenological traits are presented in Tables 2.3 and 2.4 for plant height, heading date, physiological maturity, and grain-filling period. The average heading date was 30 May 2011 (DOY 150) and the average physiological maturity date was 4 July 2011

(DOY 185). In the second growing season, average heading date was two weeks earlier (16 May 2012, DOY 137) and average physiological maturity was reached nearly 3 weeks earlier (14 June 2012, DOY 166). The mean plant height was 74.5 cm in 2010-2011 and 71.7 cm in 2011-2012. The shortfall in normal precipitation for May 2012, combined with warmer than normal temperatures, stunted plant growth and accelerated the development of the crop.

Grain yield and grain protein concentration

For all traits, the significance of correlation coefficients and fixed effects in the analyses of variance were determined at the .05 probability level. For the 20 genotypes in the 2012-2011 trial, main effects for genotype were significant for grain yield and grain protein concentration, while main effects for N rate were significant only for grain protein concentration (Table 2.5). The trial average for grain yield was 4.5 Mg ha^{-1} and for grain protein concentration was 112 g kg^{-1}. Grain yield ranged from 4.9 Mg ha^{-1} for Byrd (109.6% of trial mean) to 4.1 Mg ha^{-1} for 'Arlin' (Sears et al., 1997) (90.9% of trial mean). The unexpected result that grain yield was not responsive to N rate in 2010-2011 may have resulted from the early season drought conditions. Similarly, under the agronomic conditions of a long-term study in China, optimal N rate averaged 135 kg ha^{-1} in normal precipitation years and decreased to 45 kg ha^{-1} in dry years (Guo et al., 2012)

For the 20 genotypes in 2010-2011, grain protein concentration had a significant interaction term for genotype with N rate (Table 2.5). In a dryland field study in Colorado, grain protein concentration below 110 g kg^{-1} was proposed as a post-harvest indicator for N limitations on grain yield production of hard winter wheat (Goos et al., 1982). To understand the significant interaction, we applied this threshold to define 'low' and 'high' N conditions. Nitrogen rate was categorized as low (0 and 28 kg ha^{-1}) or high (84 and 112 kg ha^{-1}) and then adjusted means were

calculated for each genotype at the two N sufficiency levels. Trait values for 56 kg ha^{-1} level were not included in the analysis since they were not consistently categorized as high or low. In Figure 2.2, grain protein concentration for each genotype was plotted at high N rates (84 and 112 kg ha^{-1}) vs. the grain protein concentration at low N rates (0 and 28 kg ha^{-1}). Less N responsive genotypes would be closest to the bisecting line. Grain protein concentrations for all genotypes fell above the bisect line, illustrating that all genotypes were responsive to applied N. At low N, Prairie Red (Quick et al., 2001), a lower yielding genotype, was the only genotype with grain protein concentration exceeding 110 g kg^{-1}, while at high N, all genotypes were responsive to applied N, exceeding 110 g kg^{-1}. To further evaluate the nature of the significant interaction between genotype and N rate, the slopes of the N response for each genotype were compared with the N response of the genotype, Byrd. The genotype Arlin had a significantly lower slope ($p < .05$), that explained the significant genotype by N rate interaction. Relative to Byrd, Arlin accumulated 8.9 g kg^{-1} less protein per unit of increase in N supply and was the only genotype that differed significantly from Byrd. Given the limited nature of the interaction, we focused on the significant main effects of genotype and N rate ($p < .05$, Tables 2.6 and 2.7).

Despite early season drought conditions and ample residual soil N, N fertilization was required to obtain 110 g kg^{-1} grain protein concentration, a typical level required for bread wheat on the U.S. commodity markets. In 2010-2011, the grain protein concentration values, averaged across N rates, ranged from 118.5 g kg^{-1} for Prairie Red to 105.6 g kg^{-1} for Byrd. These values are 5.8% above and 5.7% below the mean (112 g kg^{-1}). Averaged across genotypes, values at each N rate increased from 95.5 g kg^{-1} to 124.3 g kg^{-1} (Table 2.6), with a linear increase over the range of N rates (Protein = 109.7 + 0.24×N rate, R^2=0.98). Below the 56 kg ha^{-1} N rate, the average grain protein concentration would have resulted in discounted pricing for this crop.

For Snowmass and Byrd, the combined analysis of variance over two growing seasons showed a significant effect for year on grain yield and grain protein concentration. The interaction terms of year × N rate × genotype and N rate × genotype were significant for grain yield. The significant interaction was largely due to low grain yield for Byrd at the 84 kg ha^{-1} N rate in 2011, where two of the three plots for Byrd had substantial plant stand gaps. Therefore, the significant interaction was not explored further and the trials were analyzed separately by year.

The analysis of variance for each year revealed significant fixed effects for genotype in both years for grain yield and for N rate in 2011-2012. For grain protein concentration, fixed effects for genotype were significant in 2010-2011 and for N rate in both years. Across all N rates in 2010-2011, the average grain yield for Snowmass and Byrd was 4.8 Mg ha^{-1}, while in 2011-2012, the average was reduced by 40% to 2.9 Mg ha^{-1}. In 2010-2011 and 2011-2012, grain yield for Byrd (5.1 and 3.1 Mg ha^{-1}) was higher than for Snowmass (4.6 and 2.7 Mg ha^{-1}). In 2010-2011 across treatments, grain protein concentration averaged 139.9 g kg^{-1}. Grain protein concentration increased linearly from the 0 N rate (88.6 and 127.3 g kg^{-1}) to the 112 kg ha^{-1} N rate (120.9 and 150.5 g kg^{-1}) in each season. In each season, Snowmass had higher grain protein concentration than Byrd, although the difference was significant only in 2010-2011 (111.1 and 102.3 g kg^{-1}).

Nitrogen uptake

A practical measure of maximal N uptake is the NY of the total aboveground biomass at anthesis (TDWaNY, Acreche & Slafer, 2009). The analysis of variance was performed for the 2010-2011 trial for post-anthesis N uptake and NY of the total aboveground biomass at anthesis (Table 2.5). The interaction term for genotype with N rate was not significant for either trait.

Main effects for genotype were significant for both traits, while for N rate were significant for NY of the total aboveground biomass at anthesis (Tables 2.5 and 2.6). The 2010-2011 trial average for NY in the total aboveground biomass at anthesis was 99.3 kg ha^{-1}. Over the range of N rates in 2011, NY in the total aboveground biomass at anthesis showed a linear increase (TDWaNY = 81.34 + 0.32×N rate, R^2=0.91). This exceeded the soil residual N by an average of 21.5 kg ha^{-1} across the 0 N rate plots, and provides an estimate of the rate of soil N mineralization. Averaged across all N rates, the NY in the total aboveground biomass at anthesis ranged from 110.6 kg ha^{-1} for Byrd to 87.8 kg ha^{-1} for CO940610, a breeding line that was not released due to inferior baking quality characteristics (Chao et al., 2007; El-Feki et al., 2015). These values ranged from 11.4% above to 11.6% below the 2011 mean. In the combined analysis for Snowmass and Byrd in two growing seasons, N rate had significant effects on NY in the total aboveground biomass at anthesis with a linear increase (TDWaNY = 81.14 + 0.28×N rate, R^2=0.98) and significant differences between genotypes (Byrd 103.2 kg ha^{-1} and Snowmass 90.8 kg ha^{-1}). Grain yield and grain protein concentration were significantly and positively correlated with NY in the total aboveground biomass at anthesis. The N in the canopy is integral to a functional photosynthetic system for C assimilation and is the source for N remobilized to the grain (Kichey et al., 2007; Aranjuelo et al., 2012).

Post-anthesis N uptake has been observed to contribute to grain N concentration in France (Bogard et al., 2010; Kichey et al., 2007). However, in Mediterranean conditions in Iran, high temperatures and high evapotranspiration rates contributed to post-anthesis N losses (Bahrani et al., 2013). In this study, the analysis of variance showed no significant differences for post-anthesis N uptake between N rates, but there were differences among genotypes. We observed losses averaging over all treatments of 5.92 kg N ha^{-1} (6% of TDWa NY) in 2010-2011

during the post-anthesis period (Table 2.5). Negative values were observed for all genotypes except CO0940610. In 2011-2012, N losses were -10.3 kg N ha^{-1} for Byrd and -5.5 kg N ha^{-1} for Snowmass, 11.1% and 6.5% of NY in the total aboveground biomass at anthesis. Grain yield, grain protein concentration, NUE, and N harvest index were not correlated with post-anthesis N uptake. Nitrogen losses from crops are known to occur through gaseous release of NH_3 during plant growth (Raun & Johnsom, 1999). In a French study, negative values for post-anthesis N uptake were observed in some environments, but when genotypes were evaluated across 27 environments in France, post-anthesis N uptake averaged 23 kg ha^{-1} (Bogard et al., 2010). Our observations may be unique to the local environment, and may not reflect the genetic potential of these genotypes if grown in an environment with greater soil moisture.

Efficiency of biomass production and N recovery

Efficiency of biomass production and N recovery were investigated to identify physiological traits that contribute to NUE and grain N concentration. Correlation coefficients describe the associations among the traits (Table 2.7).

For 2010-2011, the analysis of variance revealed significant differences for the main effects of genotype and N rate for N use, uptake, and utilization efficiencies, but only the main effect of genotype for N harvest index and harvest index (Tables 2.5 and 2.6). The interaction term for genotype × N rate was not significant. For Snowmass and Byrd, the combined analysis of variance over two growing seasons showed a significant effect for year on NUE and N utilization efficiency, but not N uptake efficiency. There were no other significant effects for N utilization efficiency. In 2010-2011, the main effects for genotype and N rate were significant for NUE and N uptake efficiency. In 2011-2012, both main effects were significant for NUE and only genotype for N uptake efficiency.

Nitrogen use efficiency

For the 20 genotypes, NUE ranged from 39.9 kg kg^{-1} for 'RonL' (PI 648020) to 46.7 kg kg^{-1} for Byrd (Table 2.5). Averaged across genotypes, as the N rate increased, NUE decreased, ranging from 74.7 kg kg^{-1} with no applied N to 29.4 kg kg^{-1} with 112 kg ha^{-1} applied N (Table 2.6). The NUE component traits, N uptake and N utilization efficiencies, also significantly decreased as N rate increased (Table 2.6). This agrees with earlier studies and lends support to hypothesized decreasing C sink strength to the ear with N supply limitation (Ehdaie & Waines, 2001; Aranjuelo et al., 2012). In the combined analysis over two seasons for two genotypes, N utilization efficiency only differed by year across treatments (2010-2011: 47.8 and 42.5 kg kg^{-1}, 2011-2012: 34.1 and 28.9 kg kg^{-1}). Averaged over N rates, NUE decreased with increased N rate (2010-2011: 72.6 and 32.2 kg kg^{-1}, 2011-2012: 69.8 and 18.0 kg kg^{-1}). In the combined analysis for two genotypes in two seasons, N utilization efficiency only differed by year across treatments (2010-2011: 45.8 kg kg^{-1}, 2011-2012: 33.5 kg kg^{-1}). In 2010-2011, N uptake efficiency differed by genotype (1.1 and 0.90 kg kg^{-1}) and N rate (ranging for increasing N rate from 1.4 to 0.7 kg kg^{-1}), and just genotype in 2011-2012 (1.0 and 0.9 kg kg^{-1}). In multi-environment yield trials, Byrd has shown superior drought stress tolerance for grain yield, with a 10% yield advantage over the 3-yr average in the 2012 to 2014 CSU dryland variety performance trials (Haley et al., 2012a). The superior NUE and N uptake efficiency for Byrd may be important characteristics for stable and superior yields in a dryland environment.

Nitrogen uptake efficiency was more strongly correlated with NUE (r = .89) than N utilization efficiency (r = .47, Table 2.7). Byrd, with the highest grain yield, was most efficient for N uptake, but ranked 16th for N utilization efficiency, while Arlin had the lowest grain yield and ranked 19th for N uptake efficiency and fourth for N utilization efficiency. These results

suggest that with more efficient N uptake, N utilization is not limiting for grain yield. Efficient N uptake was not always sufficient to drive yield. 'Goodstreak' (Baenziger et al., 2004) ranked second for N uptake efficiency, last for N utilization efficiency and 19th for grain yield. Goodstreak is the only standard-height genotype evaluated, and was ranked third for total above ground biomass at anthesis, behind Byrd and 'Denali' (Haley et al., 2012b). Harvest index is a measure of conversion of canopy biomass to harvested grain. Byrd was ranked first for harvest index, while Goodstreak was ranked last. Despite its high N uptake efficiency, low values for harvest index and N utilization efficiency are associated with relatively low grain yield in Goodstreak.

Among the 20 genotypes, the contribution of variation in component traits to variation in the resultant trait (NUE) was determined through path analysis as the sum of cross products of each component trait and the resultant trait (Moll et al., 1982). Averaged across all treatments, contributions to variation were 52% from N uptake efficiency and 48% from N utilization efficiency. Averaged across genotypes, as N rate increased, the contribution of N utilization efficiency to variation for NUE decreased from 57% at 0 kg ha^{-1} N rate to 39% at 56 kg ha^{-1} N rate, remaining stable across the higher N rates, averaging 43% for the 56, 84, and 112 kg ha^{-1} N rates (Figure 2.3a). When there is limiting N, N utilization efficiency contributes more to variation in NUE. This observation suggests that the limiter or driver for grain yield production under limiting N is N utilization efficiency, while under conditions of sufficient or excess N, it is N uptake efficiency. As observed by Moll (1982), since the component trait that contributes most to NUE differs by the level of N supply, the N level of the selection environment may favor genotypes with superior values of one or the other component trait.

Performing the path analysis, while averaging across N rates, reveals genotypic variation in the component trait contributions (Figure 2.3b). The top five ranking genotypes for NUE were Byrd, Denali, 'Winterhawk' (PI 652927), 'Above' (Haley et al., 2003) and 'Hatcher' (Haley et al., 2005) (Table 2.5). They showed a wide range of component trait proportions: N uptake efficiency (57-73%, mean 66%) and N utilization efficiency (26-42%, mean 34%). The contributions of the component traits for the five least N efficient genotypes, Snowmass, Prairie Red, Goodstreak, Arlin, and RonL also ranged widely: N uptake efficiency (72-89%, mean 79%) and N utilization efficiency (12-28%, mean 21%) (Figure 2.3b). Group mean values for component trait contributions of the five top (65.5% N utilization efficiency) and five bottom (79.4% N utilization efficiency) ranked N use efficient genotypes were significantly different. On average, N utilization efficiency contributed relatively more to variance for the genotypes with the highest NUE, as though these genotypes were under more N limiting conditions. On the basis of this observation, it may be surmised that highly efficient genotypes could have experienced N limiting conditions at an N supply that was N sufficient for genotypes with the lowest NUE. In contrast to an earlier report of winter wheat grown in high-moisture conditions in southern England which reported independence of the traits (Barraclough et al., 2010), here N utilization and uptake efficiencies were negatively correlated at individual N rates (Table 2.7, 0 to 112 kg ha^{-1} [*r:* -.68, -.84, -.71, -.58, -.55]). Detecting positive-effect genes for both N uptake and N utilization may require selection under carefully managed N levels over multiple locations and years.

Nitrogen harvest index

For N harvest index, the main effect for genotype, but not N rate, was significant over a narrow range of values (Table 2.5). Nitrogen harvest index averaged 0.91 kg kg^{-1}, ranging from

0.82 kg kg^{-1} for Byrd to 1.01 kg kg^{-1} for CO940610. Nitrogen harvest index is positively correlated with grain protein concentration and negatively correlated with NUE, NY in the total aboveground biomass at anthesis, and N remobilization, consistent with other studies (Bahrani et al., 2013; Moll et al., 1982). Degradation of the photosynthetic systems during canopy senescence provides the main source of N for remobilization to the grain (Moll et al., 1982). As such, the processes that promote NUE for grain yield oppose those that increase N harvest index and grain protein concentration.

Grain protein deviation

Genotypes were ranked for the standardized residuals of the regression of grain yield on grain protein concentration, or 'grain protein deviation' (Figure 2.4). When estimated from multi-environment trial data, grain protein deviation identifies genotypes with grain protein concentration values that reliably deviate from the expected negative relationship with grain yield (Bogard et al., 2010; Guttieri et al., 2015). To obtain the trimmed linear model relating grain yield and grain protein concentration, genotypes with standardized residuals that exceeded the trimming threshold (|z|=1.96) were removed. The genotypes left out of the trimmed model were Arlin and 'Brawl CL Plus' (Haley et al., 2012c). The Pearson's correlation for grain yield (x_i) and grain protein concentration (y_i) increased from 0.65 with all data to 0.76 with the trimmed data set. Another round of trimming (|z|=1.96) removed Snowmass from the regression, increasing the correlation to 0.80. Additional rounds of trimming did not substantially change the position of the regression line, so the second set of trimmed data was used to calculate grain protein deviation for all genotypes. The regression equation ($y_i = 174.1\text{-}13.9x_i$) was applied to predict grain protein concentration (y_i) for each genotype (i). Grain protein deviation, as standardized residuals, was calculated for each genotype from these predicted values.

Grain protein deviation is strongly correlated with the residuals calculated from the final model (r=0.99). By applying coefficients of that association, it is found that the 5% threshold value (|z|=1.64) is a deviation from the mean grain protein concentration of 4.56 g kg^{-1}. Arlin (z=-1.81, -5.02 g kg^{-1}) and Brawl CL Plus (z=2.43, 6.73 g kg^{-1}) showed the minimum and maximum observed grain protein deviation values and were the only genotypes with low and high values that exceed the 95% threshold (Figure 2.4). Hard winter wheat pricing on the Kansas City Board of Trade (KCBT) is calculated based on grain protein concentration of 110 g kg^{-1}, with this pricing structure reflected in grain pricing at the elevators. Above or below that concentration, premiums may be granted, or discounts charged, with the scale specific to market conditions in each season. Selection for high values for grain protein deviation simultaneously identifies desirable bread wheat genotypes with high values for grain yield and grain protein concentration.

For winter wheat grown in high-moisture conditions in France, post-anthesis N uptake averaged 23 kg ha^{-1} across environments and years among 27 genotypes and was the main determinant of grain protein deviation (Bogard et al., 2010). Genomic regions associated with grain protein deviation were subsequently identified in related germplasm (Bogard et al., 2013). In the present study, post-anthesis N uptake was observed as N loss after anthesis for all but three genotypes and was not significantly correlated with grain protein deviation, nor with N harvest index (Tables 2.5 and 2.7). The average post-anthesis N uptake in the second year of the study was also negative (-7.9 g kg^{-1}). Other studies showed that grain protein deviation was associated with N accumulation before anthesis and N remobilization efficiency during senescence (Monaghan et al., 2001; Slafer et al., 1990), though this was not observed in this study. The

physiological traits associated with grain protein deviation under the agronomic conditions of this study may differ from those reported in the earlier studies.

Conclusions

Optimized wheat grain production requires N use efficient cultivars that are grown under careful N fertilizer management. We identified variation for NUE and its component traits among 20 hard winter wheat genotypes that were grown under dryland agronomic conditions in the west-central Great Plains of the United States. Nitrogen use efficiency ranged from 39.9 kg kg^{-1} for RonL to 46.7 kg kg^{-1} for Byrd. Averaged across genotypes, as the N rate increased, NUE decreased. By path analysis, we determined that under N sufficiency, variation in N uptake efficiency contributed more than N utilization efficiency to variation in NUE. With limiting N, N utilization efficiency contributed more to variation in NUE. Selection under sufficient N may distinguish genotypes with improved N uptake efficiency and selection under limiting N may distinguish genotypes with improved N utilization efficiency. Among elite adapted lines, yield limiting N supply will occur at moderate N application rates.

Grain protein concentration is an important contributor to milling and baking quality. Nitrogen fertilization was required to obtain 110 g kg^{-1} grain protein concentration, despite ample residual soil N and early season drought conditions in 2010-2011. A high yielding genotype, Brawl CL Plus, delivered grain protein deviation of 6.7 g kg^{-1}, a value that would be sufficient to produce adequate grain protein concentration and to protect growers from protein discounts on the commodity market. Post-anthesis N uptake was observed as N loss and was not associated with grain protein deviation.

Table 2.1. Hard winter wheat genotypes planted at Akron, CO in the 2010-2011 and 2011-2012 growing seasons.

Genotype†	Type‡	Origin§	Release Date	PI number¶	Pedigree
Above	HRW	CSU	2001	631449	TAM 110*4/FS2
Ankor	HRW	CSU	2002	632275	Akron/Halt//4*Akron
Arlin	HWW	KSU	1992	564246	HRW/HRS bulk selection
Bill Brown	HRW	CSU	2007	653260	Yumar/Arlin
Bond CL	HRW	CSU	2004	639924	Yumar//TXGH12588-120*4/FS2
Brawl CL Plus	HRW	CSU	2011	664255	Teal 11A/Above//CO99314
Byrd	HRW	CSU	2011	664257	TAM 112/CO970547-7
CO940610	HWW	CSU	unreleased	GSTR 10702	KS87H22/MWO9
Danby	HWW	KSU	2005	648010	TREGO/JGR 8W
Denali	HRW	CSU	2011	664256	CO980829/TAM 111
Hatcher	HRW	CSU	2004	638512	Yuma/PI 372129//TAM-200/3/4*Yuma/4/KS91H184/Vista
Goodstreak	HRW	UNL	2002	632434	SD3055/KS88H164//NE89646 (=COLT*2/PATRIZANKA)
Jagger	HRW	KSU	1994	593688	KS82W418/Stephens
Platte	HWW	Syngenta	1995	596297	N84-1104/Abilene
Prairie Red	HRW	CSU	1998	605390	CO850034/PI372129//5*TAM 107
Ripper	HRW	CSU	2006	644222	CO940606/TAM107R-2
RonL	HWW	KSU	2006	648020	Trego/CO960293
Snowmass	HWW	CSU	2009	658597	KS96HW94//Trego/CO960293
TAM 112	HRW	TXAM	2005	643143	U1254-7-9-2-1/TXGH10440
Winterhawk	HRW	WestBred	2007	652927	474S10-1/X87807-26//HBK0736-3

† Twenty genotypes were planted in the first season; two genotypes, Byrd and Snowmass, were planted in both seasons.
‡ Types of winter wheat cultivars: hard red (HRW), hard white (HWW).
§ Origin of the cultivar: Colorado State University, Fort Collins, CO (CSU); Kansas State University, Manhattan, KS (KSU); University of Nebraska, Lincoln, NE (UNL); Texas AgriLife (Texas A&M System) Research and Extension Center (TXAM); Syngenta, Junction City, KS (Syngenta); WestBred, a Unit of Monsanto Company, St. Louis, MO (WestBred).
¶ Plant Introduction numbers. Source: Germplasm Resources Information Network. 2015. USDA, ARS, National Genetic Resources Program, National Germplasm Resources Laboratory, Beltsville, Maryland. (http://www.ars-grin.gov/)

Table 2.2. Soil residual nitrate (NO_3-N) in the upper 60 cm of the soil profile at Akron, CO in the 2010-2011 and 2011-2012 growing seasons. Sampling was done before planting and after harvest.

Growing season	No. genotypes	Fertilizer N kg ha^{-1}†	Sample	Date	NO_3-N kg ha^{-1}‡
2010-2011 §	20	0, 28, 56, 84, 112	Sowing	14 Sept 2010	57.2
			Harvest	11 Aug 2011	36.7
2011-2012 ¶	2	0, 28, 56, 84, 112, 140	Sowing	28 Sept 2011	43.7
			Harvest	5 July 2012	19.3

† Nitrogen applied before planting in the form of urea (46-0-0)

‡ Total nitrogen concentration in the form of nitrate in the upper 60 cm of the soil profile

§ Organic matter averaged 1%; Phosphorus averaged 19.3 kg ha^{-1}

¶ Organic matter averaged 1.5%; Phosphorus averaged 22.4 kg ha^{-1}

Table 2.3. Adjusted means and the analysis of variance for twenty genotypes grown at Akron, CO in the 2010-2011 growing season and for two genotypes grown in the 2011-2012 growing season for plant height and heading date. Genotypes are arranged in order of decreasing nitrogen use efficiency.

	Plant Height (cm)						Heading Date (DOY)†					
Genotype‡	G mean	0	28	56	84	112	G mean	0	28	56	84	112
N mean	74.5	74.4	71.4	78.6	76.3	71.8	149.7	149.4	149.6	149.5	150.5	149.5
Byrd	74.9	74.8	71.6	79.8	73	75.2	148.1	147.8	147.7	147.8	149.5	147.8
Denali	74.3	72.3	72.6	76.4	81.1	69.1	153.9	154.7	152.4	154.2	155	153.3
Winterhawk	72.1	73.2	66.4	76.5	76	68.4	150.8	151	150.6	150.3	151.1	150.9
Above	76.1	78.2	71.6	78.9	76.4	75.4	148.2	147.6	148.3	147.9	149.2	148
Hatcher	73.7	74.5	67.1	80.6	80.4	66	150.8	150.8	150.2	150.8	151.5	150.9
Ripper	74.7	73.8	76	79.2	73.8	70.6	148.6	147.7	148.9	148.1	149.5	148.7
CO940610	73.9	71	80.3	74.5	77.6	66.4	148.8	148.3	148.6	148.8	149.6	148.8
Ankor	72.9	76.7	65.4	79.2	76.9	66.5	151.2	150.8	151.1	150.4	152.4	151.5
Brawl CL Plus	75.6	75.7	76.2	77.4	73.3	75.5	147.8	147.9	148.5	147.5	148.3	146.9
Jagger	75	78.7	69.4	79.4	72.1	75.2	148.6	148.8	148.3	148	148.9	149
Danby	79.2	76.4	68.9	91.3	84.3	75.2	151.9	151.6	151.6	151.6	152.7	151.9
Platte	72.9	69.8	73.2	74.5	68.9	78.1	152.5	151.8	152.7	152.7	153.2	152.4
Bond CL	73.4	71	70.8	80.1	77.9	67.2	148.5	148	148.6	148	149.6	148.3
Bill Brown	75.4	72.1	70.3	83	75.7	75.9	149.7	149.9	150	148.9	150	149.6
TAM 112	72.1	73.7	68.4	72.5	75.5	70.6	148.4	147.7	148.3	147.8	150.1	147.9
Snowmass	73.9	74.5	71	79.9	80.9	63.3	150.1	149.9	150.3	150.3	151	149.2
Prairie Red	75.9	77.2	72.3	73.7	72.8	83.5	147.5	147.1	147.2	147	149	147
Goodstreak	76.9	81.8	73.8	80.8	80.4	67.7	151.5	150.6	151.6	151	152.2	152.1
Arlin	74.2	74.7	71.6	74.5	75.2	75	147	146.9	147	147	147.1	146.8
RonL	73.1	68.7	72	79.4	74.2	71.1	150.4	150.1	149.9	152.5	150.6	149.1
lsd_G	ns§					lsd_G	**0.26**§					
lsd_N	ns					lsd_N	ns					
lsd_{GxN}	ns					lsd_{GxN}	ns					
2011 N mean	74.4	74.7	71.3	79.8	77.0	69.3	149.1	148.8	148.8	149.0	150.2	148.5
Byrd	74.9	74.8	71.6	79.8	73.0	75.2	148.1	147.7	147.7	147.7	149.3	148.0
Snowmass	73.9	74.5	71.0	79.9	80.9	63.3	150.1	150.0	150.0	150.3	151.0	149.0
2012 N mean	71.7	72.1	70.8	71.4	72.2	71.8	137.0	137.0	136.8	137.0	137.1	137.3
Byrd	70.7	70.2	68.9	71.8	72.1	70.8	136.5	136.5	136.3	136.5	136.8	136.5

	Plant Height (cm)						Heading Date (DOY)†					
Genotype‡	G mean	0	28	56	84	112	G mean	0	28	56	84	112
Snowmass	72.6	74.0	72.7	71.1	72.4	72.7	137.6	137.5	137.3	137.5	137.5	138.0
year¶		**2011**	**2012**				year¶	**2011**	**2012**			
lsd_G		ns	ns				lsd_G	**0.30**	**0.21**			
lsd_N		ns	ns				lsd_N	ns	ns			
lsd_{GxN}		**2.2**	**1.7**				lsd_{GxN}	ns	ns			

†Day of year, from 1 January (DOY).

‡Within columns, when genotype is a significant effect, mean comparison is done according to lsd (α =0.05).

§Bold font for the lsd values indicate significance at .05 probability level, respectively; ns is not significant.

¶Year was significant at $P < .05$.

Table 2.4. Adjusted means and the analysis of variance for twenty genotypes grown at Akron, CO in the 2010-2011 growing season and for two genotypes grown in the 2011-2012 growing season for physiological maturity and grain filling period. Genotypes are arranged in order of decreasing nitrogen use efficiency.

	Physiological Maturity (DOY)†						Grain Filling Period (days)‡					
Genotype‡	G mean	0	28	56	84	112	G mean	0	28	56	84	112
N mean	184.9	184.0	184.5	185.4	185.2	185.7	35.5	35.0	35.4	35.4	35.3	36.7
Byrd	186.3	185.2	186.7	187	185.7	186.8	38.5	38.4	39.3	39	36.3	39.7
Denali	186.4	186.2	185.9	187.4	185.6	186.7	32.9	32	34.7	33	31.3	33.6
Winterhawk	187.2	187.3	185.8	187.8	186.9	188	36.5	36.3	35.3	36.6	36.3	37.7
Above	183.4	182.5	182.8	184.1	183.8	183.6	35.6	35.3	35	36	35.7	36
Hatcher	186.8	186.2	183.5	187.5	187.4	189.3	36.3	36	33.7	36	37	38.7
Ripper	183.6	182.6	182.9	183.5	184.3	184.8	35.5	35.3	35	35.4	35.3	36.3
CO940610	186	184.7	185.7	186.3	186.4	186.9	37.8	37	38.3	37	37.7	39
Ankor	184.1	182.5	184.3	184.3	184.4	184.9	33.3	32.3	33.7	34	32.6	34
Brawl CL Plus	184.5	181.9	185.8	184.7	185.7	184.4	36.7	34.3	37.3	36.3	37.6	38
Jagger	183.1	181.9	183.6	182.8	183.7	183.8	34.9	33.7	35.7	34	35.7	35.7
Danby	185.7	185.9	185	187.1	184.2	186.4	34.2	34.7	34	35	32	35.3
Platte	185	182.9	184.5	184.6	186	186.9	32.5	30.8	32.7	31.7	32.6	34.6
Bond CL	185.2	183.6	184.4	185.9	185.5	186.5	36.9	36	36	37.7	36	39
Bill Brown	186.7	186.2	186	187.8	186.3	187.1	37.1	36.7	36.3	38	36.7	38
TAM 112	184.3	182.5	184.4	184	184.8	185.8	36.3	35.3	36.7	35.7	35	38.7
Snowmass	185	185.4	184	185.5	185.5	184.8	35.3	36.4	34.3	34.7	35	36
Prairie Red	182.9	180.5	183	183.8	184.6	182.7	35.7	34	36	36.3	36.3	36
Goodstreak	186.3	186.5	185.6	186.7	186.4	186.2	35.1	36.3	34.3	35.3	35	34.7
Arlin	182.6	181.5	182.1	181.9	183.7	183.9	35.8	35	35.3	34	37	37.7
RonL	184	183.2	183.3	184.8	184	184.8	33.8	33.7	33.7	32.3	33.7	35.7
lsd_G	**0.4**¶						**0.4**					
lsd_N	ns¶						ns					
lsd_{GxN}	ns						ns					
2011 N mean	186.0	186.2	185.7	185.8	185.8	186.3	36.9	37.3	36.8	36.8	35.7	37.8
Byrd	186.6	186.0	187.0	186.7	185.7	187.7	38.5	38.3	39.3	39.0	36.3	39.7
Snowmass	185.3	186.3	184.3	185.0	186.0	185.0	35.3	36.3	34.3	34.7	35.0	36.0
2012 N mean	165.6	165.0	165.0	165.8	166.5	165.8	28.6	28.0	28.3	28.8	29.4	28.5
Byrd	165.2	164.8	164.5	165.3	166.3	165.3	28.7	28.3	28.3	28.8	29.5	28.8

	Physiological Maturity (DOY)†						Grain Filling Period (days)‡					
Genotype‡	G mean	0	28	56	84	112	G mean	0	28	56	84	112
Snowmass	166.0	165.3	165.5	166.3	166.8	166.3	28.5	27.8	28.3	28.8	29.3	28.3
year#	**2011**	**2012**				year#	**2011**	**2012**				
lsd_G	**0.4**	**0.3**				lsd_G	**0.3**	ns				
lsd_N	ns	ns				lsd_N	ns	ns				
lsd_{GxN}	ns	ns				lsd_{GxN}	ns	ns				

† Day of year, from 1 January (DOY).

‡ Within columns, when genotype is a significant effect, mean comparison is done according to lsd (α =0.05).

§ Grain filling period was calculated from plot level data as the difference between physiological maturity and heading date.

¶ Bold font for the lsd values indicate significance at .05 probability level, respectively; ns is not significant

#Year was significant at $P < .05$.

Table 2.5. Adjusted means, summary statistics and the analysis of variance for twenty genotypes grown at Akron, CO in the 2010-2011 growing season for biomass and nitrogen yield and efficiencies. Genotypes are arranged in order of decreasing nitrogen use efficiency.

				Nitrogen Yield¶					Biomass Production and N Recovery Efficiencies#							
Genotype†	Grain Yield	GPRO ‡	TDWaY §	Shoot	Grain	Straw	NREM	PANU	NUE††		NUpE††		NUtE††		NHI	HI
	Mg ha^{-1}	g kg^{-1}	kg ha^{-1}	kg ha^{-1}					kg kg^{-1}							
Byrd	4.9	105.6	7363	110.6	92.0	2.12	102.1	-12.18	3.84	46.7	0.02	1.02	3.82	45.7	0.82	0.45
Denali	4.7	107.0	7140	106.5	89.4	0.99	102.4	-14.42	3.81	45.3	-0.03	0.97	3.82	45.8	0.86	0.41
Winterhawk	4.8	109.7	6683	100.9	94.0	1.61	93.8	-2.56	3.81	45.3	-0.08	0.93	3.86	47.4	0.91	0.43
Above	4.6	110.6	6522	93.8	90.7	0.97	90.7	0.42	3.79	44.4	-0.15	0.86	3.93	51.1	0.98	0.45
Hatcher	4.6	107.3	6659	108.9	88.4	1.78	101.0	-13.43	3.79	44.4	0.00	1.00	3.78	44.0	0.83	0.44
Ripper	4.6	113.1	6613	102.6	93.0	1.42	95.1	-3.77	3.78	44.0	-0.08	0.92	3.85	46.9	0.93	0.45
CO940610	4.5	112.5	5946	87.8	90.1	0.98	84.7	3.56	3.77	43.5	-0.21	0.81	3.96	52.6	1.01	0.44
Ankor	4.5	111.1	6715	101.3	89.1	1.74	98.4	-12.14	3.77	43.4	-0.07	0.93	3.82	45.7	0.88	0.43
Brawl CL Plus	4.5	118.4	6633	100.0	94.7	1.36	94.4	-1.08	3.77	43.2	-0.11	0.90	3.84	46.5	0.98	0.42
Jagger	4.4	117.1	6712	100.6	93.9	1.11	97.2	-3.33	3.76	43.2	-0.09	0.91	3.84	46.3	0.94	0.43
Danby	4.6	110.2	6460	94.3	89.8	1.22	90.6	-1.82	3.76	43.1	-0.15	0.86	3.91	49.8	0.96	0.42
Platte	4.5	113.1	6325	97.1	91.0	1.41	93.4	-3.57	3.76	43.1	-0.11	0.89	3.86	47.3	0.94	0.43
Bond CL	4.6	107.8	6426	102.0	87.6	1.76	94.8	-10.59	3.76	43.0	-0.07	0.93	3.81	45.2	0.85	0.43
Bill Brown	4.4	109.6	6089	100.3	85.9	1.48	95.8	-11.20	3.75	42.7	-0.09	0.92	3.83	46.0	0.87	0.42
TAM 112	4.5	112.4	6764	103.4	89.0	1.20	101.2	-11.41	3.75	42.4	-0.06	0.94	3.80	44.5	0.89	0.43
Snowmass	4.3	110.7	6604	93.0	86.4	1.93	86.8	-2.26	3.72	41.3	-0.15	0.86	3.85	47.1	0.91	0.41
Prairie Red	4.3	118.5	6463	94.9	91.0	1.30	90.1	-0.08	3.71	41.0	-0.14	0.87	3.84	46.5	0.97	0.43
Goodstreak	4.2	116.5	7132	108.1	88.4	1.10	106.6	-18.24	3.70	40.5	-0.01	0.99	3.70	40.4	0.83	0.38
Arlin	4.1	112.9	5797	89.0	83.5	1.39	81.5	1.72	3.69	40.1	-0.19	0.83	3.87	48.1	0.94	0.42
RonL	4.2	115.1	5993	91.0	85.8	1.27	85.3	-2.03	3.69	39.9	-0.18	0.84	3.83	46.2	0.93	0.42
mean	4.5	112.0	6552	99.3	89.7	1.41	94.3	-5.92	3.76	43.0	-0.10	0.91	3.84	46.6	0.91	0.43
min	4.1	105.6	5797	87.8	83.5	0.97	81.5	-18.24	3.69	39.9	-0.21	0.81	3.70	40.4	0.82	0.38
max	4.9	118.5	7363	110.6	94.7	2.12	106.6	3.56	3.84	46.7	0.02	1.02	3.96	52.6	1.01	0.45

				Nitrogen Yield¶					Biomass Production and N Recovery Efficiencies#							
Genotype†	Grain Yield	GPRO ‡	TDWaY §	Shoot	Grain	Straw	NREM	PANU	NUE††		NUpE††		NUtE††		NHI	HI
	Mg ha^{-1}	g kg^{-1}	kg ha^{-1}	kg ha^{-1}					kg kg^{-1}							
N rate	ns	***	ns	*	*	**	ns	ns	***		***		**		ns	ns
Genotypes (G)	***	***	***	***	***	***	***	***	***		***		***		***	***
N rate x G	ns‡‡	*	ns	ns	ns	ns	ns	ns	ns		ns		ns		ns	ns
lsd_G#	0.07	1.2	177.1	3.0	1.4	0.19	3.43	3.46	0.02		0.03		0.03		0.03	0.01

* Significant at .05 probability level.
** Significant at .01 probability level.
*** Significant at .001 probability level.
† Within columns, when genotype is a significant effect, mean comparison is done according to LSD (α=0.05).
‡ Grain protein concentration (GPRO).
§ Yield of above ground green biomass collected at anthesis (TDWaY).
¶ Nitrogen yield in the aboveground biomass at anthesis (Shoot), the harvested grain, the straw at maturity, remobilized N (NREM), and post-anthesis N uptake in the grain and straw (PANU).
# Trait values are N use efficiency (NUE), N uptake efficiency (NUpE), N utilization efficiency (NUtE), N harvest index (NHI), and harvest index (HI).
†† To stabilize variance before analysis, trait values were transformed by taking the natural logarithm. Those values are in the first column with back-transformed values in the second column for each trait.
‡‡ ns, not significant

Table 2.6. Adjusted means within and across five nitrogen rates (0, 28, 56, 84, 112 kg ha^{-1}) for grain protein concentration and N yields and recovery efficiencies at Akron, CO in the 2010-2011 growing season.

Trait†	Units	combined	0	28	56	84	112	lsd_N
Grain Protein	g kg^{-1}	112.0	95.5	106.2	112.2	121.6	124.3	***
Shoot NY	kg ha^{-1}	99.3	78.7	95.9	98.8	103.2	120.0	*
Straw NY	kg ha^{-1}	1.4	0.6	1.2	1.2	2.1	2.0	**
NUE‡	kg kg^{-1}	43	74	55	39	31	29	***
NUpE‡	kg kg^{-1}	0.9	1.3	1.1	0.9	0.7	0.7	***
NUtE‡	kg kg^{-1}	47	54	48	47	43	41	**

*,**,*** Significant at .05, .01, .001 probability level, respectively, ns not significant.
† Traits and use efficiencies are: grain protein concentration, shoot biomass N yield, straw biomass N yield, N use efficiency (NUE), N uptake efficiency (NUpE), and N utilization efficiency (NUtE).
‡ To obtain normally distributed values, trait values were transformed by the natural logarithm. Back-transformed values are presented, as well as significance of the least significant difference (lsd) between N rates.

Table 2.7. Correlation coefficients for nitrogen related traits at Akron, CO for 20 hard winter wheat genotypes in the 2010-2011 growing season at five nitrogen rates.

	Applied nitrogen (kg ha^{-1})					
Trait†	0	28	56	84	112	Combined
	Correlation coefficient					
NUE vs.						
NUpE (*r*)	0.80***	0.30**	0.73***	0.86***	0.57***	0.89***
NUtE (r_s)	ns	ns	ns	ns	ns	0.47***
NHI (r_s)	ns	ns	-0.39**	-0.41**	ns	-0.14**
NUtE vs.						
NUpE (*r*)	-0.68***	-0.84***	-0.71***	-0.58***	-0.55***	ns
NHI (*r*)	-0.16***	0.11***	-0.36***	-0.48***	0.05***	-0.17***
NHI vs.						
GN (*r*)	ns	ns	0.85**	0.85**	ns	0.70*
TDWaNY (*r*)	-0.65***	-0.76***	-0.82***	-0.83***	-0.74***	-0.65***
NREM (*r*)	-0.33*	ns	ns	-0.28*	-0.35*	-0.26***

*,**,*** Significant at 0.05, 0.01, 0.001 probability level, respectively; ns not significant.
† Pearson's (*r*) and Spearman's (r_s) coefficients are listed, as indicated. Trait name codes: NUE, N use efficiency; NUpE, N uptake efficiency; NUtE, N utilization efficiency; NHI, N harvest index; GN , grain N concentration; TDWaNY, biomass N yield; NREM, N remobilization efficiency.

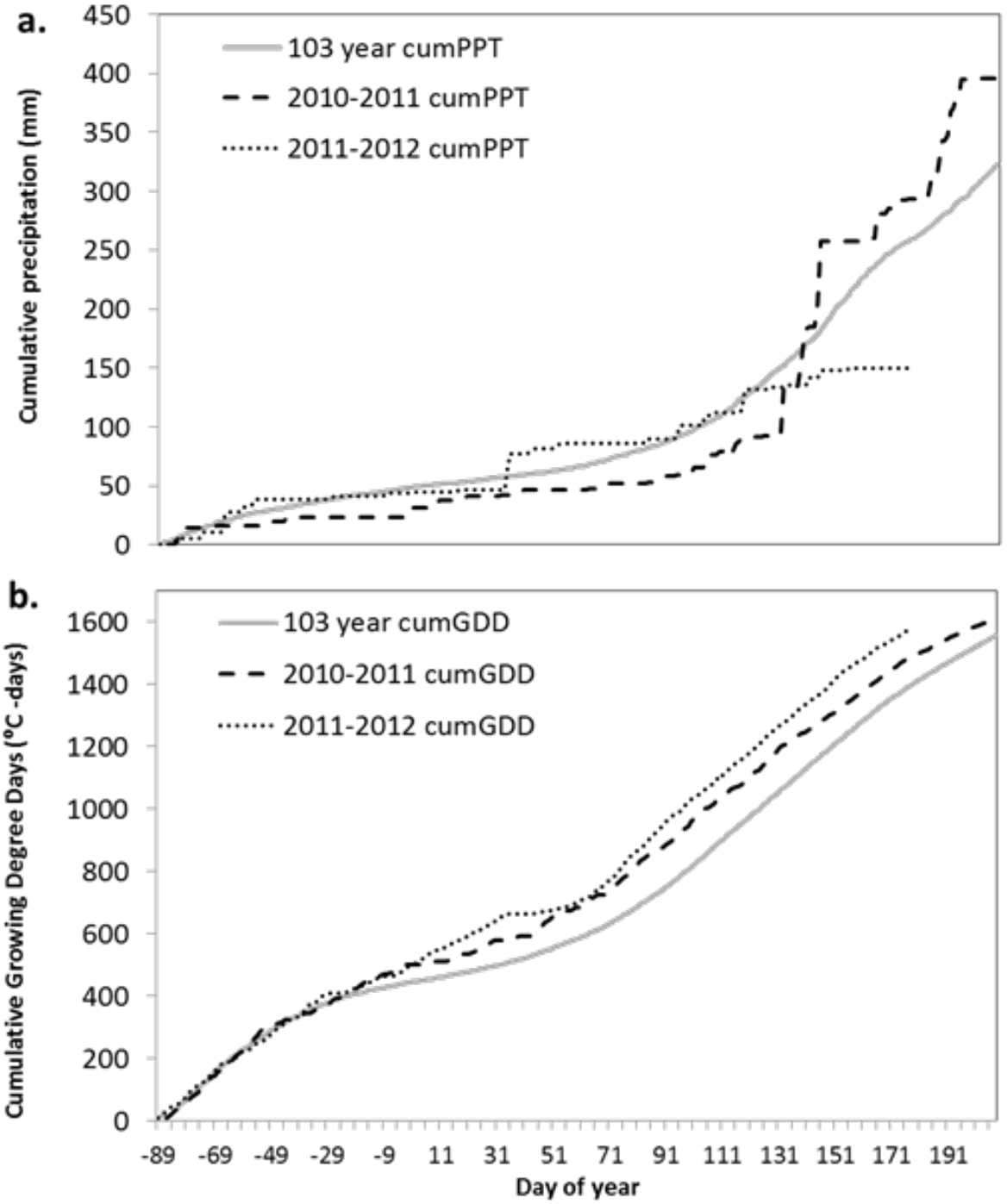


Figure 2.1. (a) Cumulative precipitation (cumPPT) and (b) growing degree days (cumGDD) at Akron, CO during 2010-2011, 2011-2012 and the 103 year cumulative averages for 1910-2012 (103 year cumPPT and 103 year cumGDD). Day of year (DOY) for average heading dates were 30 May 2011 (DOY 150) and 16 May 2012 (DOY 137). Average dates of physiological maturity were 4 July 2011 (DOY 185) and 14 June 2012 (DOY 166). The growing season begins with the planting date in the prior year (negative values for DOY) and extends into the harvest season in the following year (positive values for DOY).

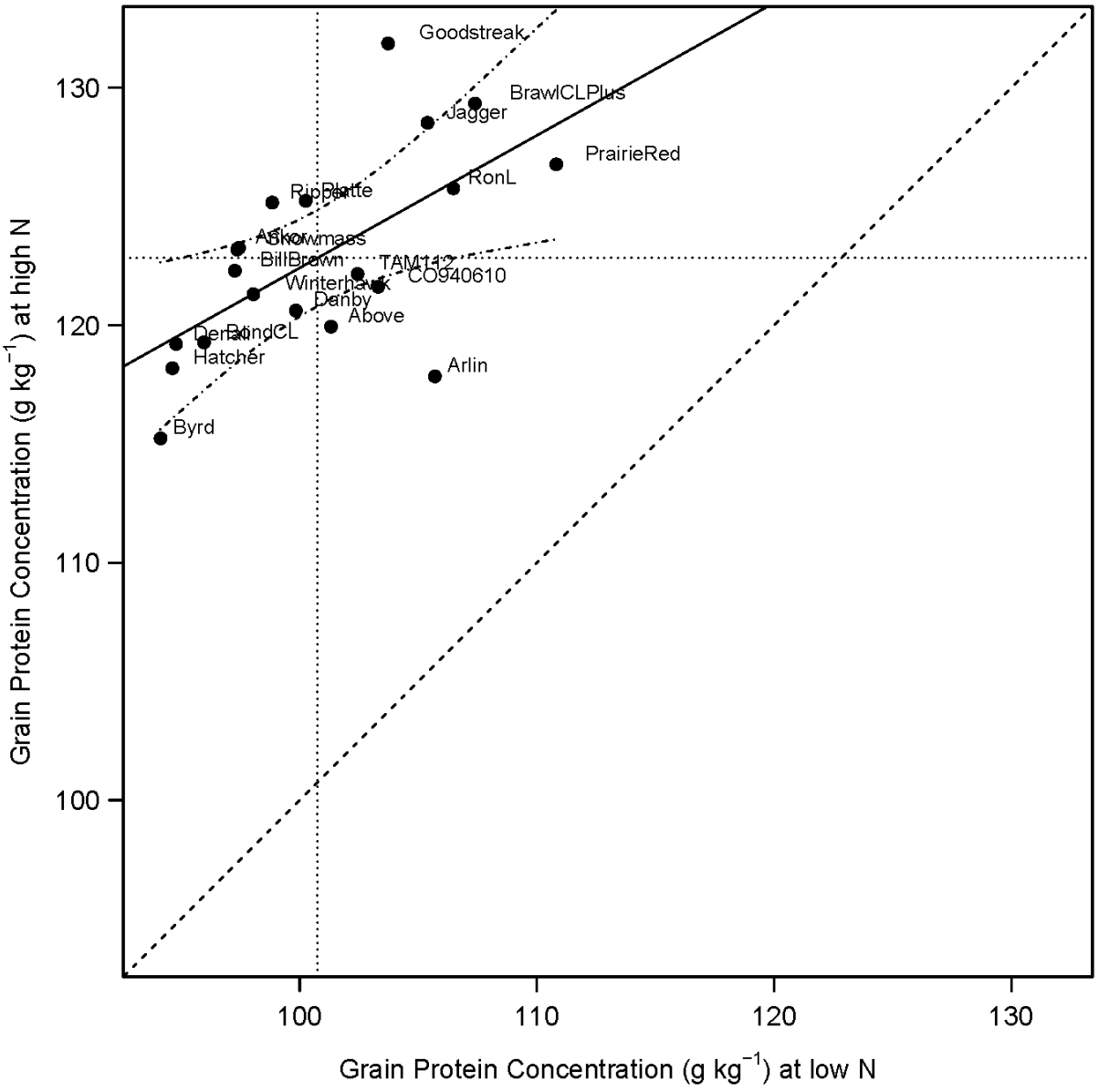


Figure 2.2. Interaction plot for grain protein concentration (GPRO, g kg^{-1}) adjusted means for 20 hard winter wheat genotypes grown at low (0 and 28 kg ha^{-1}) and high (84 and 112 kg ha^{-1}) applied N rates at Akron, CO in the 2010-2011 growing season. The least significant difference (α=0.05) for grain protein concentration was 1.4 g kg^{-1}. The dashed line drawn diagonally across the plot bisects the high and low values, with the intercept constrained to zero; the solid line is the least squares regression line (high N GPRO=66.6 + 0.56 × low N GPRO). Dash-dotted lines flank the 95% confidence interval. The horizontal and vertical dotted lines indicate the mean grain protein concentration at the high and low N rates.

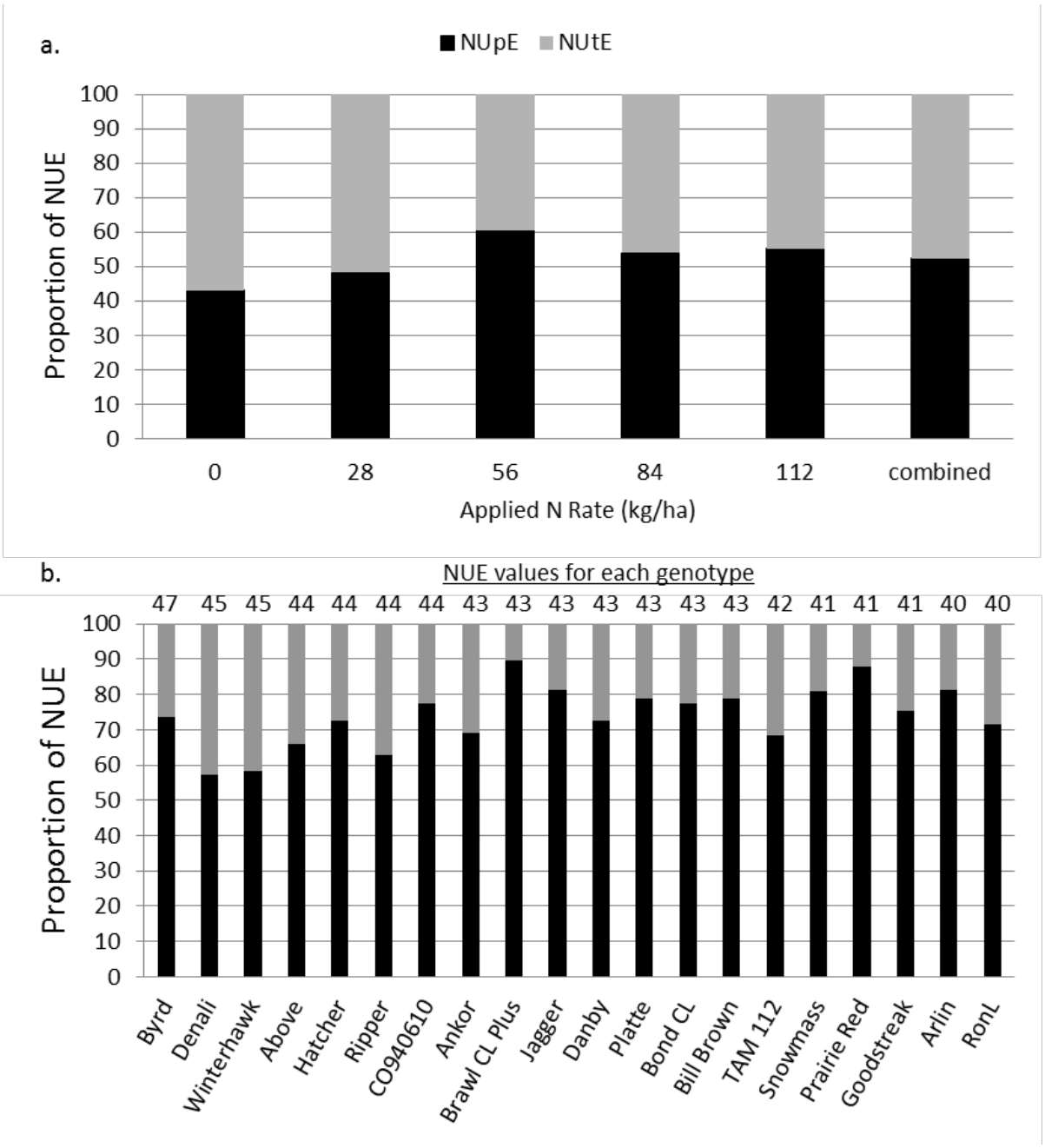


Figure 2.3. Contributions of nitrogen uptake (NUpE, kg kg^{-1}) and nitrogen utilization (NUtE, kg kg^{-1}) efficiency to variation in nitrogen use efficiency (NUE, kg kg^{-1}) in hard winter wheat (a) during the 2010-2011 growing season in Akron, CO among all genotypes within nitrogen rates or across all plots and (b) for each of 20 genotypes. Genotypes are arranged from highest to lowest values for nitrogen use efficiency.

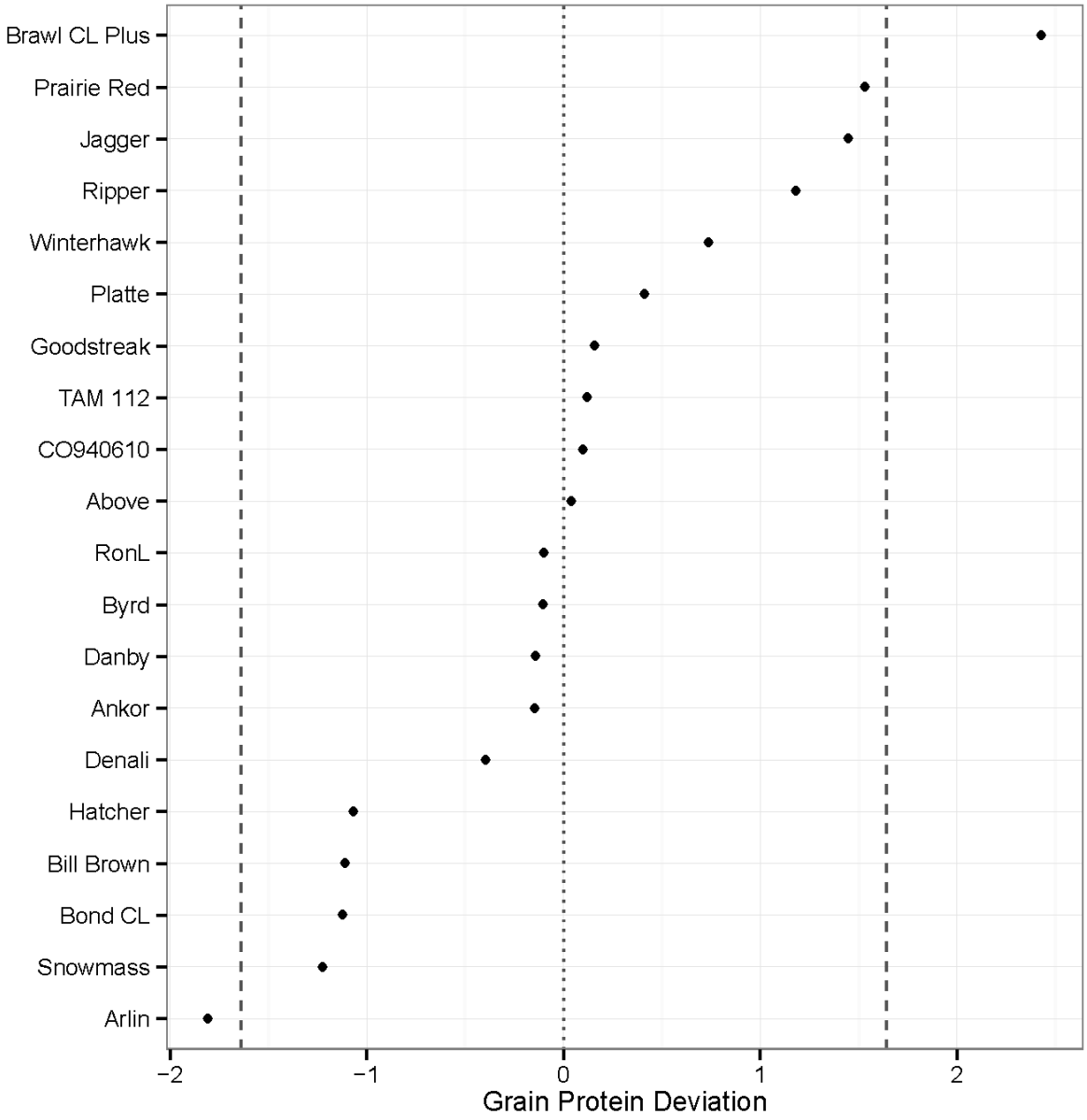


Figure 2.4. Grain protein deviation of 20 hard winter wheat genotypes grown in Akron, CO during the 2010-2011 growing season. Scores are positions on the standard normal distribution (z). The 95% threshold (|z|=1.64, 4.55 g kg^{-1}) is indicated by dashed lines.

References

AACC Approved Methods of Analysis, 11th Ed. Method 46-30.01. Crude Protein -- Combustion Method. Cereals & Grains Association, St. Paul, MN, U.S.A. http://dx.doi.org/10.1094/AACCIntMethod-46-30.01

Acreche, M. M., & Slafer, G. A. (2009). Variation of grain nitrogen content in relation with grain yield in old and modern spanish wheats grown under a wide range of agronomic conditions in a Mediterranean region. *Journal of Agricultural Science, 147*(6), 657.

Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics, 16*(1), 3-14.

Aranjuelo, I., Cabrera-Bosquet, L., Araus, J. L., & Nogues, S. (2013). Carbon and nitrogen partitioning during the post-anthesis period is conditioned by N fertilisation and sink strength in three cereals. Plant Biology, 15(1), 135-143. https://doi.org/10.1111/j.1438-8677.2012.00593.x

Baenziger, P. S., Beecher, B., Graybosch, R. A., Baltensperger, D. D., Nelson, L. A., Krall, J. M., Mcvey, D. V., Watkins, J. E., Hatchett, J. H., & Chen, M.-S. (2004). Registration of 'Goodstreak' wheat. *Crop Science, 44*(4), 1473-1474. doi:10.2135/cropsci2004.1473

Bahrani, A., Abad, H. H. S., & Aynehband, A. (2013). Nitrogen remobilization in wheat as influenced by nitrogen application and post-anthesis water deficit during grain filling. *Afr. J. Biotechnol., 10*(52), 10585-10594.

Barraclough, P. B., Howarth, J. R., Jones, J., Lopez-Bellido, R., Parmar, S., Shepherd, C. E., & Hawkesford, M. J. (2010). Nitrogen efficiency of wheat: Genotypic and environmental variation and prospects for improvement. *European Journal of Agronomy, 33*(1), 1-11.

Barraclough, P. B., Lopez-Bellido, R., & Hawkesford, M. J. (2014). Genotypic variation in the uptake, partitioning and remobilisation of nitrogen during grain-filling in wheat. *Field Crops Research, 156*, 242-248.

Bogard, M., Allard, V., Brancourt-Hulmel, M., Heumez, E., Machet, J. M., Jeuffroy, M. H., Gate, P., Martre, P., & Le Gouis, J. (2010). Deviation from the grain protein concentration–grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. *Journal of Experimental Botany, 61*(15), 4303-4312.

Bogard, M., Allard, V., Martre, P., Heumez, E., Snape, J., Orford, S., Griffiths, S., Gaju, O., Foulkes, J., & Le Gouis, J. (2013). Identifying wheat genomic regions for improving grain protein concentration independently of grain yield using multiple inter-related populations. *Molecular Breeding, 31*(3), 587-599. doi:10.1007/s11032-012-9817-5

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological), 26*(2), 211-252.

Chao, S., Zhang, W., Dubcovsky, J., & Sorrells, M. (2007). Evaluation of genetic diversity and genome-wide linkage disequilibrium among US wheat (*Triticum aestivum* L.) germplasm representing different market classes. *Crop Science, 47*(3), 1018-1030.

Cox, M. C., Qualset, C. O., & Rains, D. W. (1985). Genetic variation for nitrogen assimilation and translocation in wheat. I. Dry matter and nitrogen accumulation. Crop Science, 25(3), 430-435.

Desai, R. M., & Bhatia, C. R. (1978). Nitrogen uptake and nitrogen harvest index in durum wheat cultivars varying in their grain protein concentration. *Euphytica, 27*(2), 561-566. doi:10.1007/bf00043182

Dhugga, K., & Waines, J. (1989). Analysis of nitrogen accumulation and use in bread and durum wheat. *Crop Science, 29*(5), 1232-1239.

Ehdaie, B., & Waines, J. (2001). Sowing date and nitrogen rate effects on dry matter and nitrogen partitioning in bread and durum wheat. *Field Crops Research, 73*(1), 47-61.

El-Feki, W. M., Byrne, P. F., Reid, S. D., & Haley, S. D. (2015). Registration of CO940610/'Platte' wheat doubled haploid mapping population. *J. Plant Reg., 9*(3), 419-423. doi.org/10.3198/jpr2014.12.0085crmp

Gaju, O., Allard, V., Martre, P., Snape, J. W., Heumez, E., Le Gouis, J., Moreau, D., Bogard, M., Griffiths, S., Orford, S., Hubbart, S., & Foulkes, M. J. (2011). Identification of traits to improve the nitrogen-use efficiency of wheat genotypes. *Field Crops Research, 123*(2), 139-152. doi:10.1016/j.fcr.2011.05.010

Goos, R. J., Westfall, D. G., Ludwick, A. E., & Goris, J. E. (1982). Grain protein content as an indicator of N sufficiency for winter wheat. *Agronomy Journal, 74*(1), 130-133.

Guo, S., Zhu, H., Dang, T., Wu, J., Liu, W., Hao, M., Li, Y., & Syers, J. K. (2012). Winter wheat grain yield associated with precipitation distribution under long-term nitrogen fertilization in the semiarid Loess Plateau in China. Geoderma, 189–190, 442-450.

Guttieri, M. J., Baenziger, P. S., Frels, K., Carver, B., Arnall, B., & Waters, B. M. (2015). Variation for Grain Mineral Concentration in a Diversity Panel of Current and Historical Great Plains Hard Winter Wheat Germplasm. Crop Science, 55(3), 1035-1052.

Haley, S. D., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Heaton, E. E., Seifert, S. A., Kottke, R. A., Rudolph, J. B., Martin, T. J., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Seifers, D. L., Chen, M.-S., & Seabourn, B. W. (2011). Registration of 'Snowmass' wheat *J. Plant Reg., 5*(1), 87-90. doi:10.3198/jpr2010.03.0175crc

Haley, S. D., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Hudson, E. E., Seifert, S. A., Kottke, R. A., Valdez, V. A., Rudolph, J. B., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Chen, M.-S., & Seabourn, B. W. (2012a). Registration of 'Byrd' wheat. *J. Plant Reg., 6*(3), 302-305. doi:10.3198/jpr2011.12.0672crc

Haley, S. D., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Hudson, E. E., Seifert, S. A., Kottke, R. A., Valdez, V. A., Rudolph, J. B., Martin, T. J., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Chen, M.-S., & Seabourn, B. W. (2012b). Registration of 'Denali' wheat. *J. Plant Reg., 6*(3), 311-314. doi:10.3198/jpr2011.12.0675crc

Haley, S. D., Johnson, J. J., Westra, P. H., Peairs, F. B., Stromberger, J. A., Hudson, E. E., Seifert, S. A., Kottke, R. A., Valdez, V. A., Rudolph, J. B., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Chen, M.-S., & Seabourn, B. W. (2012c). Registration of 'Brawl CL Plus' wheat. *J. Plant Reg., 6*(3), 306-310. doi:10.3198/jpr2011.12.0673crc

Haley, S. D., Lazar, M. D., Quick, J. S., Johnson, J. J., Peterson, G. L., Stromberger, J. A., Clayshulte, S. R., Clifford, B. L., Pester, T. A., Nissen, S. J., Westra, P. H., Peairs, F. B., & Rudolph, J. B. (2003). Above winter wheat. *Can J Plant Sci, 83*(1), 107-108. doi:10.4141/P02-014

Haley, S. D., Quick, J. S., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Clayshulte, S. R., Clifford, B. L., Rudolph, J. B., Seabourn, B. W., Chung, O. K., Jin, Y., & Kolmer, J. (2005). Registration of 'Hatcher' wheat. *Crop Science, 45*(6), 2654-a-2656. doi:10.2135/cropsci2005.0030

Hanft, J. M., & Wych, R. (1982). Visual indicators of physiological maturity of hard red spring wheat. Crop Science, 22(3), 584-588.

Hawkesford, M. J. (2011). An overview of nutrient use efficiency and strategies for crop improvement. In M. J. Hawkesford & P. Barraclough (Eds.), *The molecular and physiological basis of nutrient use efficiency in crops* (pp. 5-19). Chichester, West Sussex: Wiley-Blackwell.

Hawkesford, M. J. (2014). Reducing the reliance on nitrogen fertilizer for wheat production. *Journal of Cereal Science, 59*(3), 276-283.

Joppa, L., & Cantrell, R. (1990). Chromosomal location of genes for grain protein content of wild tetraploid wheat. *Crop Science, 30*(5), 1059-1064.

Justes, E., Mary, B., Meynard, J.-M., Machet, J.-M., & Thelier-Huche, L. (1994). Determination of a Critical Nitrogen Dilution Curve for Winter Wheat Crops. Annals of Botany, 74(4), 397-407.

Keeney, D. R., & Nelson, D. W. (1982). Nitrogen—inorganic forms. In R. L. Westerman (Ed.), *Methods of soil analysis. Part 2. Chemical and microbiological properties.* (2 ed., Vol.

Agronomy Monograph, pp. 676-682). Madison, WI: American Society of Agronomy, Soil Science Society of America.

Kichey, T., Hirel, B., Heumez, E., Dubois, F., & Le Gouis, J. (2007). In winter wheat (*Triticum aestivum* L.), post-anthesis nitrogen uptake and remobilisation to the grain correlates with agronomic traits and nitrogen physiological markers. *Field Crops Research, 102*(1), 22-32. doi:http://dx.doi.org/10.1016/j.fcr.2007.01.002

Large, E. C. (1954). Growth Stages in Cereals: Illustration of the Feekes Scale. Plant Pathology, 3(4), 128-129. https://doi.org/10.1111/j.1365-3059.1954.tb00716.x

Le Gouis, J., Béghin, D., Heumez, E., & Pluchard, P. (2000). Genetic differences for nitrogen uptake and nitrogen utilisation efficiencies in winter wheat. *European Journal of Agronomy, 12*(3-4), 163-173. doi:10.1016/s1161-0301(00)00045-9

Littell, R. C., Stroup, W. W., Milliken, G. A., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models*.: SAS institute.

Marinciu, C., & Saulescu, N. (2009). Grain yield and protein concentration in winter wheat cultivars tested with and without nitrogen fertilizer. *Romanian Agricultural Research, 26*, 13-19.

McMaster, G. S., & Wilhelm, W. (1997). Growing degree-days: One equation, two interpretations. *Agricultural and Forest Meteorology, 87*(4), 291-300.

Moll, R. H., Kamprath, E. J., & Jackson, W. A. (1982). Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization. *Agronomy Journal, 74*(May-June), 562-564.

Monaghan, J. M., Snape, J. W., Chojecki, A. J. S., & Kettlewell, P. S. (2001). The use of grain protein deviation for identifying wheat cultivars with high grain protein concentration and yield. *Euphytica, 122*(2), 309-317.

Muurinen, S., Slafer, G., & Peltonen-Sainio, P. (2006). Breeding effects on nitrogen use efficiency of spring cereals under northern conditions. *Crop Science, 46*(2), 561-568.

Ortiz-Monasterio R., J. I., Sayre, K. D., Rajaram, S., & McMahon, M. (1997). Genetic progress in wheat yield and nitrogen use efficiency under four nitrogen rates. *Crop Science, 37*(3), 898-904. doi:10.2135/cropsci1997.0011183X003700030033x

Oury, F.-X., & Godin, C. (2007). Yield and grain protein concentration in bread wheat: How to use the negative relationship between the two characters to identify favourable genotypes? *Euphytica, 157*(1-2), 45-57. doi:10.1007/s10681-007-9395-5

Quick, J. S., Stromberger, J. A., Clayshulte, S., Clifford, B., Johnson, J. J., Peairs, F. B., Rudolph, J. B., & Lorenz, K. (2001). Registration of 'Prairie Red' wheat. *Crop Science, 41*(4), 1362-a-1363. doi:10.2135/cropsci2001.4141362-ax

R Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Version R x64 2.15.0. Vienna, Austria. Retrieved from http://www.R-project.org/

Raun, W. R., & Johnson, G. V. (1999). Improving nitrogen use efficiency for cereal production. *Agronomy Journal, 91*(3), 357-363.

Sadras, V., & Lawson, C. (2013). Nitrogen and water-use efficiency of Australian wheat varieties released between 1958 and 2007. *European Journal of Agronomy, 46*, 34-41.

SAS Institute Inc. (2010). SAS/STAT software, version 9.3 TS Level 1M0 of the SAS System for Windows x64 (Version 9.3 TS Level 1M0 of the SAS System for Windows x64). Cary, NC: SAS Institute Inc.

SAS Institute Inc. (2013). JMP® Pro version 11.0.0 Windows 7 Enterprise x64 (Version 11.0.0) [64-bit]. Cary, NC: SAS Institute Inc.

Sears, R. G., Martin, T. J., Cox, T. S., Chung, O. K., Curran, S. P., Heer, W. F., & Witt, M. D. (1997). Registration of 'Arlin' wheat. *Crop Science, 37*(2), 627-627. doi:10.2135/cropsci1997.0011183X003700020055x

Sinclair, T. R., & de Wit, C. T. (1975). Photosynthate and nitrogen requirements for seed production by various crops. *Science, 189*(4202), 565-567. doi:10.2307/1740987

Slafer, G. A., Andrade, F. H., & Feingold, S. E. (1990). Genetic improvement of bread wheat (*Triticum aestivum* L.) in Argentina: Relationships between nitrogen and dry matter. *Euphytica, 50*(1), 63-71. doi:10.1007/bf00023162

Soil Survey Staff (1999). Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys. 2nd ed. Natural Resources Conservation Service. USDA Handbook 436. USDA, Washington, DC.

Sosulski, F. W., & Imafidon, G. I. (1990). Amino acid composition and nitrogen-to-protein conversion factors for animal and plant foods. J*ournal of Agricultural and Food Chemistry, 38*(6), 1351-1356. doi:10.1021/jf00096a011

Uauy, C., Brevis, J. C., & Dubcovsky, J. (2006a). The high grain protein content gene Gpc-B1 accelerates senescence and has pleiotropic effects on protein content in wheat. *Journal of Experimental Botany, 57*(11), 2785-2794.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., & Dubcovsky, J. (2006b). A NAC gene regulating senescence improves grain protein, Zn, and Fe content in wheat. *Science, 314*(5803), 1298-1301. doi:10.1126/science.1133649

Van Sanford, D. A., & MacKown, C. T. (1986). Variation in nitrogen use efficiency among soft red winter wheat genotypes. TAG Theoretical and Applied Genetics, 72(2), 158-163. https://doi.org/10.1007/bf00266987

Wang, L., Cui, F. A., Wang, J., Jun, L. I., Ding, A., Zhao, C., Li, X., Feng, D., Gao, J., & Wang, H. (2012). Conditional QTL mapping of protein content in wheat with respect to grain yield and its components. *Journal of Genetics, 91*(3), 303-312. doi:10.1007/s12041-012-0190-2

Wang, R. F., An, D. G., Hu, C. S., Li, L. H., Zhang, Y. M., Jia, Y. G., & Tong, Y. P. (2011). Relationship between nitrogen uptake and use efficiency of winter wheat grown in the North China Plain. *Crop and Pasture Science, 62*(6), 504-514.

CHAPTER 3

STRATEGIES FOR SIMULTANEOUS IMPROVEMENT OF GRAIN YIELD AND GRAIN PROTEIN CONCENTRATION IN HARD WINTER WHEAT

Summary

Simultaneous improvement of grain yield (GY) and grain protein concentration (GPRO) is a central challenge for hard winter wheat (*Triticum aestivum* L.) breeding. The priority in this study was to counter the troublesome negative association between GY and GPRO by identifying an effective strategy for simultaneous selection. Within a population of hard winter wheat breeding lines and varieties grown across the target environment for the Colorado State University hard winter wheat breeding program during the 2012 to 2015 growing seasons, significant genetic variation for GY and GPRO was observed and a high negative association between these traits was confirmed. This motivated evaluation of selection strategies that may distinguish those lines that are high yielding and retain desirable GPRO levels. A set of protein-yield selection indices combined GY and GPRO trait values, including grain protein yield (GPY), grain protein deviation (GPD), and several equally weighted indices that summed pairs of standardized single trait values. Correlation analysis showed that selection on index values focused to differing extents on GY or GPRO. Genomic selection applied to index values (direct method) provided forward prediction accuracy ranging between r = .21 to .44 for the 2013 validation set. Index values were also calculated from univariate or bivariate genomic estimated breeding values (reverse method) and applied for selection. Selection strategies were compared

by comparing means and distributions of GY and GPRO phenotypic values in the selected fractions. Genomic selection strategies based on index values provided selection for improved GPRO without antagonistic selection on GY.

Introduction

Grain yield (GY) is the primary selection target in commercial hard wheat (*Triticum aestivum* L.) breeding programs, with milling and bread-making quality as important secondary selection targets. High-input agricultural systems optimize productivity of high yielding wheat varieties (Graybosch et al., 2014; Lollato & Edwards, 2015), but this often comes at the cost of decreased grain protein concentration (GPRO) due to a negative genetic correlation with GY (Nuttall et al., 2017; Simmonds, 1995). For hard wheat, flour primary constituents on a sample weight basis are carbohydrates in the form of starch (~ 70-75%), water (~ 14%), and proteins (~ 10-12%), with glutens, the main storage proteins, composing 80-85% of the total grain protein (Goesaert et al., 2005). Gluten proteins are the major determinants of bread-making quality through the visco-elastic structure that forms during dough mixing (Goesaert et al., 2005). The positive correlation of GPRO with overall baking quality includes significant genetic correlation with individual dough rheological traits (Michel et al., 2018). Bread-making quality is assessed using standardized methods for rheological measurements of dough strength, stability, and extensibility and by baking characteristic ratings such as water absorption, loaf volume, and crumb structure (Shelton et al., 2008). With low throughput and high unit costs, these assessments are most commonly done in later breeding cycles for relatively few advanced breeding lines. Grain protein concentration is an indirect selection criterion for end-use quality.

Relatively fast and inexpensive assessment methods are available that require seed amounts suitable for early generation testing of GPRO.

Simultaneous selection for negatively correlated traits requires careful breeding strategies to ensure that desired standards are met for all selection targets, within the resource capacity of the program and the desired rate of genetic improvement (Bernardo, 2010). Additional challenges for accurate assessment during early generation testing are capacity constraints as a function of limited replication and environment sampling (Vikas Belamkar et al., 2018). Breeding strategies commonly implemented for selection of negatively correlated traits include independent culling, where selection is applied to each trait in each breeding cycle, and tandem selection, where individual traits are selected in sequential cycles, with order and number of cycles based on relative importance (Hazel & Lush, 1942). Independent culling applied to negatively correlated traits slows simultaneous selection for both traits, since the highest scoring lines for one trait may be culled upon selection for the other trait. Similarly, when tandem selection is applied to one trait in the initial cycles, the improvement of the second trait is consequently limited in subsequent cycles (Schulthess et al., 2016). Alternatively, selection indices may enable efficient simultaneous selection for multiple targeted traits in each cycle, including those with negative correlations (Bernardo, 2010). Prior to selection, a single value is calculated for total or economic merit, with weighting factors determined through foreknowledge of relative heritability and economic or inherent values (Hazel & Lush, 1942). Genetic and environmental correlations require adjustments to the weights and the final index may balance favorable and unfavorable values across all traits (Hazel & Lush, 1942; Schulthess et al., 2016). Efficiency of genetic improvement is greater with index selection, but independent culling

presents a practical advantage when it is desirable to cull individuals based on traits expressed early in the life cycle, prior to availability of data on all traits (Hazel & Lush, 1942).

Protein-yield selection indices have been applied in cereal grain breeding with the objective to overcome the negative genetic correlation of GY and GPRO to accomplish simultaneous selection. Grain protein yield (GPY) is a measure of the total harvested seed N per unit area, calculated as the product of GY and GPRO. In cereal grain crops, maximal levels for GPY are typically achieved when GY is maximized, often at the cost of low GPRO (Simmonds, 1995). Grain protein yield is positively correlated with GY and, less strongly, with GPRO (Koekemoer et al., 1999; Simmonds, 1995) and may select lines that attain high GPY through high GY or high GPRO (Michel et al., 2019b). Applied to a diversity panel of triticale (× *Triticosecale*) lines, the correlation of GPY with GPRO was negative and the selected fraction trended to lines with lower GPRO and high GY, perhaps reflecting a breeding history that prioritized GY (Neuweiler et al., 2021). The correlation of GPY with N uptake and remobilization makes it an interesting target for indirect selection for improved nitrogen use efficiency (NUE) (Cormier et al., 2013).

Restricted selection indices hold one trait stable while providing selection pressure for the other trait (Kempthorne & Nordskog, 1959). Grain protein deviation (GPD) is a commonly reported protein-yield composite trait that is expressed as the standardized residual of the regression of GPRO on GY (Oury & Godin, 2007). It functions similarly to a restricted selection index by identifying genotypes that retain higher than expected GPRO across a range of GY, without penalizing GY (Guttieri et al., 2015; Iqbal et al., 2007; Monaghan et al., 2001; Oury & Godin, 2007; Rapp et al., 2018; Thorwarth et al., 2018). Post-anthesis N uptake is positively correlated with GPD in some production environments, again providing an avenue for indirect

selection for a difficult-to-measure component of NUE (Bogard et al., 2010; Guttieri et al., 2017). In a study of durum wheat (*Triticum turgidum* L. ssp. *durum*), selection on GPD among breeding lines was observed to select more strongly for GPRO than for GY, resulting in selection for a high proportion of lines with high GPRO, but low GY, while selecting on GPY had the opposite effect (Rapp et al., 2018). The effects of restricted selection indices were monitored through response to selection for GY, GPRO, and GPY (Michel et al., 2019a). The indices incorporated index weights derived from phenotypic or genomic covariances in combination with observed values for GY, GPRO, or GPY to distinguish lines that deviated from linear regression of pairs of traits, thus being equivalent in structure to GPD. When top-ranked lines from complementary indices were combined, Michel et al. (2019) observed only a marginal reduction in the trade-off between GY and GPRO for phenotypic selection, whereas genomics-based methods promoted positive responses for GY, GPRO, and GPY. The work was extended to restricted selection indices that included additional factors for dough rheological traits (Michel et al., 2019a). Index selection that combined GY with dough rheological traits reduced the negative selection pressure on GY relative to an index that included GY and GPRO. Similarly, application of a restricted selection index of GPRO and GPY combined with dough rheological parameters held GY stable while advancing end-use quality.

In Rapp et al. (2018), selection on GPD among durum wheat lines strongly emphasized GPRO at the cost of GY. To overcome this undesirable outcome, they evaluated equally weighted protein-yield indices that combined standardized values for GY and GPRO or GPD and yield deviation. These indices achieved the desired outcome of identifying genotypes that exceeded culling levels for both GY and GPRO (Rapp et al., 2018). Phenotypic correlations of index values with both GY and GPRO were positive, with indices differing in the influence on

one trait or the other. Following Rapp et al. (2018), application of equally weighted protein-yield indices to a diversity panel and bi-parental populations in triticale identified lines with high GY and GPRO (Neuweiler et al., 2021). A Hazel-Smith Index (Hazel, 1943; Smith, 1936) may be developed for correlated traits by estimation of total merit through the phenotypic and genotypic covariances, with consideration of economic weights. Such an index was compared to several restricted selection indices in two rye (*Secale cereale*) test cross populations (Schulthess et al., 2016). The Hazel-Smith index and the restricted selection indices each strongly influenced culling levels for GY and GPRO, providing opportunity to tailor the selection strategy to the breeding program objectives. In an effort to develop early maturing varieties with high GY and acceptable GPRO in a hard spring wheat program, two Hazel-Smith indices included GY and GPRO, plus either maturity or anthesis date (Iqbal et al., 2007). Except for anthesis date, the heritability of the indices was higher than for individual traits, indicating that genetic gain will be superior for index selection. The index that included anthesis date provided positive selection pressure for GY and GPRO, while the index that included maturity delivered positive selection for GPRO and negative selection for GY. In a simulation study to predict usefulness of cross combinations among a set of hard winter wheat breeding lines, parent selection for improving GY and quality traits was performed by the use of a Hazel-Smith index (Yao et al., 2018). In response to index selection, progeny populations were observed that had positive correlations between traits that showed a negative correlation in the parent population. Use of the selection index improved quality and GY and produced greater genetic variance in the progeny generation (Yao et al., 2018).

Many genes of small effect underlie quantitative traits such as GY and GPRO. Successful breeding strategies for quantitative traits enrich positive effect alleles in a population.

Historically, the enrichment was effected through phenotypic selection, but the advent of genomic-scale DNA sequencing has enabled marker-based selection for all chromosome segments that contribute to trait variation (Meuwissen et al., 2001). These methods use statistical modeling of genome-wide DNA marker effects to estimate breeding values for prediction of modeled phenotypes (Heffner et al., 2009). Genomic selection will complement conventional breeding strategies through optimization of field phenotyping resources and through earlier recycling of lines as parents in the next breeding cycle (Vikas Belamkar et al., 2018; Lado et al., 2018; Lozada et al., 2019). Similarly, culling of lines based on genomic predictions may be applied prior to end-use quality assessments, thus focusing laboratory resources on lines with the best predicted quality (Hayes et al., 2017). Enhanced genetic gain through improved assessment accuracy and reduced cycle times are additional potential benefits of genomic selection (Bassi et al., 2016).

Breeding programs are moving towards an era where genomics assisted breeding will be routine as supportive technologies become economic and available (Poland, Endelman, et al., 2012; Rasheed & Xia, 2019). Supportive technologies include cost efficiencies for genotyping relative to phenotyping, accelerated computing power, powerful statistical modeling tools, and development of reference genome sequences. Application of genomic selection in wheat breeding programs for agronomic, disease and insect resistance, productivity, and quality targets has been reported for both single trait models (Asoro et al., 2013; Vikas Belamkar et al., 2018; Charmet et al., 2020; Michel et al., 2018; Rutkoski et al., 2015) and multi-trait models (Hayes et al., 2017; Lozada & Carter, 2019; Michel et al., 2019a; Schulthess et al., 2016). Application of genome-wide molecular markers for simultaneous improvement of GY and GPRO in breeding

populations has been reported for durum wheat (Rapp et al., 2018), hard winter wheat (Michel et al., 2016; Michel et al., 2019a, 2019b), and rye (Schulthess et al., 2016).

Objectives of this study were to develop breeding strategies to enable simultaneous selection for improved GY and acceptable GPRO in the Colorado State University (CSU) winter wheat breeding program. Using field data from multi-environment trials conducted over four field seasons, phenotypic and genetic variance, heritability and correlation among the traits GY, GPRO, GPY, and GPD were determined. These data were used in selection strategies for simultaneous selection of GY and GPRO. We observed the response to selection for GY and GPRO after applying phenotypic single-trait selection, independent culling, and several protein-yield selection indices. We extended the methods to single-trait and multi-trait genomic selection.

Materials and methods

Environments and genotypes

Thirty-two environments (ENV) represented the CSU wheat breeding and genetics program target population of environments, including rainfed and irrigated nurseries during the 2012 to 2015 growing seasons (Table 3.1). Data were also derived from agronomic studies including N fertilization trials with limiting or replete N application and a drought study with limited irrigation. The combination of harvest year, location, and trial type produced 32 datasets (Table 3.2). These ENV sampled diverse agronomic conditions across the target environment for the CSU wheat breeding and genetics program. Plot level GY and GPRO were collected for a population of 790 breeding lines and released varieties (Supplementary table 1). The breeding lines were at several stages of development (Table 3.2): elite lines in 3rd year of evaluation

(CSU Elite), lines in 2nd year of evaluation (Advanced Yield Nursery, AYN), doubled haploid lines in the AYN (DH), and diverse lines from earlier CSU cohorts (training panel, TP). Each genotype was assigned a breeding cycle cohort based on its initial entry year in the AYN, as follows: the TP and lines in the 2012 CSU Elite trials into the '2012' cohort, the lines in the 2013 AYN into the '2013' cohort, and those in the 2014 AYN and 2015 CSU Elite in the '2014' cohort (Table 3.2).

Experimental design

Plots were planted in six rows with 23 cm spacing between rows and 30 cm spacing between adjacent plots. Experimental units (plots) for the N trials were 1.8 m long and 1.5 m wide, with 2.7 m^2 harvested area. All other trials were planted in the same row spacing, with plot length of 3.7 m and 5.5 m^2 harvested area. In the 2012 N trials, two replications of each entry were randomized in an augmented incomplete block, latinized row-column design (John & Williams, 1995) with varieties 'Byrd' (Haley et al., 2012a), and 'Denali' (Haley et al., 2012b) included as checks in about 8 percent of plots (4 percent for each variety). In the 2013 N trials, the row-column randomization of entries augmented a partially-replicated alpha-lattice design (Williams et al., 2011) where all repeated Byrd checks (about 6% of the plots), replicated checks, and 29% of the entries were included in both replications. Three released varieties (Byrd; 'Brawl CL Plus', (Haley et al., 2012c); 'Antero', (Haley et al., 2014) were included as replicated checks. The entries in the 2013 CSU Elite trials were randomized in two complete blocks arranged in a latinized row-column design (Williams, 1986). A set of 85 genotypes included in the 2014 cohort were initially tested in the 2014 AYN with 15 selected genotypes advanced for further testing in 2015. Plots with both GY and GPRO data recorded in a single replicate at one to three

2014 locations and a single replicate at three 2015 CSU Elite trial locations were included in the analysis. Trait values for the 85 genotypes were calculated as means across ENV.

Phenotypes

Grain weight and moisture concentration were recorded by an on-board measuring system (HarvestMaster, Logan UT) during combine harvest. Grain weight was adjusted to 120 g kg^{-1} (12%) moisture concentration for calculation of grain yield on a per area basis (GY, Mg ha^{-1}). Grain N concentration as a percentage of grain dry weight was estimated by near infrared reflectance spectroscopy (NIR) using a Foss DS2500TM Feed and Forage analyzer (Foss North America, Eden Prairie, MN)). A conversion factor of 5.7 was applied to calculate percent grain protein (Baker, 1979), followed by conversion to metric units for grain protein concentration (GPRO, g kg^{-1}).

At the plot level for all genotypes, GY and GPRO values were combined to calculate two protein-yield indices. Grain protein deviation (g kg^{-1}) values were calculated as the Studentized residuals of the linear regression of GPRO on GY by rank-based estimation in the *Rfit* package v.0.23.0 (Kloke & McKean, 2012) in R (R Core Team, 2016). This robust regression method iteratively ranks residuals via a dispersion function, and then re-fits the model until the fit converges. The method reduces the influence of extreme values on the estimation of GPD, analogous to the trimming algorithm applied in an earlier study (Oury & Godin, 2007). Grain protein deviation enables selection for lines that have higher than expected GPRO at a given GY level, across the range of GY. On a plot level basis, GPY (Mg ha^{-1}) was calculated as the product of GY (Mg ha^{-1}) and GPRO (g kg^{-1}).

Best linear unbiased predictors (BLUPs) and variance components for each trait were estimated within each of the 20 ENV that contained the 2012 and 2013 cohorts with the following mixed model:

$$y_{imn} = \mu + g_i + r_m + c_n + \varepsilon_{imn}$$

Here, the notation y_{imn} is the observed trait value for genotype(i) in row(m) and column(n), μ is the intercept, g_i is the random residual effect for genotype(i), r_m and c_n are row and column random effects that are independently and identically distributed (*iid*), with r_m and $c_n \sim N(0,\sigma^2)$, and ε_{imn} is the random error for genotype(i) in row(m) and column(n) with $\varepsilon_{imn} \sim N(0,\sigma_\varepsilon^2)$. Variance components were estimated for each environment using the restricted maximum likelihood algorithm (REML). The average number of replications per entry within a location was calculated (rep) and then broad sense heritability (H) on an entry mean basis was calculated according to Fehr (1987):

$$H = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{rep}}$$

Under the assumption that the trial environments are a random sample from a population of target environments, a two-stage procedure was applied to estimate 'true' genotypic values. First, within each of the 20 ENV that contained the 2012 and 2013 cohorts, best linear unbiased estimates (BLUEs) were derived for each ENV by modeling genotype as a fixed effect (Supplementary table 2). A set of homogeneous and heterogeneous residual variance structures (Table 3.3) were used to fit the spatial field variation for each ENV (Butler et al., 2009), with model selection guided by convergence and the Akaike Information Criteria difference (ΔAIC) between models (Burnham, 2010). Weighting factors for each ENV (w_p) were calculated as the

ratio of the average error mean square across ENV ($\overline{EMS}$) and the error mean square within an ENV (EMS_p), multiplied by the number of reps within an ENV (rep_p). In the case of partially replicated designs, the average number of reps across all entries was used.

$$w_p = rep_p * \frac{\overline{EMS}}{EMS_p}$$

In a second stage, the BLUEs and weights were subject to a combined analysis with genotypes and ENV as random effects to obtain best linear unbiased predictors (BLUPs) across ENV for each genotype:

$$y_{ijk} = \mu + g_i + t_j + l_k + w_p + \varepsilon_{ijk}$$

In this weighted model, y_{ijk} is the ENV BLUE, μ is the intercept, g_i is the random effect (BLUP) for genotype(i) with iid $g_i \sim N(0, \sigma_g^2)$, t_j and l_k are random effects for year and trial, with iid t_j and $l_k \sim N(0,\sigma^2)$, w_p is the single ENV weighting factor covariate, and ε_{ijk} is the random residual effect for genotype(i) in year(j) and trial(k) with iid $\varepsilon_{ijk} \sim N(0,\sigma_\varepsilon^2)$. Phenotypic correlations among traits were computed from the BLUPs. Mixed model analyses were performed using the package *ASReml-R* (Version 3, Butler et al., 2009) in the statistical software R (R Core Team, 2016).

Means comparisons for groups containing the top or bottom 20 ranked genotypes for GPD and GPY BLUPs were performed using the Student's *t*-test. Data visualizations, the Shapiro-Wilk test for Normality, and Levene's test for equality of variances were performed in preparation for choosing methods for means comparisons between groups. Equally weighted protein-yield selection indices were calculated according to the method of Rapp et al. (2018) by summing 'z-scores' for included traits. First, GY, GPRO, GPY, and GPD phenotypes (x_{ij}), as

BLUPs for the 2012 and 2013 cohort and as entry means for the 2014 cohort, were standardized and centered using the base R function *scale* (R Core Team, 2016). For individual *i* and trait *j* with trait mean μ_j and trait standard deviation $stdv_j$, 'z-scores' (z_{ij}) were calculated as:

$$center_{ij} = x_{ij} - \mu_j$$

$$z_{ij} = \frac{center_{ij}}{stdv_j}$$

Equally weighted selection indices, with weights equal to 1 or 3, were then constructed by summing scaled trait values in the following combinations: GY + GPRO, 3*GY + GPRO, GY + GPD, GY + GPY, GPY + GPD, and GPY + GPRO. The approximation of economic values and of phenotypic and genotypic correlations of included traits is contained within the index weights (Hazel and Lush, 1942). Genotypes that ranked in the top 20% of index values were retained in the selected fraction. The selection threshold was based on index values, while the breeding target was to identify genotypes with high GY and GPRO. To compare selection strategies, the selection response was determined from the selection differentials for standardized values for GY and GPRO. These were calculated as the difference of means between the selected fraction and the population.

Marker genotypes

Genome-wide markers were generated according to the genotyping-by-sequencing method (GBS, Elshire et al., 2011), as modified for a two-enzyme system (Poland, Brown, et al., 2012). During GBS library preparation, DNA is treated with methylation sensitive restriction endonucleases (*Pst*I and *Msp*I) which preferentially target gene-rich lower copy regions. Genomic DNA was extracted from bulked leaf tissue from 10 seedlings per line at the single leaf stage in a 96-well format using King Fisher 96 magnetic bead extraction kits on the King Fisher Flex Purification System (ThermoFisher Scientific Inc., Waltham, MA, USA). Multiplexing of

barcoded libraries was done at 96-plex or 192-plex, with a single blank well in each plate assigned at random for a sample tracking and cross-contamination control. Sequencing was carried out in a single lane for each multiplex library on an Illumina HiSeq2000 instrument (Informatics Research Core Facility, University of Missouri, Columbia, MO). Reference sequence-based single nucleotide polymorphism calls were made according to the TASSEL-GBSv1 Pipeline (Glaubitz et al., 2014). Single nucleotide polymorphism markers (SNPs) were anchored on the wheat reference genome, IWGSC RefSeq v1.0 (Alaux et al., 2018).

The GBS method randomly samples the genome, resulting in a sparse genotype matrix with missing values for individual SNPs across a population. Statistical models do not support missing values, necessitating marker imputation. The multivariate normal expectation maximization (EM) algorithm (Poland, Endelman, et al., 2012) was applied for marker imputation, retaining all markers with fewer than 30% missing values and a minor allele frequency that exceeded 3%. The assumption of multivariate normal distribution of marker genotypes is incorrect for biallelic markers, but the algorithm provides improved estimation of missing marker genotypes over simple mean estimation. Marker imputation and estimation of the realized relationship matrix (G matrix) was performed with the function *A.mat* from the package *rrBLUP* (Endelman, 2011) within the statistical software R (R Core Team, 2016). Principal components analysis (PCA) of the G matrix was performed within the base R function *eigen* (R Core Team, 2016).

Genomic mixed models

Univariate (UV) or bivariate (BV) genomic best linear unbiased prediction (G-BLUP) models (VanRaden, 2008) fitted BLUPs to predict genomic estimated breeding values (GEBV) for GY, GPRO, and several protein-yield indices. Prediction modeling was performed using the

package *ASReml-R* (Version 4, Butler et al., 2017) in the statistical software R (R Core Team, 2016). Multi-trait models for correlated traits may improve model accuracy through information sharing in the genotypic and phenotypic covariance matrices (Jia & Jannink, 2012). Bivariate models separately estimated GEBVs for pairs of traits, as follows: GY and GPRO, GY and GPD, GY and GPY, and GPY and GPD.

The G-BLUP model is analogous to the 'animal model' (Henderson, 1984), but with the pedigree relationship matrix (A matrix) replaced by the G matrix:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Zu} + \boldsymbol{e}$$

The input $\boldsymbol{y}$ is an $n_i \; x \; q_j$ matrix of BLUPs for genotype(i) and trait(j) obtained in the phenotypic mixed model analysis. The univariate model is a special case where $q = 1$. The design matrix $\boldsymbol{X}$ relates the fixed effects $\boldsymbol{b}$ to observations $(1_N b_i)$. The design matrix $\boldsymbol{Z_i}$ relates random genetic effects to observations. The vector $\boldsymbol{u}$ contains additive genetic effects (GEBVs) for all traits for each entry, with genetic variance $\sigma_u^2 \sim N(0, \boldsymbol{G})$. Residual errors $\boldsymbol{e}$ follow a multivariate normal distribution $\sim MN(0, \boldsymbol{I_n}\sigma_e^2)$, where the diagonal is a vector of zeros and the residuals are $n \; x \; n$ identity sub-matrices for each trait. The count of phenotypic records is the summation of all observations for each individual for each trait ($MN = \sum_1^q n_i$). The G matrix captures the correlation of random genetic effects among related entries through its off-diagonal elements (g_{ij}). Correlation and covariance structures which allow non-zero variance for the random genetic effects were evaluated, with model selection guided by the likelihood ratio test and model convergence status. In the bivariate models, unstructured heterogeneous covariance structures fitted the genotypic and residual covariance matrices. In the univariate models, a diagonal residual structure was fitted.

Variance components were estimated using the restricted maximum likelihood algorithm (REML) and were applied to calculate narrow sense heritability (h^2) for each trait as follows:

$$h^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

Model performance was assessed within repeated 5-fold cross-validation of the training data and by forward prediction of a validation data set. For cross-validation, the training data were randomly divided into five folds by applying the *sample* function in R (R Core Team, 2016). Phenotypes were masked in one fold and predicted by model parameterization using the other four folds. Five repeats of cross-validation were performed, each with a new random sampling for cross-validation folds. Model predictive ability within each cross-validation fold was calculated as the Pearson's correlation between the BLUPs and the GEBVs. Additionally, I calculated forward predictive ability in each fold for genotypes in a validation set. Correlation coefficients were averaged across five folds and then averaged across five repeats of cross-validation.

During advancement in the breeding program, pedigrees typically advance across years with parents preceding progeny in the testing nurseries. In 12 of the included 280 pedigrees, initial testing of a few of the sibling lines took place in a later growing season. To maintain a greater degree of independence of testing and validation datasets (Runcie & Cheng, 2019), the earliest tested year for a pedigree was assigned as the testing cohort for each sibling line, regardless of its actual initial entry in the trials (designated in column GRP_2 in Supplementary table 1). This was done to minimize biasing of model prediction accuracies through non-independence of training and validation data (Runcie & Cheng, 2019). To prepare a correlation matrix for visualizing relationships among GEBVs determined by different models, Pearson's

correlation coefficients were also calculated by the 'hold' method (Zhou et al., 2016) where the correlation was calculated after all the phenotypes had been predicted, rather than within each cross-validation fold ('Instant' method, Zhou et al., 2016). All analyses were performed within the statistical software R (R Core Team, 2016).

Selection strategies

The priority in this study was to counter the troublesome negative association of GY and GPRO by identifying an effective strategy for simultaneous selection. An optimal selection strategy would apply positive selection pressure for both GY and GPRO. Accordingly, after applying a selection method to rank genotypes, the top 20% were retained and then the method was evaluated according to its impact on the selection response for GY and GPRO. To compare selection strategies, the selection response was determined from the selection differentials for GY and GPRO. Independent culling is a selection strategy where thresholds are applied to each trait in each selection cycle to achieve the overall desired selection intensity. Independent culling was applied by first ranking genotypes for GY, retaining the top 40%, and then retaining the top 50% based on GPRO within the selected fraction to identify the top 20% overall. Equally weighted protein-yield indices were calculated from BLUPs prior to genomic selection modeling (univariate selection indices) and were also calculated from GEBVs derived from bivariate models (bivariate selection indices). In Schulthess et al. (2016), these methods were termed 'direct' and 'reversed', respectively. A selection threshold of the top 20%, based on percentile ranks of BLUPs or GEBVs, was applied for each single trait and protein-yield index. The Kruskal-Wallis test and the Dunnett's multiple comparison ($p < .05$) were applied to evaluate whether the distribution of standardized scores for GY and GPRO were the same among the selection cohorts for each single trait and protein-yield index. Dunnett's method contrasts to the

base population or to the selection cohorts obtained by independent culling were done to identify selection strategies that may enable simultaneous selection. The subset of methods which showed significant differences were then all compared by applying the Tukey procedure. Significance was tested at an alpha probability level of .05 for all means comparisons.

Results

Environment Characterization

Management practices in the context of seasonal weather conditions are strong determinants of GY and GPRO for the wheat crop produced in the Great Plains (Peairs & Armenta, 2010). Among the test environments in this study, substantial phenotypic variation for all traits reflected both genotypic differences and climatic impacts on crop productivity across the typically variable Colorado conditions. Detailed descriptive summaries of agronomic conditions are available in the annually published "Making Better Decisions" guides, accessed at the Colorado Variety Testing Program website (https://csucrops.agsci.colostate.edu/) or the Mountain Scholar repository (https://mountainscholar.org/handle/10217/100000). The 2012 wheat crop had a good start due to adequate soil moisture during stand establishment in the fall, but drought set in during the winter months with continuing drought conditions through harvest. Harvest was two to three weeks early across Eastern Colorado due to hot spring and summer conditions, resulting in reduced grain yield. The 2012 environmental averages for GY ranged from 2.5 Mg ha^{-1} to 4.1 Mg ha^{-1} and average GPRO ranged from 100 to 151 g kg^{-1}. In 2013, yields were also limited, but by a different pattern of environmental stressors. Overall, the conditions at planting were extremely dry, limiting crop establishment. Conditions remained dry through the winter and spring, but a cool spring improved the crop outlook. However, summer

drought again suppressed yields at the dryland sites, resulting in a range for average GY of 1.3 Mg ha^{-1} to 6.2 Mg ha^{-1} and average GPRO from 113 to 172 g kg^{-1}. Abundant September rains over much of Eastern Colorado provided good soil moisture for the 2014 crop establishment. The winter was dry and windy, but cool spring conditions, with adequate summer rains, conditioned high yields in most locations, with average GY ranging from 3.3 Mg ha^{-1} to 7.9 Mg ha^{-1} and average GPRO ranged from 108 to 128 g kg^{-1}. Warm and moist conditions favored crop establishment in the fall of 2014. Temperatures fluctuated widely during the winter and spring, promoting winterkill and freeze damage across the production area. Conditions were wet during May and hot during June through to harvest of the 2015 crop. Average grain yield across entries included in this study in the 2015 CSU Elite trial ranged from 4.0 Mg ha^{-1} to 6.5 Mg ha^{-1} and average GPRO ranged from 116 to 141 g kg^{-1} for the three locations.

Box and whiskers plots display the distribution of trait values recorded during harvest years 2012 to 2015 for genotypes that initially entered the performance trials during the 2012 to 2014 growing seasons (Figure 3.1). Summary statistics for plot level phenotypic data for measured and calculated trait values for GY, GPRO, GPY and GPD are recorded in Tables 4 and 5. All trait distributions within each ENV approximated the normal distribution. Within environments, the line of best fit for the ranked regression of GPRO on GY showed a negative slope in all but four ENV (Table 3.4). Across ENV (Figure 3.2), GY showed a negative correlation ($r = -0.48$, $p < .05$) with GPRO and a strong positive correlation with GPY ($r = 0.93$, $p < .05$), but was not correlated with GPD. The correlation coefficient for GPRO was positive with GPD ($r = 0.34$, $p < .05$) and negative with GPY ($r = -0.16$, $p < .05$), though with a lesser magnitude than for GPRO and GY. These data provided the platform to explore strategies for simultaneous selection for improved grain yield and acceptable grain protein concentration.

Mixed model analysis of phenotypic values

A two-stage mixed model analysis was used to estimate trait values for each of 676 check varieties and breeding lines from the 2012 and 2013 trialing cohorts. In the first stage, single ENV were analyzed in mixed models to obtain best linear unbiased estimates (BLUEs) for fixed effects of all included genotypes for GY, GPRO, GPY, and GPD (Supplementary table 2). The best models in all ENV included separate residual terms, estimated as two-dimensional spatial variance structures (Butler et al., 2009) to account for field variation within the row-column design of each nursery (Table 3.6). The distribution of BLUEs within ENV reflects the high genotypic variation for all traits (Table 3.6). Mean GY BLUEs ranged from 1.32 Mg ha^{-1} in the 2013 Akron DH-AYN to 6.94 Mg ha^{-1} in the 2014 Fort Collins CSU Elite trial. The variation for GY BLUEs within ENV illustrates the wide range of productivity environments included in this study (Figure 3.3). Mean GPRO BLUEs ranged from 100.0 g kg^{-1} in the 2012 Fort Collins irrigated NL-TP trial to 152.1 g kg^{-1} in the 2012 Akron NH-TP trial. Variance components, heritability estimates and weighting factors for each ENV are reported (Table 3.7). Heritability ranged from 0.12 to 0.81 for GY, 0.10 to 0.92 for GPRO, 0.06 to 0.89 for GPY and 0.12 to 0.79 for GPD. In the second stage, BLUPs for each genotype were determined within a combined analysis across 20 environments (Supplementary table 3). This analysis approximates peak performance for each genotype within the target environment, as represented by the sample of ENV included in this study (de la Vega et al., 2007). A bivariate plot of GY and GPRO BLUPs illustrates the strong negative relationship between these traits (Figure 3.4), a characteristic known to include an underlying genetic component (Simmonds, 1995). The scatter of the data around the regression line and the low R-squared value display the substantial variation in GPD, across the range of GY values.

Simultaneous phenotypic selection strategies for GY and GPRO

When independent culling is applied to traits with significant negative genetic correlation, selection may be mutually antagonistic, where the unit change in one trait will result in the opposite direction unit change for the other, scaled by the strength of the correlation (Bernardo, 2010). Consequently, given their strong negative correlation, selection for the highest yielding lines as a first step during independent culling will eliminate the lines with the highest GPRO levels and vice versa (Schulthess et al., 2016). The protein-yield index, GPD, distinguishes genotypes that have higher GPRO than expected at a given GY level and when calculated across diverse environments reliably classify genotypes as having positive or negative GPD (Oury & Godin, 2007). Genotypes with the highest and lowest rankings for GPD BLUPs were consistently negative or positive across diverse ENV (Figure 3.5). The variety Brawl CL Plus was ranked eleventh for positive GPD in this study among tested lines, consistent with the variety characteristics of acceptable GY and superior milling and baking quality (Haley et al., 2012c). Grain protein yield is a protein-yield index that considers productivity for GPRO as a function of GY. The variety 'Ripper' (Haley et al., 2007) was ranked thirteenth for GPY among tested lines, consistent with release notes that identified key characteristics of superior GY, baking and milling qualities (Haley et al., 2007). The protein-yield indices, GPD and GPY, showed opposite patterns for selection response for GY and GPRO (Figure 3.6). Selection on high positive values for GPD showed a negative selection response for GY and a strong positive response for GPRO (Figure 3.6, panels a and b). After selection based on GPD, the mean standardized score for the top 20 ranked lines for GY was slightly reduced (-0.161 Mg ha^{-1}) while GPRO was increased (9.23 g kg^{-1}). In contrast, selection on GPY (Figure 3.6, panels d and e) identified genotypes with superior GY (0.374 Mg ha^{-1}), while holding GPRO stable (0.58 g kg^{-1}). Grain protein deviation and GPY have complementary patterns of selection response for

GY and GPRO, but also show positive selection responses for each other (Figure 3.6, panels c and f). To further enhance simultaneous selection for GY and GPRO, additional selection strategies were explored that combined traits in additional protein-yield selection indices.

The objective of exploring protein-yield indices is to identify genotypes with superior values for both GY and GPRO through their correlated response to selection on the trait or index values. Trait values for 661 genotypes from the 2012 and 2013 cohorts were standardized and centered to scale to standard deviation units to produce 'z-scores'. Equally weighted selection indices were calculated by summing pairs of scaled traits: GY+GPRO, 3*GY+GPRO, GY+GPY, GY+GPD, GPY+GPD, and GPY+GPRO. The top 20 % of genotypes based on rankings of z-scores for single traits or selection indices identified the selected fraction. Selection differentials were calculated as the difference of the mean z-scores of the selected fraction and the selection candidates. The selection strategies were compared through the correlated response of GY and GPRO. The correlated response for GY (z = -0.62) after selection on GPRO or for GPRO (z = -0.68) after selection on GY shows the expected antagonistic responses and strong negative correlations (Table 3.8). Independent culling was applied by identifying the top ranked 20% for net merit by first culling the bottom 60% for GY, followed by culling the bottom 50% for GPRO among the retained lines. The mean standard scores for the selected fraction for GY (z = 0.78) and GPRO (z = 0.31) after independent culling showed a positive selection response for both, but the highest value lines for GY and GPRO were eliminated. Selection on GPY held GPRO stable (z= 0.08) and provided positive selection for GY (z = 1.09). On the other hand, GPD selected for increased GPRO (z = 1.34) and for decreased GY (z = -0.35). Both GPD and GPY included the individual with the highest GPRO value (z = 4.36) and GPY also retained the individual with the highest GY value (z = 2.87). Two

indices which summed GY with a second trait (GY + GPY and 3*GY + GPRO) showed positive selection for GY (z = 1.33 or 1.30), with negative selection on GPRO (z = -0.31 or -0.17). Combining GY with either GPD or GPRO created an index with positive selection response for both GY (z = 0.77 or 0.63) and GPRO (z = 0.65 or 0.84). Combining GPY with GPD or GPRO exerted strong selection pressure on GPRO (z = 0.95 or 1.06), with moderate selection pressure on GY (z = 0.40 or 0.30). Each of the equal weighted selection indices retained the individuals with the highest values for GY and GPRO. In Table 3.8, the selection strategies are ordered by the relative response to selection for GY and GPRO. At the extremes are selection on GY or GPRO alone, and in the middle are the indices that provide more uniform selection pressure for GY and GPRO. Choice of which protein-yield index to apply during selection will depend on breeder objectives during the breeding cycle. For example, it may be that early generation selection would emphasize simultaneous selection to retain diversity, with later cycles emphasizing selection for superior performance for GY and other traits essential to growers, millers, and bakers.

Marker genotypes

Marker genotypes were obtained for 775 breeding lines and released varieties, representing 280 unique pedigrees (Table 3.9). Markers with a minor allele frequency that exceeded 3% and with fewer than 30% missing values were retained, totaling 53,649 SNPs. The resulting *n x m* matrix of imputed marker scores was used to calculate the *n x n* realized relationship matrix (G, Supplementary table 4). The diagonal elements (g_{ii}) of the realized relationship matrix (G) approximate $1 + F_i$ with the inbreeding coefficient F_i approaching 1 for fully inbred lines (Endelman, 2011). The diagonal elements of the G matrix (g_{ii}) showed a bimodal distribution with a secondary peak that approached zero and a main peak centered on 2 (Figure 3.7). The secondary peak suggests the presence of heterozygosity for some lines, while

the peak centered on 2 would represent inbred lines approaching homozygosity. In the bottom panels of Figure 3.7, the distribution are displayed in overlapping histograms shaded by selection cohort. The genotypes with diagonal elements approaching zero are from the 2012 growing season, which primarily included breeding lines developed through modified bulk selection and inbreeding, mostly developed as F_3-derived lines. With this breeding history, these lines retain some level of residual heterozygosity. The 2013 pedigrees and 75% of the 2014 pedigrees were doubled haploid breeding lines, as reflected in peaks for the diagonal elements of those groups approaching 2. In the absence of genetic relationships between pairs of genotypes, the expectation for the off-diagonal elements of the G matrix (g_{ij}) is $\frac{1(1+f_i)}{n} \approx 0$. The observed distribution of the off-diagonal elements suggests the presence of shared ancestry, as expected, given the presence of sets of half-sib families among the 280 pedigrees (Supplementary table 5).

Principal components analysis of the G matrix provides graphical display tools for the genetic covariance among genotypes. The first seven principal components (PC) explained 80% of the variance contained in the G matrix (Figure 3.8). Together, the first two PCs explain 42.4% of the variance. A biplot of PC1 and PC2 with color-coding by year of entry in the trialing network shows some clustering of the genotypes within each selection cohort (Figure 3.9). Further examination of the content of the 2012 cluster centered in the biplot identified several pedigrees containing single crosses to parents from the Kansas State University breeding program. The dispersed cluster in the upper right quadrant was composed of several double haploid families present in the 2014 validation set and in the upper left quadrant from the 2013 validation set. The scatter of conventionally derived lines did not display obvious patterns by year. Released varieties from the Colorado State University breeding program and check varieties were present in all four quadrants of the PCA biplot.

To gain insight into the variance explained by that PC1 and PC2, the degree of linear relationship of the PCs with the trait BLUPs from the 2012 and 2013 selection cohorts was observed through correlation analysis (Figure 3.10). Principal component 1 was positively correlated with GPRO ($r = .7$, $p < .05$) and GPD ($r = .6$, $p < .05$) and negatively correlated with GY ($r = -.63$, $p < .05$) and GPY ($r = -.32$, $p < .05$). It represents variation that captures the negative association between GY and GPRO. On the other hand, PC2 is significantly and positively correlated with GY ($r = .42$, $p < .05$), GPRO ($r = .44$, $p < .05$), GPY ($r = .68$, $p < .05$), and GPD ($r = .57$, $p < .05$).

Squared cosines (cos2) show the importance of a PC for the position of a given genotype (Figure 3.11). The top 15 cos2 vectors cluster into 3 sets of half- or full-sib pedigrees. The underlying pedigree clusters contribute to target market characteristics, adaptation, and productivity. The cluster with high PC1 values includes the variety Snowmass (Haley et al., 2011) and its half-sib progeny and may share features that contribute to dough mixing strength. Snowmass is marketed within the Colorado Wheat Research Foundation (CWRF) Ardent Mills Ultragrain® Premium Program (Haley et al., 2017) and has superior dough strength due its allelic combination at the Glu-B1 and Glu-D1 high molecular weight glutenin genes (Cooper et al., 2016). The opposing clusters along the PC2 axis represent half-sib families related to the varieties Denali and Byrd. Each of these varieties was released due to the combination of superior yield, sufficient milling and baking quality, and stability across the target environment. Their opposing positions on the PC2 scale suggest that diverse allelic combinations accomplished the selection targets.

Prediction accuracy

Univariate and bivariate models for measured traits and protein-yield indices were used to derive genomic estimated breeding values (GEBVs) through cross-validation and forward prediction. Model training and validation sets included: 1) 405 lines that first entered evaluation trials in 2012 predicted the 2013 cohort (N=255), and 2) 660 lines evaluated in 2012 and 2013 predicted the 2014 cohort (N=85). For each model, five repeated cycles of cross-validation (k=5) produced GEBVs that were averaged within cycles, then across repeats. For the univariate models, predictive ability in cross-validation was high, ranging from r = .67 to .79 (Table 3.10). Narrow sense heritability ranged from h^2 = .20 to .39 for the univariate traits. The highest h^2 was observed for the protein-yield index, GPY, when predicted within the 2012 or 2012+2013 training set (r = .39 and .34), and the lowest for GPRO in the 2012 training set (r = .20), with marked improvement with the addition of the 2013 data (r = .31). Forward predictive ability for GY and GPRO of the 2013 validation set was substantially higher (r = .29 and .45) than for predictions of the 2014 validation set (r = .16 and .19). Genomic estimated breeding values for protein-yield selection index values were predicted in univariate models using index values calculated from phenotypes as the vector of observations ('direct' method, Schulthess et al., 2016). Predictive ability for the 2013 validation set ranged from r = .21 to .44, but predictive ability approached null or was negative for the 2014 validation set (r = .-.19 to .08).

Correlated traits in multi-trait models are theorized to improve predictive ability, particularly when a low heritability trait is combined with a high heritability trait (Jia & Jannink, 2012), as demonstrated for sensor-based traits applied in multi-trait models to improve wheat GY predictive ability (Crain et al., 2018). Other studies in wheat have shown little or no improved prediction accuracy for multi-trait models that apply balanced data in the training set to predict

validation lines (Lado et al., 2018; Ward et al., 2019). Multi-trait models showed superior predictive ability in scenarios where the validation set contained phenotypes for the second trait (Lado et al., 2018; Ward et al., 2019), or where balanced training data for a heritable correlated trait enabled imputation of missing target trait phenotypes (Schulthess et al., 2016). We compared predictive ability obtained for a target trait in a bivariate model to that obtained in a univariate model by calculating their ratio ('FP ratio', Table 3.11). For forward prediction of the 2013 validation set, predictive ability in the bivariate models was improved over univariate model prediction for GY when the paired trait was GPY (1.024), but not GPRO (0.958) or GPD (0.961). The predictive ability for GPY was improved when the paired trait was GY (1.021), but not GPD (0.975). For prediction of 3*GY paired with GPRO in the bivariate model, predictive ability improved for 3*GY (1.028), but did not change from the univariate model for GPRO (1.000). For forward prediction of the 2014 validation set, predictive ability for GY was slightly better in the bivariate model when the paired trait was GPD (1.006), as was GPY in the bivariate model with GY (1.013), while all other bivariate models showed no improvement. For the bivariate model combining GY and GPY or GPD, genetic covariance was positive in sign for prediction of the 2013 validation set and negative for prediction of the 2014 validation set. Heritability and genetic covariance contribute to efficiency of multi-trait predictions with some simulated data sets (Jia & Jannink, 2012). Here, the trait heritability estimates varied over a narrow range across all bivariate models (Table 3.11, h^2 = .19 to .40). Genetic covariance between traits included in the bivariate models ranged between -.37 to .41, except when modeling GY+ GPY in the 2012+2013 training set where we observed genetic covariance of .88. The GY+GPY bivariate model was superior to the univariate model ('FP ratio' = 1.024). In the observed ranges of genetic covariance and trait heritability, in some, but not all cases, the

bivariate model provided superior predictive ability. Another consideration for the efficiency of bivariate models is their output of a vector of GEBVs for all modeled traits. The obtained GEBVs can be applied for selection on individual traits, for selection by independent culling, or may be combined into selection indices. In the context of equivalent or improved prediction ability, obtaining GEBVs for correlated traits in a multi-trait model is an opportunity to simplify computing and data management processes by running fewer prediction models, with the opportunity to improve prediction accuracy.

Correlation analysis of predicted values

In preparation for choosing simultaneous selection strategies applied to the GEBVs, the 'hold' method (Zhou et al., 2016) was applied to calculation of Pearson's correlations between BLUPs and GEBVs obtained in univariate and bivariate models. All of the univariate GEBVs are highly correlated with the related BLUPs. Grain yield-related GEBVs (Figure 3.12) were all moderately and positively correlated with GY BLUPs (r = .56 to .66, p < .0001). The GY GEBVs obtained in the bivariate models were very nearly the same as those obtained in the univariate model (r = .97 to 1, p < .0001). Grain protein concentration-related GEBVs, but not GPY (Figure 3.13) were all moderately and positively correlated with GPRO BLUPs (r =.69 to.73, p < .0001). The GPRO GEBVs obtained in the bivariate models were very nearly the same as those obtained in the univariate model (r = .95 to 1, p < .0001). Grain protein yield GEBVs obtained in the univariate model (Figure 3.14) were highly correlated with GY GEBVs (r = .77, p < .0001) and weakly correlated with GPRO GEBVs (r = .07, p < .01). Similarly, when GPY was predicted with bivariate models, it was moderately correlated with GY GEBVs (Figure 3.12, r = .56, p < .0001) and weakly correlated with GPRO GEBVs (Figure 3.13, r = .14, p < .0001). Grain protein deviation GEBVs obtained in the univariate model (Figure 3.14) were

negatively correlated with GY GEBVs (r = -.33, p < .0001) and highly correlated with GPRO GEBVs (r = .97, p < .0001). Similar results were obtained when GPD was predicted with bivariate models, it was highly correlated with GPRO GEBVs (Figure 3.13, r = .70, p < .0001) and negatively correlated with GY GEBVs (r = -.31 or r = -.28, p < .0001, data not displayed). The GEBVs for GY and GPRO are negatively correlated (Figure 3.14, r = -0.52, p < .0001), so selection on these alone would be mutually antagonistic. Grain protein yield and GPD can be seen as restricted selection indices, where for GPY, strong selection pressure is applied for GY, while GPRO is held steady, and for GPD, selection pressure is strong for GPRO while GY is held steady (Michel et al., 2019a). Our results support this assertion (Figure 3.14): correlation of GPY with GY (r = 0.77, p < .0001) and GPRO (r = 0.07, p < .01) and of GPD with GY (r = -0.33, p < .0001) and GPRO (r = 0.97, p < .0001) applied positive selection for one trait, while holding the other stable, or reducing the negative selection pressure.

When applied in phenotypic selection, several equal weighted selection indices were found to retain the selection candidates with the highest values for GY and GPRO, but each differed for the relative response to selection for GY and GPRO (Table 3.8). The index values were entered in univariate genomic prediction models to evaluate their performance for simultaneous prediction of GY and GPRO. Correlation analysis of the GEBVs for the index values with univariate single trait GEBVs is presented in Figure 3.14. The GEBVs for selection index GY+GPRO were positively correlated with all four single trait GEBVs (GY, r = .32; GPRO, r = .64; GPY, r = .77; GPD, r = .77, all p < .0001). The GEBVs for selection index GY+GPD were positively correlated with all four single trait GEBVs (GY, r = .45; GPRO, r = .52; GPY, r = .85; GPD, r = .69, all p < .0001). The GEBVs for selection index GPY+GPD were positively correlated with all four single trait GEBVs (GY, r = .22; GPRO, r = .70; GPY, r = .76;

GPD, r = .83, all p < .0001). The GEBVs for selection index 3*GY+GPRO were positively correlated with three single trait GEBVs (GY, r = .91; GPY, r = .93, all p < .0001; and GPD, r = .09, p < .01) and negatively correlated with GPRO (r = -.11, p < .001). As was seen for phenotypic selection, this index performed as a restricted index, similarly to GPY, by being highly correlated with GY and holding GPRO relatively stable.

Simultaneous genomic selection strategies for GY and GPRO

The core interest in this study was to develop strategies for simultaneous selection for GY and GPRO. To this end, the performance of the genomic predictions was further evaluated based on the indirect selection response for GY and GPRO in response to selection on predicted values. The distribution of GY and GPRO within the selected fraction were determined for each selection strategy. Standardized phenotypic values (z-scores) were calculated for GY and GPRO BLUPs in the validation set to convert the values to the same scale. Summary statistics for the GY and GPRO z-scores are reported for the 2013 validation set and selected fractions (Table 3.12) and the 2014 validation set and selected fractions (Table 3.13). The Kruskal-Wallis test (Kruskal & Wallis, 1952) (α = .05) was applied to compare GY and GPRO z-score distributions in the selected fractions to the base distribution in the validation set. The distribution of GY in the selected fractions did not differ from the validation set, while the distribution of GPRO differed from the validation set for at least one selection strategy. Next, the Dunnett's multiple comparison procedure (α < .05) (Dunnett, 1955) was applied to identify which selection strategies impacted GPRO by contrasting the mean GPRO for each selected fraction to the mean of the validation set. For selections applied to the 2013 validation set (Table 3.12), three selection strategies showed a selection response for GPRO. The univariate selection index GPY+GPD increased mean z-score for GPRO from -0.19 in the validation set to 0.60 in the

selected fraction ($D = 0.79$, $\alpha = .001$). Bivariate GEBVs for GPY and GPD were summed to construct a vector of selection index values that were applied to obtain the 20% selected fraction. The mean z-score for GPRO increased from -0.19 in the validation set to 0.52 in the selected fraction ($D = 0.71$, $\alpha = .001$). When bivariate GEBVs for GPY and GPD were used for independent culling to obtain the 20% selected fraction, the mean z-score for GPRO increased from -0.19 in the validation set to 0.34 in the selected fraction ($D = 0.53$, $\alpha = .01$). For selections applied to the 2014 validation set (Table 3.13), the univariate selection index 3*GY+GPRO reduced GPRO in the selected fraction ($D = -0.86$, $\alpha = .05$). Forward predictive ability of all indices (Table 3.10) and bivariate traits (Table 3.11) for the 2014 validation set was consistently low. The low predictive ability for the 2014 validation set may stem from not adequately accounting for impacts of environmental factors on trait values. The 2014 growing cycle was highly productive, with minimal drought, unlike the prior two cycles. The 2015 nurseries experienced winter conditions that resulted in winterkill and late spring freezes that limited yield productivity. The magnitude of the negative relationship of GY and GPRO can be strongly influenced by environmental conditions (Haile et al., 2018; Iqbal et al., 2007; Michel et al., 2019b; Neuweiler et al., 2021; Rapp et al., 2018). We observed three environments in the 2014 and 2015 seasons with slopes that did not differ from zero (not significant, Table 3.4). These environments may have effected changed genotypic responses that impacted both GY and GPRO, but were not captured in prediction models from 2012 and 2013 season data for the 85 lines in the 2014 validation set.

To identify the best of the alternative selection strategies, means comparisons by the Tukey method were carried out among the subset of 2013 models that showed significant differences for GPRO in paired comparisons to the validation set ($\alpha < .05$, Table 3.14). Based

on results of the Dunnett's test, the selection strategies compared in Tukey means comparisons included GPY, univariate and bivariate selection indices GPY+GPD, and bivariate independent culling for GY & GPRO and GPY & GPD. Relative to the base population, the difference of mean GPRO was significantly different for the univariate selection index GPY+GPD (Δz-score = 0.79, $p < .001$), bivariate selection index GPY+GPD (Δz-score = 0.71, $p < .001$), and the bivariate independent culling GPY+GPD (Δz-score = 0.53, $p < .01$), but they did not differ among the three strategies ($p < .05$). These results suggest that combining GPY and GPD into protein-yield indices enables selection for GPRO without antagonistic selection on GY. The use of bivariate GEBVs simplifies the calculation of the selection indices by obtaining needed trait values within a single genomic selection model. During the very short timeline between harvest and planting in a winter wheat breeding program, such resource efficiencies are desirable.

Discussion

Simultaneous improvement of GY and GPRO is a central challenge for wheat breeders if they are to meet the 21st century productivity imperatives while maintaining milling and baking quality standards. This is difficult to accomplish given the strong negative relationship of GY and GPRO in cereal grain crops (Simmonds, 1995). Within a diverse population of breeding lines and varieties representative of the Colorado State University wheat breeding program, (1) index values were calculated based on several protein-yield indices, (2) correlations were evaluated between individual traits and selection indices, (3) the selection differential for GY and GPRO in response to index selection was determined, and (4) the selection strategies were extended to use in genomic predictions.

Phenotypic selection for simultaneous improvement of grain yield and protein concentration

Best linear unbiased predictors for GY and GPRO phenotypes were calculated across 676 lines grown in 20 ENV during the 2012 and 2013 growing seasons. A strong negative relationship was observed between the traits (Figure 3.4) and genotypes were observed with GPRO values that fell outside the range expected for their measured GY. Grain protein deviation represents the residuals of the regression of GPRO on GY (Oury & Godin, 2007). Genotypes were observed with high and low GPD that consistently expressed high or low values across ENV (Figure 3.5). The level of diversity in the population enabled development and testing of simultaneous selection strategies for GY and GPRO. Phenotypic selection using the protein-yield indices, GPD and GPY, identified complementary sets of breeding lines (Figure 3.6). These indices were positively correlated with each other, but GPD was negatively correlated with GY and positively correlated with GPRO, while GPY had the opposite relationships (Figure 3.10). In a study of a European hard winter wheat breeding population, (Michel et al., 2019b) the response to selection by several selection indices was evaluated for simultaneous selection for GY, GPRO, and GPY. No index produced a pool of selected lines that included both high GPRO lines with acceptable GY and high GY lines with acceptable GPRO. However, they could achieve this objective by splitting the selection decisions between two complementary indices, to obtain a combined selected fraction that included a broad range of the best performing lines for each trait. Another solution to a similar observation in durum wheat was presented by Rapp et al. (2018) to develop equally weighted selection indices that combined GPD and grain yield deviation (GYD) or combined GY and GPRO to apply for simultaneous selection in two European wheat breeding populations. Grain yield deviation is derived in the same way as GPD, but with the dependent and independent variables switched,

resulting in an index that selects high-yielding lines over the range of GPRO, similar to GPY. Both equally weighted indices showed simultaneous selection responses for GY and GPRO, but also retained lines in the selected fraction that did not meet minimum standards for GY or GPRO. Each index characteristically focused more or less strongly on selection for GY or GPRO. Similarly, index selection within a triticale breeding population, resulted in considerable differences within sets of selected individuals, depending on which of several protein-yield indices were applied.

The breeder's best choice for a selection index will relate to the grain composition desired for the targeted market (Neuweiler et al., 2021; Rapp et al., 2018). Colorado wheat growers profit most from varieties that deliver high GY under highly variable annual weather patterns, but obtain the best pricing for grain that meets minimum GPRO standards for milling and flour end-use markets. Strategies to obtain simultaneous selection of GY and GPRO were evaluated, including independent culling and several equally weighted selection indices (Table 3.8). The selected fraction for independent culling neither contained the highest values for GY nor GPRO, but did provide selection pressure for improved GPRO over selection based on GY alone. The apparent targeted grain composition for each selection index is approximated by its relative rank for the resultant selection differential (deltaS rank, Table 3.8) for GY or GPRO. For example, an index that nearly equally emphasized selection for GY and GPRO was the GY+GPD index that retained individuals with higher GY (deltaS = 0.77) equivalent to those retained by independent culling (deltaS = 0.78), but obtained a higher mean value for GPRO (deltaS = 0.65) compared to independent culling (deltaS = 0.31). Depending on the desired emphasis of the applied index, the distribution of GY and GPRO values in the selected fraction will differ and can be targeted for the breeding program objectives.

Genomic selection for simultaneous improvement of grain yield and protein concentration

A number of authors extended the work on phenotypic protein-yield selection indices to genomic selection (Haile et al., 2018; Michel et al., 2019a, 2019b; Rapp et al., 2018; Schulthess et al., 2016; Yao et al., 2018). Predictive ability for the target trait or index, correlations of index values with BLUPs and GEBVs for all traits and index values, and response to selection for GY and GPRO provided comparisons among index selection by direct (GEBVs for selection indices) and reversed (selection indices calculated from GEBVs) methods (Schulthess et al., 2016). Univariate or multi-trait GEBVs may be input to calculate selection index values by the reversed method. Multi-trait GS does not break undesirable negative trait correlations, but rather is useful for prediction of missing phenotypes and for improving prediction accuracy through genetic covariance (Jia & Jannink, 2012). Difficult or expensive to phenotype traits can be predicted in multi-trait GS in models that include partial phenotypes for highly correlated traits and may promote resource efficiency in a breeding program (Lado et al., 2018). Multi-trait GS utilizes correlated traits to enhance prediction accuracy for the target trait, but results have been mixed for superiority of multi-trait GS models outside of simulation studies (as reviewed in: Schulthess et al., 2016). These authors found marginal or no prediction accuracy improvement of multi-trait GS over single trait models for GY and GPRO in rye. When training population size is limited, and includes unbalanced data for a low heritability trait supported by balanced data for a highly heritable trait, multi-trait GS prediction accuracy improved. For unbalanced data, calculation of the selection index from multi-trait GEBVs provides a means to predict performance in advance of having complete phenotypic data (Schulthess et al., 2016). In this study, univariate and bivariate models delivered equivalent predictive ability for forward prediction of GY and GPRO (FP ratio, Table 3.11). Single trait GEBVs for GY, GPRO, GPY, and GPD were applied in

construction of selection indices and were used for independent culling. In the context of equivalent predictive ability, use of bivariate GEBVs to construct equally weighted selection indices reduces the computing steps required for index calculation by output of a single vector containing all traits and provides the opportunity for imputation of missing trait values (Lado et al., 2018).

Selection indices were calculated from BLUPs and then GEBVs for index values were obtained in univariate genomic selection models. Predicted index values included GPY, GPD and a set of indices calculated by summing paired traits (direct method, Table 3.10). The relative strength of the correlation with GY and GPRO indicated the emphasis of each selection strategy (Figure 3.14). The emphasis of the applied genomic selection index closely tracked that obtained with the same index applied to phenotypic selection (Table 3.8). Index values were also calculated from GEBVs obtained in univariate and bivariate models (reversed method, Table 3.11). The selection strategies were compared through analysis of the mean and distribution of the GY and GPRO among the lines retained in the selected fraction as compared to the 2013 and the 2014 validation sets (Tables 12 and 13). As observed in earlier studies (Neuweiler et al., 2021; Rapp et al., 2018; Schulthess et al., 2016), the distribution of GY and GPRO values in response to index selection extends beyond the desired selection thresholds for each trait, but each selection index enriches the selected fraction for high value lines, and delivers a characteristic emphasis on GY or GPRO. In this study, three genomic selection strategies were effective in emphasizing selection for GPRO while maintaining selection pressure for GY (Table 3.14): independent culling for GEBVs for GPY and GPD predicted in a bivariate model; index selection based on GEBVs predicted in a bivariate model, then summed to produce index values;

and index selection based on GEBVs predicted in a univariate model from index values calculated from phenotypic BLUPs.

Principal components analysis of the realized relationship matrix (G matrix) clustered allelic effects that captured the negative association between GY and GPRO in PC1 and captured effects that were positively correlated with both GY and GPRO in PC2 (Figures 10 and 11). These results are consistent with the underlying genetic architecture of the negative association between GY and GPRO. Genome wide association mapping and QTL analysis revealed a complex genetic architecture for protein-yield selection indices and the related single traits that included QTL with antagonistic pleiotropy and QTL with joint positive association for GY and GPRO (Thorwarth et al., 2019). With antagonistic pleiotropy, the authors recommend selection for the stronger allele at one locus, coupled with selection for a counteracting allele at a different locus. Given its reasonably high genomic prediction accuracy, they proposed the use of GPD for enrichment of alleles with positive effects on GY and GPRO during recurrent selection cycles. Cross prediction through simulated variances under an additive model demonstrated that when negative GY and GPRO correlation exists in the parental population, it is possible to have positive correlations in some progeny populations (Yao et al., 2018). The quantitative nature of the genetic architecture under-lying these traits and derived indices predicts limited utility of marker-assisted selection.

Conclusions

This work contributes to the evidence supporting use of protein-yield selection indices for simultaneous selection for GY and GPRO. Among the set of protein-yield selection indices evaluated in this study, the emphasis of the selection response for GY or GPRO is diverse and

provides an opportunity to tailor the selection response to breeding program objectives. A set of protein-yield selection indices were identified that generate sufficient predictive ability in genomic selection to be effective tools for simultaneous selection of GY and GPRO. The accuracy of genomic prediction of protein-yield index values obtained for the 2013 validation set encourages expanding the work to optimize simultaneous selection throughout a breeding cycle. Genomic selection may be applied during different stages of the breeding cycle (Bassi et al., 2016) and optimization may include selection on different indices at different stages.

Table 3.1. Summary of Colorado State University wheat breeding trial locations and harvest years.

Location Name	Management	Latitude (N)	Longitude(W)	Elevation (m)	No. of trials	No. of plots	No. of entries	Harvest Years
Akron	dryland	40.149	-103.136	1383	8	2197	676	2012- 2015
Arapahoe	dryland	38.840	-102.129	1213	1	150	75	2012
Burlington	dryland	39.187	-102.300	1295	4	150	75	2012, 2014, 2015
Dailey	dryland	40.598	-102.686	1251	2	433	275	2013, 2014
Fort Collins	irrigated	40.650	-104.999	1557	9	2605	663	2012-2015
Julesburg	dryland	40.801	-102.365	1169	4	433	306	2012-2014
Lamar	dryland	37.761	-102.482	1265	1	150	75	2012
Roggen	dryland	40.070	-104.302	1493	1	148	75	2014
Walsh	dryland	37.431	-102.315	1212	2	149	75	2014

Table 3.2. Environment code definitions, modeling cohorts and numbers of entries for Colorado State University wheat breeding program trials included in this study. An environment is represented by the combination of harvest year, location, and trial. Selection cohort (GRP_2) designates the first year that sets of genotypes entered evaluation trials.

Environment code[a]	Harvest year	Location name	Trial name[b]	Cohort (GRP_2)	Number of entries
12AK-NH	2012	Akron	TP-NH	2012	399
12AK-NL	2012	Akron	TP-NL	2012	399
12FC-NH	2012	Fort Collins	TP-NH	2012	399
12FC-NL	2012	Fort Collins	TP-NL	2012	399
12AK	2012	Akron	CSU Elite	2012	75
12AR	2012	Arapahoe	CSU Elite	2012	75
12BU	2012	Burlington	CSU Elite	2012	75
12JL	2012	Julesburg	CSU Elite	2012	75
12LM	2012	Lamar	CSU Elite	2012	75
13AK-DH	2013	Akron	AYN-DH	2013	228
13DL-DH	2013	Dailey	AYN-DH	2013	231
13FC-DHNH	2013	Fort Collins	AYN-DH NH	2013	234
13FC-DHNL	2013	Fort Collins	AYN-DH NL	2013	231
13FC-DH	2013	Fort Collins	AYN-DH	2013	231
13JL-DH	2013	Julesburg	AYN-DH	2013	234
14AK	2014	Akron	CSU Elite	2013	75
14DL	2014	Dailey	CSU Elite	2013	75
14FC	2014	Fort Collins	CSU Elite	2013	75
14RG	2014	Roggen	CSU Elite	2013	75
14WA	2014	Walsh	CSU Elite	2013	75
14AK1	2014	Akron	AYN	2014	17
14AK-DH	2014	Akron	AYN-DH	2014	33
14BU1	2014	Burlington	AYN1	2014	33
14BU3	2014	Burlington	AYN3	2014	15
14FC-dry	2014	Fort Collins	AYN-dry	2014	15
14FC2	2014	Fort Collins	AYN2	2014	7
14JU	2014	Julesburg	AYN	2014	17
14JU1	2014	Julesburg	AYN1	2014	32
14WA1	2014	Walsh	AYN	2014	17
15AK	2015	Akron	CSU Elite	2014	15
15BU	2015	Burlington	CSU Elite	2014	15
15FC	2015	Fort Collins	CSU Elite	2014	15

[a] Environment name encoding: doubled haploids (DH), limiting nitrogen fertilizer (NL), and sufficient nitrogen fertilizer (NH). Colorado location codes: Akron (AK), Arapahoe (AR), Burlington (BU), Dailey (DL), Fort Collins (FC), Julesburg (JL), Lamar (LM), Roggen (RG), Walsh (WA).

[b]Trial name abbreviations: Check varieties with breeding lines in 3rd year of evaluation (CSU Elite), check varieties with breeding lines in 2nd year of evaluation (Advanced Yield Nursery, AYN), check varieties and earlier cohorts of breeding lines (training panel, TP), doubled haploid populations (DH).

Table 3.3. Descriptions of spatial variance models tested for each environment (further details in: Table B.1, Butler et al., 2009).

Model	Variance model type	Spatial variance model function name	*ASReml-R* function	Measurement error term (nugget)	ASReml-R (Version 3) code
model 1	correlation	scaled identity structure	iid	no	asreml(trait~ID, random = ~ ROW + COLUMN, na.method.Y = 'include', na.method.X = 'include', data = x)
model 2	correlation	1st order autoregressive (row1)	ar1()	no	asreml(trait~ID, random = ~ ROW + COLUMN, rcov = ~ ar1(ROW1):COLUMN1, na.method.X = 'include', data = x)
model 3	correlation	1st order autoregressive (col1)	ar1()	no	asreml(trait~ID, random = ~ ROW + COLUMN, rcov = ~ ROW1:ar1(COLUMN1), na.method.X = 'include', data = x)
model 4	correlation	1st order autoregressive (col1)	ar1()	yes	asreml(trait~ID, random = ~ units + ROW + COLUMN, rcov = ~ ROW1:ar1(COLUMN1), na.method.X = 'include', data = x)
model 5	correlation	1st order autoregressive (row1 & col1)	ar1()	no	asreml(trait~ID, random = ~ ROW + COLUMN, rcov = ~ ar1(ROW1):ar1(COLUMN1), na.method.X = 'include', data = x)
model 6	correlation	1st order autoregressive (row1 & col1)	ar1()	yes	asreml(trait~ID, random = ~ units + ROW + COLUMN, rcov = ~ ar1(ROW1):ar1(COLUMN1), na.method.X = 'include', data= x)
model 7	2-dimensional irregularly spaced power	isotropic exponential	iexp()	yes	asreml(trait~ID,random= ~ units + ROW + COLUMN, rcov= ~iexp(COLUMN1,ROW1,init=0.9), data=x,control=asreml.control(maxiter=50))
model 8	2-dimensional irregularly spaced power	anisotropic exponential	aexp()	yes	asreml(trait~ID,random= ~ units + ROW + COLUMN,rcov=~aexp(COLUMN1,ROW1,init=c(summary(model 12)$varcomp[4,1],summary(model7)$varcomp[4,1])), data=x,control=asreml.control(maxiter=25))
model 9	2-dimensional irregularly spaced power	isotropic euclidean	ieuc()	yes	asreml(trait~ID,random= ~ ROW + COLUMN + units,rcov=~ieuc(COLUMN1,ROW1,init=0.9),data=x,control=asreml.control(maxiter=100))
model 10	2-dimensional irregularly spaced power	Matérn	mtrn()	yes	asreml(trait~ID,random= ~ ROW + COLUMN + units,rcov=~mtrn(COLUMN1,ROW1,phi=-1/log(summary(model9)$varcomp[4,1])), data=x,control=asreml.control(maxiter=25))

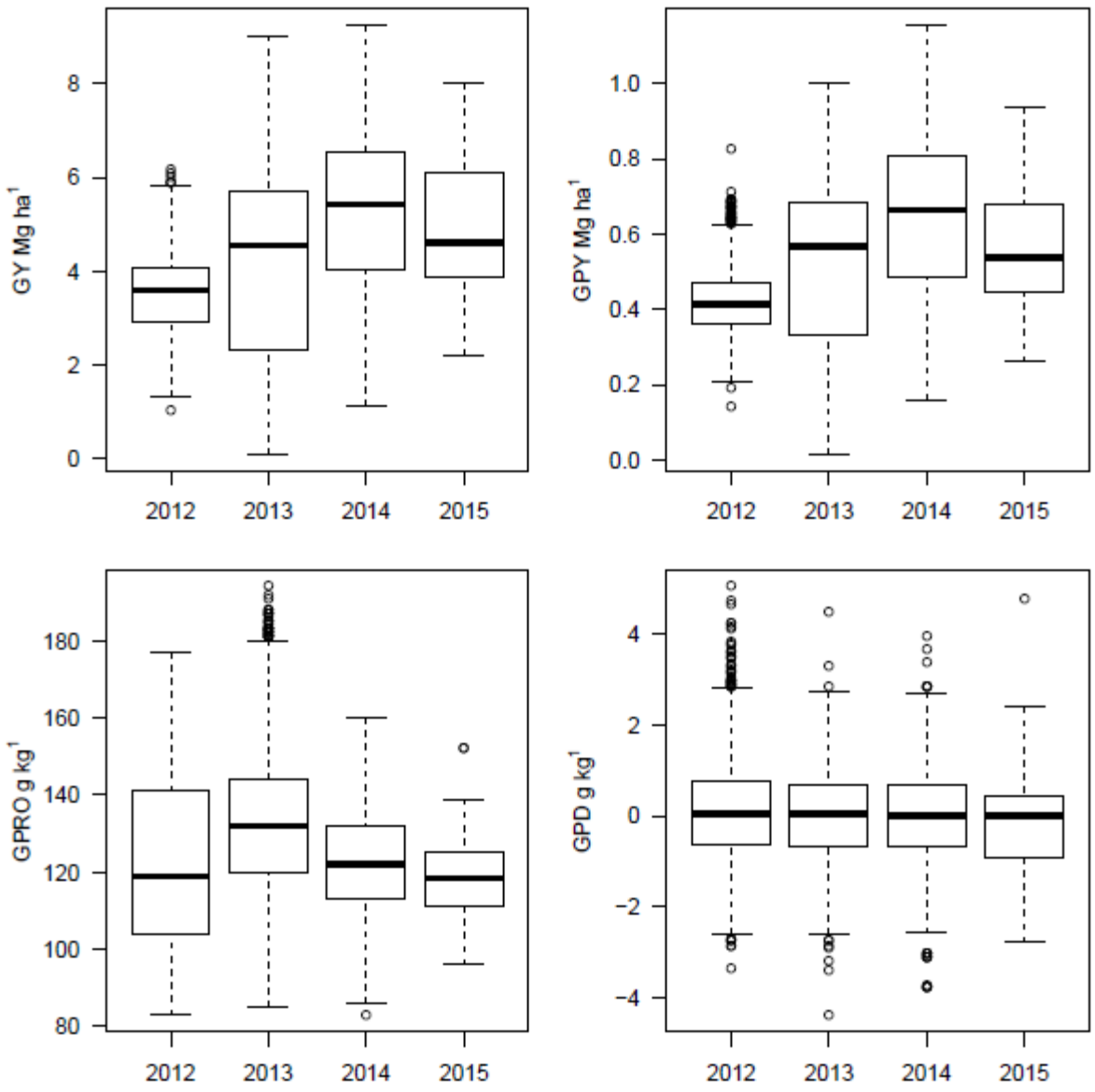


Figure 3.1. Box-and-whiskers plots for the 2012 to 2015 growing seasons of plot level grain yield (GY), grain protein concentration (GPRO), grain protein yield (GPY), and grain protein deviation (GPD). Data were recorded for 761 genotypes across 32 environments in Colorado, defined by harvest year, location, and type of trial.

Table 3.4. Phenotypic summary statistics for plot level wheat grain yield (GY, Mg ha^{-1}) and grain protein concentration (GPRO, g kg^{-1}) in 32 environments for 761 checks and genotypes that first entered performance nurseries in 2012 to 2014. Intercept and slope for the ranked regression of GPRO on GY.

Environment code[a]	GY N	GY mean	GY sd	GY min	GY max	GY quantile 25	GY quantile 75	GPRO mean	GPRO sd	GPRO min	GPRO max	GPRO quantile 25	GPRO quantile 75	Intercept	Slope
12AK	75	3.9	0.69	1.6	5.7	3.5	4.3	144	15.2	108	174	133	156	207.4	-16.2
12AK-NH	399	2.5	0.52	1.4	4.4	2.2	2.9	152	7.8	128	177	146	157	170.6	-7.6
12AK-NL	399	3.5	0.47	2.1	5.2	3.2	3.8	129	9.8	102	156	122	136	164.1	-10.1
12AR	75	2.5	0.47	1.3	3.9	2.2	2.9	144	8.7	122	171	138	149	178.9	-13.9
12BU	75	3.7	0.65	2.1	5.6	3.3	4.2	121	13.1	93	155	112	129	164.6	-12.0
12FC-NH	399	4.1	0.55	2.3	5.9	3.8	4.5	105	8.5	86	140	99	111	109.4	-1.3
12FC-NL	399	4.0	0.63	2.4	6.2	3.6	4.4	101	9.9	83	146	94	106	86.0	3.1
12JL	75	3.9	0.32	3.0	4.8	3.7	4.1	109	6.8	92	126	105	114	134.2	-6.4
12LM	75	2.8	0.55	1.0	4.3	2.5	3.2	118	12.4	94	147	109	126	156.0	-14.0
13AK-DH	227	1.3	0.54	0.1	3.0	0.9	1.6	172	8.5	143	194	167	179	187.2	-11.7
13DL-DH	230	4.8	0.55	3.3	6.4	4.4	5.2	128	7.0	98	146	123	133	135.5	-1.5
13FC-DHNH	231	6.2	0.84	3.7	8.3	5.6	6.8	115	11.2	92	147	107	123	165.7	-8.3
13FC-DHNL	230	6.0	0.92	3.8	8.4	5.3	6.6	113	12.2	85	146	104	122	134.4	-3.8
13FC-DH	232	4.5	1.15	2.4	9.0	3.8	5.0	134	8.6	102	160	129	139	145.1	-2.5
13JL-DH	234	2.3	0.31	1.5	3.3	2.1	2.5	142	5.1	130	163	139	145	147.7	-2.6
14AK1	25	6.9	0.70	5.5	8.0	6.5	7.5	125	9.4	109	145	119	130	186.6	-8.7
14AK-DH	35	7.4	0.75	5.9	8.7	7.0	7.9	122	7.5	109	139	115	127	163.2	-5.5
14AK	75	6.2	0.71	4.3	8.3	5.9	6.7	128	8.4	107	150	122	134	179.4	-8.3
14BU1	35	3.4	1.02	2.0	5.6	2.5	4.0	114	12.3	86	139	106	125	129.7	-5.4
14BU3	23	3.3	0.63	1.7	4.6	3.1	3.8	113	11.0	96	134	104	119	147.0	-10.5
14DL	75	4.8	0.91	2.7	7.2	4.1	5.3	108	9.6	83	129	100	115	107.0	0.0^{ns}
14FC-dry	23	6.2	0.43	5.5	7.0	5.9	6.6	119	8.0	105	136	114	126	134.6	-2.4^{ns}
14FC2	11	7.9	0.50	7.2	9.2	7.7	8.1	118	3.2	113	122	115	121	142.6	-3.0
14FC	75	7.0	0.64	5.3	8.5	6.6	7.4	117	8.3	95	150	111	121	152.2	-5.2
14JU	26	4.8	0.54	3.4	6.1	4.5	5.1	141	6.2	132	160	137	145	166.9	-5.5
14JU1	34	5.5	0.61	3.5	6.5	5.3	5.8	136	9.5	113	156	130	143	189.3	-9.4
14RG	75	5.4	0.65	3.9	7.7	4.9	5.8	109	7.4	90	130	105	114	132.6	-4.3
14WA1	26	3.4	0.75	1.3	4.4	3.0	4.0	136	6.4	124	150	130	141	149.7	-4.0
14WA	75	3.6	0.72	1.1	4.9	3.4	4.1	128	6.4	114	149	123	132	143.0	-4.4
15AK	24	4.0	1.11	2.2	6.1	3.3	5.0	120	9.7	97	139	113	126	137.9	-3.6

Environment code[a]	GY N	GY mean	GY sd	GY min	GY max	GY quantile 25	GY quantile 75	GPRO mean	GPRO sd	GPRO min	GPRO max	GPRO quantile 25	GPRO quantile 75	Intercept	Slope
15BU	24	4.1	0.66	2.8	5.5	3.8	4.5	120	11.9	101	152	110	127	158.2	-9.1
15FC	24	6.5	0.91	4.4	8.0	6.0	7.0	116	10.5	96	152	111	119	134.3	-2.9[ns]
Overall	761	4.1	1.55	0.1	9.2	3.1	5.0	125	20.5	83	194	109	139	207.4	-16.2

[ns]Not significant at the .05 probability level.

[a] Trial name encoding: doubled haploids (DH), limiting nitrogen fertilizer (NL), and sufficient nitrogen fertilizer (NH). Colorado location codes: Akron (AK), Arapahoe (AR), Burlington (BU), Dailey (DL), Fort Collins (FC), Julesburg (JL), Lamar (LM), Roggen (RG), Walsh (WA).

Table3. 5. Phenotypic summary statistics for plot level wheat grain protein yield (GPY, Mg ha^{-1}) and grain protein deviation (GPD, g kg^{-1}, as Studentized residuals of a ranked regression) for 32 environments for 761 checks and genotypes which first entered advanced testing in 2012 to 2014.

Environment code[a]	GPY N	GPY mean	GPY sd	GPY min	GPY max	GPY quantile 25	GPY quantile 75	GPD mean	GPD sd	GPD min	GPD max	GPD quantile 25	GPD quantile 75
12AK	75	0.554	0.072	0.283	0.710	0.520	0.592	-0.037	0.963	-2.850	2.065	-0.588	0.683
12AK-NH	399	0.382	0.070	0.217	0.602	0.335	0.424	0.071	0.959	-2.851	3.212	-0.631	0.736
12AK-NL	399	0.449	0.054	0.296	0.635	0.411	0.485	-0.010	0.973	-2.725	2.572	-0.714	0.672
12AR	75	0.362	0.054	0.190	0.515	0.327	0.396	0.096	1.124	-3.338	3.610	-0.668	0.708
12BU	75	0.449	0.061	0.270	0.648	0.410	0.489	0.167	1.037	-2.173	2.738	-0.585	0.861
12FC-NH	399	0.431	0.066	0.244	0.687	0.384	0.476	0.169	0.976	-2.122	4.225	-0.513	0.808
12FC-NL	399	0.408	0.084	0.235	0.825	0.346	0.453	0.242	1.135	-2.189	5.038	-0.576	0.892
12JL	75	0.428	0.038	0.315	0.513	0.405	0.455	0.004	1.016	-2.738	2.730	-0.611	0.715
12LM	75	0.325	0.053	0.142	0.429	0.293	0.366	0.094	0.915	-1.914	2.225	-0.580	0.776
13AK-DH	227	0.218	0.084	0.016	0.496	0.161	0.281	-0.009	1.059	-2.913	2.733	-0.674	0.663
13DL-DH	230	0.616	0.076	0.412	0.830	0.559	0.675	-0.040	0.977	-4.371	2.463	-0.711	0.634
13FC-DHNH	231	0.706	0.076	0.478	0.923	0.652	0.755	0.071	0.914	-2.061	2.510	-0.622	0.730
13FC-DHNL	230	0.672	0.106	0.441	0.974	0.582	0.756	0.104	0.896	-2.028	2.288	-0.594	0.834
13FC-DH	232	0.604	0.139	0.367	1.001	0.505	0.682	0.059	0.985	-2.544	2.835	-0.710	0.673
13JL-DH	234	0.325	0.043	0.203	0.452	0.295	0.354	0.071	1.049	-2.585	4.466	-0.627	0.716
14AK1	25	0.864	0.069	0.714	1.031	0.813	0.905	-0.119	0.934	-1.643	1.567	-0.685	0.384
14AK-DH	35	0.903	0.083	0.788	1.153	0.837	0.951	-0.070	1.033	-2.284	2.852	-0.728	0.633
14AK	75	0.792	0.068	0.598	0.973	0.752	0.839	0.048	0.925	-2.565	2.240	-0.601	0.758
14BU1	35	0.378	0.104	0.235	0.652	0.290	0.446	0.202	0.965	-2.066	2.310	-0.565	0.902
14BU3	23	0.374	0.058	0.230	0.502	0.345	0.413	0.182	0.989	-1.548	2.350	-0.612	0.828
14DL	75	0.513	0.106	0.264	0.835	0.442	0.589	0.063	0.931	-2.316	2.127	-0.675	0.773
14FC-dry	23	0.740	0.067	0.632	0.884	0.692	0.784	-0.068	0.859	-1.596	1.667	-0.725	0.515
14FC2	11	0.933	0.055	0.838	1.062	0.903	0.954	-0.419	1.250	-3.094	1.083	-0.936	0.593
14FC	75	0.810	0.070	0.606	0.972	0.765	0.861	0.070	1.138	-3.702	3.967	-0.646	0.723
14JU	26	0.678	0.070	0.544	0.890	0.641	0.711	0.115	0.711	-1.040	1.567	-0.490	0.710
14JU1	34	0.744	0.067	0.549	0.862	0.709	0.793	-0.261	1.177	-3.743	2.203	-0.983	0.455
14RG	75	0.588	0.067	0.423	0.854	0.545	0.631	-0.038	0.966	-2.351	2.504	-0.811	0.649
14WA1	26	0.464	0.098	0.195	0.639	0.398	0.528	0.021	0.831	-1.506	2.069	-0.673	0.763
14WA	75	0.459	0.087	0.161	0.602	0.423	0.520	0.108	0.992	-2.515	3.661	-0.540	0.757
15AK	24	0.478	0.125	0.261	0.742	0.403	0.576	-0.290	0.844	-2.150	0.971	-0.973	0.305

Environment code[a]	GPY N	GPY mean	GPY sd	GPY min	GPY max	GPY quantile 25	GPY quantile 75	GPD mean	GPD sd	GPD min	GPD max	GPD quantile 25	GPD quantile 75
15BU	24	0.486	0.066	0.357	0.651	0.443	0.510	-0.117	1.117	-1.757	2.392	-0.971	0.476
15FC	24	0.749	0.100	0.564	0.937	0.669	0.826	0.057	1.396	-2.780	4.788	-0.563	0.684
Overall	761	0.497	0.176	0.016	1.15	0.374	0.598	0.068	1.000	-4.370	5.040	-0.640	0.725

[a] Trial name encoding: doubled haploids (DH), limiting nitrogen fertilizer (NL), and sufficient nitrogen fertilizer (NH). Colorado location codes: Akron (AK), Arapahoe (AR), Burlington (BU), Dailey (DL), Fort Collins (FC), Julesburg (JL), Lamar (LM), Roggen (RG), Walsh (WA).

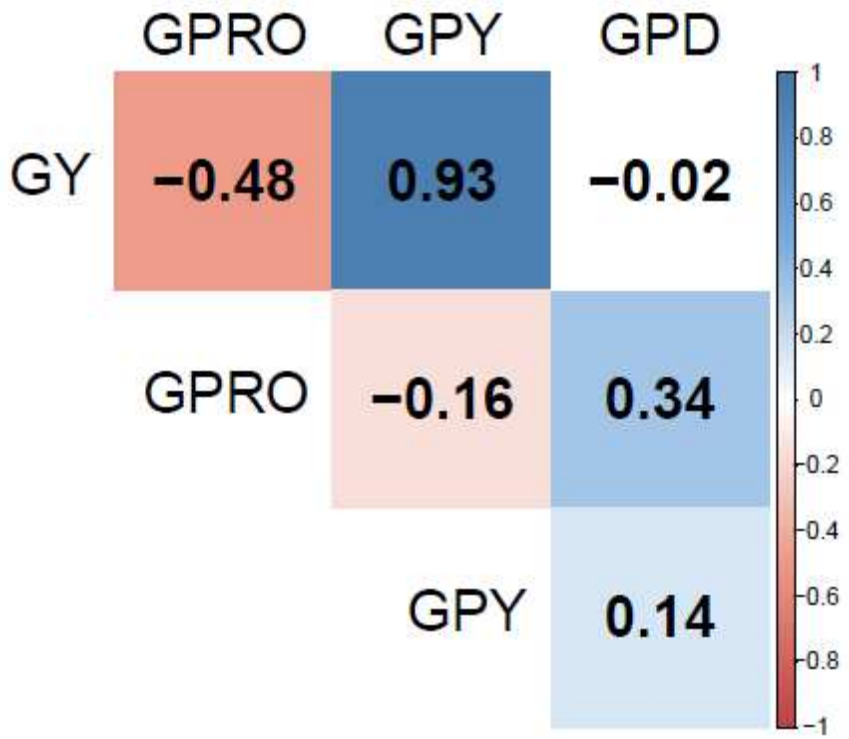


Figure 3.2. Matrix of the Pearson's correlation coefficients between wheat grain yield (GY), grain protein concentration (GPRO), grain protein yield (GPY), and grain protein deviation (GPD) phenotypes collected from 761 breeding lines and varieties. Plot level phenotypes were recorded across 32 environments encompassing four harvest years (2012-2015). All correlations are significant ($p < .05$) except for that which relates GY to GPD, which is non-significant. Intensity of color shading scales with the value of the correlation coefficient.

Table 3.6. Number of reps and best spatial variance models, determined by ΔAIC method (Burnham & Anderson, 2003), for each of 20 trials included in the combined analysis over three growing seasons (2012-2014) in the Colorado State University winter wheat breeding program for 676 breeding lines and varieties. Summary of the model output, including: within environment adjusted means and standard error of the mean (sem) for grain yield, grain protein concentration, grain protein yield, and grain protein deviation.

		Grain Yield Mg ha^{-1}			Grain Protein Concentration g kg^{-1}			Grain Protein Yield Mg ha^{-1}			Grain Protein Deviation g kg^{-1}		
Trial code[a]	Mean reps	Model[b]	Mean	sem	Model	Mean	sem	Model	Mean	sem	Model	Mean	sem
12AK	1.9	6	3.44	1.16	9	149.6	17.6	5	0.55	0.03	2	-0.044	0.076
12AK-NH	2.2	8	2.52	0.45	9	152.1	2.9	9	0.38	0.08	7	0.082	0.042
12AK-NL	2.2	8	3.45	0.19	7	129.3	4.3	6	0.44	0.02	7	-0.036	0.031
12AR	1.9	5	2.54	0.21	1	144.2	3.8	5	0.36	0.03	1	0.096	0.114
12BU	1.9	6	3.67	0.25	7	122.4	8.5	2	0.45	0.02	2	0.144	0.060
12FC-NH	2.2	8	4.10	0.24	8	106.5	3.7	7	0.43	0.03	8	0.310	0.031
12FC-NL	2.2	8	4.01	0.30	6	100.0	4.6	7	0.40	0.04	7	0.191	0.037
12JL	1.9	5	3.91	0.13	9	110.8	4.5	9	0.44	0.18	7	0.227	0.076
12LM	1.9	6	2.79	0.23	2	117.4	5.6	7	0.31	0.03	2	0.085	0.068
13AK-DH	1.3	5	1.32	0.22	9	171.8	3.2	9	0.23	0.03	5	-0.031	0.068
13DL-DH	1.3	9	4.75	0.27	7	128.8	16.2	9	0.61	0.04	1	0.056	0.056
13FC-DHNH	1.3	8	6.18	0.29	3	114.8	3.5	3	0.70	0.03	3	0.037	0.053
13FC-DHNL	1.3	7	6.38	0.82	3	112.9	4.1	6	0.70	0.08	3	0.086	0.046
13FC-DH	1.3	6	4.93	1.40	2	134.2	2.6	2	0.61	0.04	5	0.062	0.060
13JL-DH	1.3	5	2.28	0.09	2	142.2	1.3	1	0.33	0.01	7	0.104	0.066
14AK	1.9	6	6.18	0.46	2	127.9	2.8	7	0.79	0.04	3	0.044	0.077
14DL	1.9	6	5.32	0.55	5	107.6	3.5	7	0.57	0.06	5	0.062	0.081
14FC	1.9	5	6.94	0.28	2	116.7	2.2	2	0.81	0.03	1	0.070	0.120
14RG	1.9	5	5.41	0.33	9	108.3	19.4	5	0.59	0.04	9	-0.257	0.072
14WA	1.9	6	3.61	0.27	1	127.8	2.5	2	0.46	0.03	1	0.110	0.100

[a] Trial name encoding: doubled haploids (DH), limiting nitrogen fertilizer (NL), and sufficient nitrogen fertilizer (NH). Colorado location codes: Akron (AK), Arapahoe (AR), Burlington (BU), Dailey (DL), Fort Collins (FC), Julesburg (JL), Lamar (LM), Roggen (RG), Walsh (WA).

[b]Spatial variance model definitions appear in Table 3.

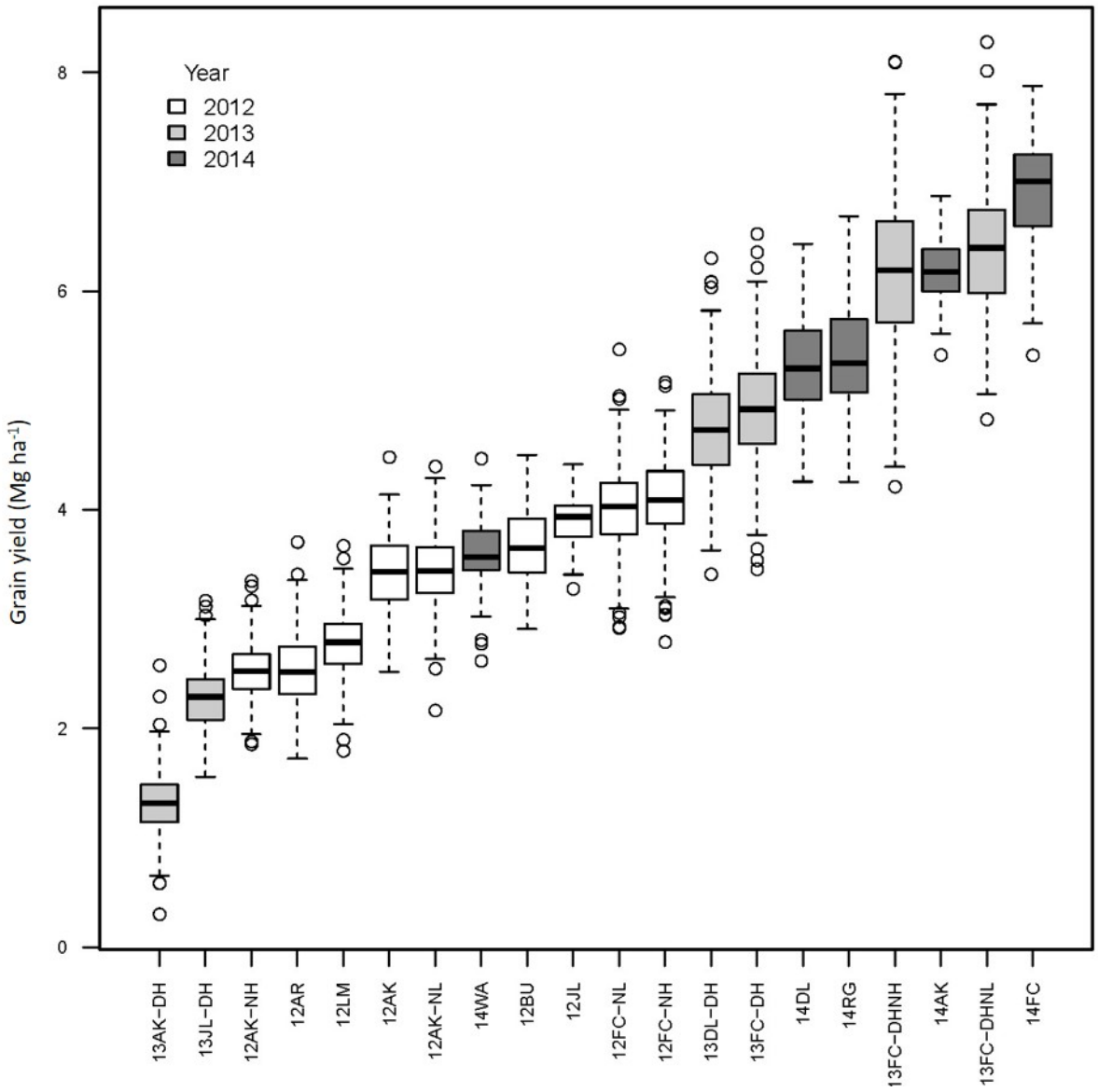


Figure 3.3. Box and whiskers plots of adjusted entry means (BLUEs) for grain yield (Mg ha^{-1}) at 20 environments, ordered by environmental averages for 676 hard winter wheat breeding lines and varieties grown during the 2012-2014 seasons in Colorado. Trial name encoding: doubled haploids (DH), limiting nitrogen fertilizer (NL), and sufficient nitrogen fertilizer (NH). Colorado location codes: Akron (AK), Arapahoe (AR), Burlington (BU), Dailey (DL), Fort Collins (FC), Julesburg (JL), Lamar (LM), Roggen (RG), Walsh (WA).

Table 3.7. Variance components, heritability on an entry mean basis, and weighting factors obtained in mixed model analysis for wheat grain yield, grain protein concentration, grain protein yield and grain protein deviation for 676 hard winter wheat breeding lines and varieties grown in 20 Colorado environments during the 2012 to 2014 growing seasons. Model weighting factors (w_p) for each environment to apply in a combined analysis across environments were calculated as the quotient of the average error mean square over all ENV and the error mean square within an ENV(p), multiplied by the average number of reps.

	Grain Yield Mg ha^{-1}				Grain Protein Concentration g kg^{-1}				Grain Protein Yield 10^2 x Mg ha^{-1}				Grain Protein Deviation 10^2 x g kg^{-1}			
Environment code[a]	σ_g^2	σ_e^2	h^2	w_p	σ_g^2	σ_e^2	h^2	w_p	σ_g^2	σ_e^2	h^2	w_p	σ_g^2	σ_e^2	h^2	w_p
12AK	0.679	1.066	0.56	0.47	26.2	47.8	0.52	0.25	1.18	1.72	0.58	2.22	0.25	0.47	0.35	3.19
12AK-NH	0.377	0.699	0.52	2.64	31.5	17.6	0.78	7.25	0.66	1.34	0.50	1.01	0.53	0.16	0.77	8.95
12AK-NL	0.828	0.999	0.62	8.66	21.7	43.8	0.50	2.68	1.01	1.39	0.59	8.08	0.19	0.55	0.26	3.19
12AR	0.578	1.138	0.50	5.61	45.9	26.1	0.78	4.78	0.40	1.70	0.32	3.97	0.70	0.53	0.57	2.84
12BU	0.684	1.648	0.45	2.92	4.1	74.6	0.10	0.80	0.60	1.59	0.43	3.12	0.08	0.62	0.12	1.65
12FC-NH	0.865	0.989	0.64	20.6	21.7	23.6	0.65	5.14	0.65	1.54	0.46	8.76	0.25	0.39	0.38	4.21
12FC-NL	0.649	1.531	0.46	5.63	16.8	41.2	0.45	2.71	0.10	3.33	0.06	2.50	0.36	0.50	0.41	3.14
12JL	0.400	0.292	0.73	22.8	11.4	20.6	0.53	3.98	0.40	0.44	0.64	0.27	0.18	0.67	0.21	2.69
12LM	0.615	0.866	0.59	7.68	28.0	52.1	0.52	1.76	0.51	1.23	0.45	4.20	0.17	0.46	0.27	3.33
13AK-DH	0.720	1.041	0.12	2.96	17.1	26.6	0.56	3.79	0.24	2.51	0.16	1.21	0.67	0.34	0.66	3.19
13DL-DH	0.353	1.471	0.32	3.87	18.7	18.2	0.67	0.33	0.55	2.36	0.32	1.64	0.31	0.41	0.43	0.16
13FC-DHNH	1.806	1.986	0.65	3.33	50.1	27.8	0.78	2.86	0.43	1.00	0.46	3.43	0.19	0.25	0.43	3.96
13FC-DHNL	1.086	2.458	0.47	0.60	44.8	41.8	0.68	1.52	0.51	2.50	0.29	0.55	0.21	0.21	0.50	3.52
13FC-DH	0.688	1.305	0.51	0.19	37.7	16.9	0.82	4.69	0.62	2.05	0.38	1.91	0.46	0.36	0.56	3.66
13JL-DH	0.523	0.240	0.81	23.6	18.7	6.7	0.85	11.12	1.02	0.41	0.83	14.5	0.75	0.22	0.77	10.8
14AK	0.715	0.892	0.62	2.17	19.2	13.7	0.74	7.01	0.72	1.21	0.54	3.93	0.31	0.24	0.56	5.79
14DL	1.619	1.424	0.69	0.94	43.7	24.4	0.78	2.75	0.21	2.62	0.14	0.78	0.45	0.45	0.50	3.53
14FC	1.502	1.499	0.67	4.47	56.7	9.6	0.92	12.07	1.45	2.00	0.59	3.86	0.93	0.25	0.79	6.36
14RG	1.650	2.306	0.59	3.45	25.6	13.0	0.80	0.36	0.16	3.19	0.09	2.19	0.17	0.70	0.19	0.67
14WA	0.552	1.028	0.52	5.80	21.0	9.2	0.82	13.82	0.83	1.36	0.55	4.54	0.59	0.23	0.72	6.47

[a] Trial name encoding: doubled haploids (DH), limiting nitrogen fertilizer (NL), and sufficient nitrogen fertilizer (NH). Colorado location codes: Akron (AK), Arapahoe (AR), Burlington (BU), Dailey (DL), Fort Collins (FC), Julesburg (JL), Lamar (LM), Roggen (RG), Walsh (WA).

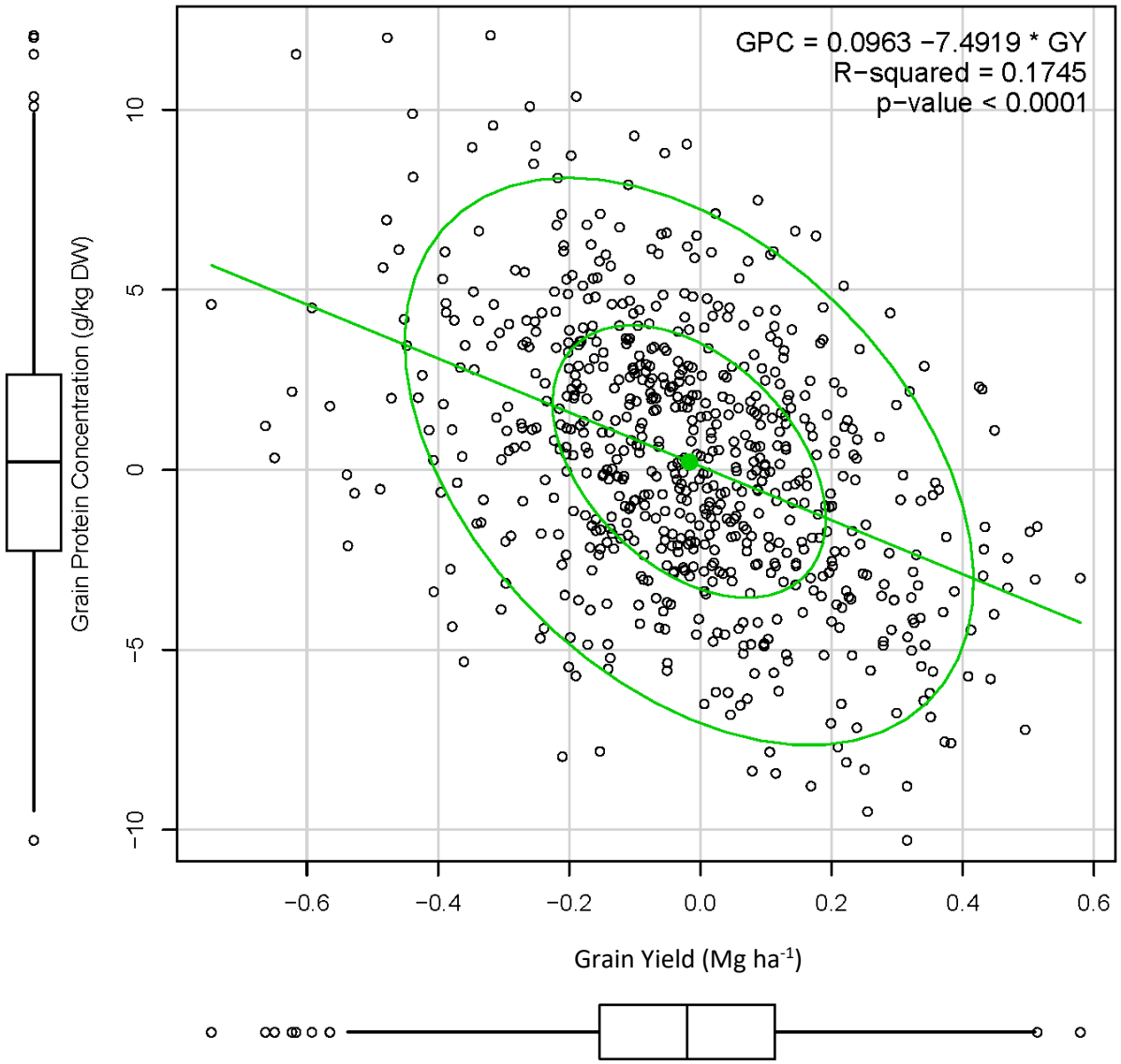


Figure 3.4. Scatterplot and line of best fit relating best linear unbiased predictors for grain protein concentration (GPRO, g kg^{-1}) to grain yield (GY, Mg ha^{-1}) for 676 entries grown in 20 environments in Colorado during 2012, 2013, and 2014. Positions of trait values on the standard *Normal* distributions are marked by concentric ovals and by a green dot for the mean value $N[0,0]$. Quantile boxplots, regression coefficients and significance of the fit display the significant inverse relationship between these traits.

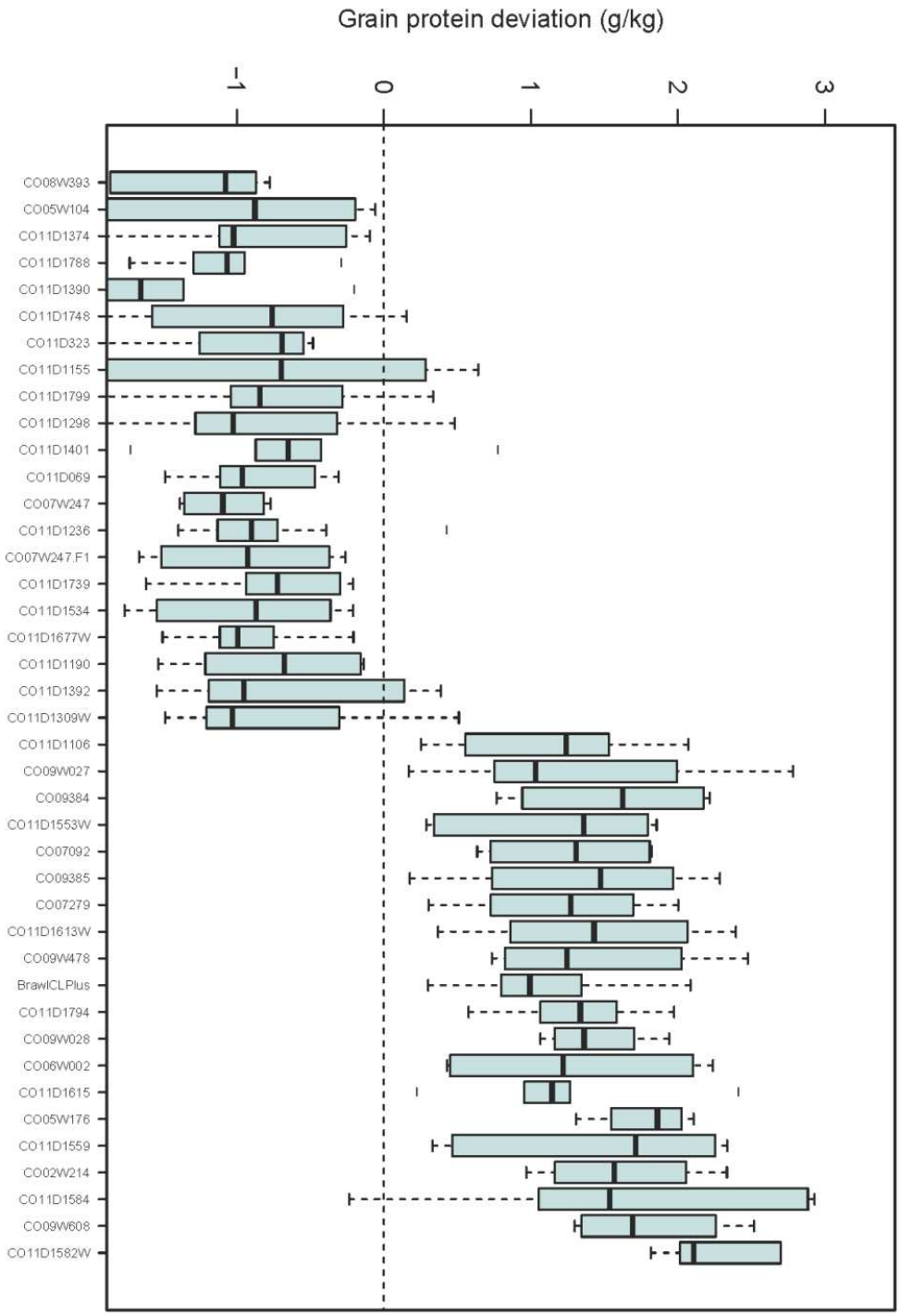
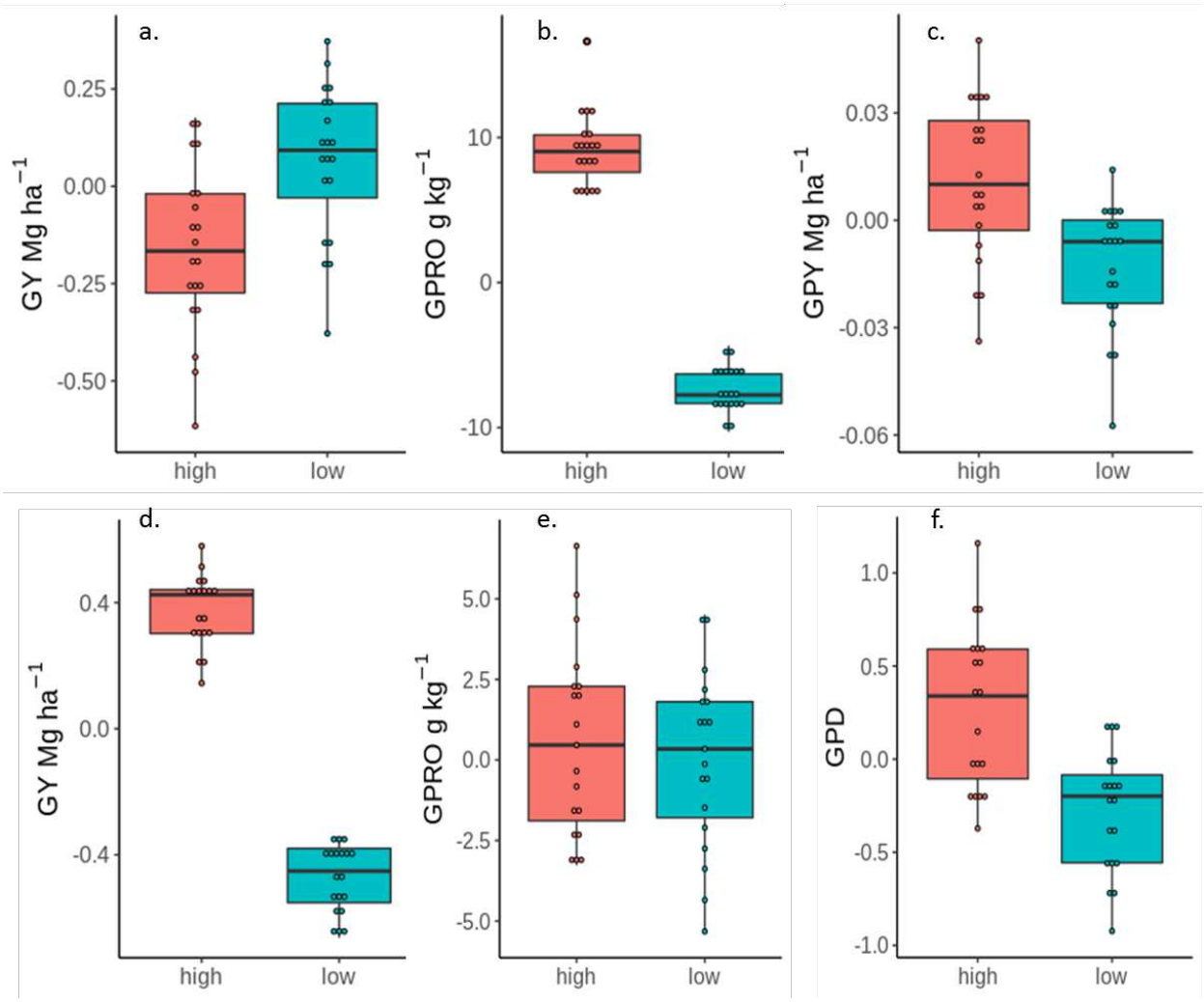


Figure 3.5. Box and whiskers plots of grain protein deviation BLUEs calculated within 20 environments during the 2012 to 2014 Colorado growing seasons. Values for the twenty highest and lowest ranked genotypes among 676 hard winter wheat breeding lines and varieties from the Colorado State University breeding program are plotted in rank order of the mean.



Selection target	Selection category	Grain yield mean (Mg ha^{-1})	Grain yield std dev	Student's 2-sample t(df) and significance	Grain protein conc. mean (g kg^{-1})	Grain protein conc. std dev	Student's 2-sample t(df) and significance
GPD	High	-0.161	0.215	-3.50 (37.66)***	9.23	2.58	25.01 (30.23)***
a,b,c	Low	0.065	0.195		-7.38	1.48	
GPY	High	0.374	0.113	23.68 (35.95)***	0.58	2.95	0.60 (35.80)^{ns}
d,e,f	Low	-0.475	0.108		0.02	2.74	

p-value significance level: 0.05*, 0.01**, 0.001***

Figure 3.6. Box and whiskers plots of BLUPs calculated across 20 environments during the 2012 to 2014 Colorado growing seasons. Best linear unbiased predictors (BLUP) for GY and GPRO and GPY or GPD for the top (pink fill) and bottom (blue fill) 20 genotypes ranked by GPD (a, b, c) or GPY (d, e, f). Open circles display BLUPs for the selected genotypes. Summary statistics and means comparison tests are tabulated below the plots. Small p-values indicate group means that are not the same.

Table 3.8. Response to independent culling compared to index selection and to indirect selection for grain yield (GY) and grain protein concentration (GPRO) among 661 genotypes from the 2012 and 2013 cohorts. A selection threshold for single traits or protein-yield selection indices identified the top 20 % of genotypes based on rankings of standardized and centered z-scores calculated from BLUPs. Selection response is reported in standard deviation units (z-score). Selection differentials (deltaS) are the difference between the mean z-scores of the selected fraction and that of the population for GY or GPRO. Summary statistics for z-scores for GY and GPRO of the selected fraction are reported.

Trait or Protein-yield Index [a]	r (GY)[b]	r (GPRO)[c]	deltaS (GY)	deltaS (GPRO)	deltaS rank (GY)	deltaS rank (GPRO)	sd (GY)	sd (GPRO)	min (GY)	min (GPRO)	max (GY)[d]	max (GPRO)[e]
IC GY, then GPRO			0.78	0.31			0.44	0.52	0.22	-0.43	2.25	1.93
GY		-0.42	1.39	-0.68	1	10	0.46	0.87	0.79	-2.80	2.87	1.67
GY + GPY	0.95	-0.18	1.33	-0.31	2	9	0.54	1.04	0.01	-2.80	2.87	4.36
3*GY + GPRO	0.94	-0.09	1.30	-0.17	3	8	0.57	1.06	-0.01	-2.40	2.87	4.36
GPY	0.81	0.08	1.09	0.08	4	7	0.73	1.09	-1.17	-2.08	2.87	4.36
GY + GPD	0.62	0.43	0.77	0.65	5	6	0.85	0.96	-1.45	-1.16	2.87	4.36
GY + GPRO	0.54	0.54	0.63	0.84	6	5	0.94	0.94	-2.21	-1.13	2.87	4.36
GPY + GPD	0.36	0.65	0.40	0.95	7	4	1.01	0.93	-2.88	-0.96	2.87	4.36
GPY + GPRO	0.27	0.73	0.30	1.06	8	3	1.03	0.88	-2.88	-0.93	2.87	4.36
GPD	-0.23	0.95	-0.35	1.34	9	2	0.98	0.63	-3.50	0.32	2.15	4.36
GPRO	-0.42		-0.62	1.40	10	1	0.87	0.58	-3.50	0.84	1.48	4.36

[a] IC, independent culling; GY, grain yield; GPRO, grain protein concentration; GPY, grain protein yield; GPD, grain protein deviation; '+', summed standardized and centered trait values to calculate selection index values.

[b] Pearson's correlations (r, $p < .05$) between trait or index values with grain yield (GY) or [c] grain protein concentration (GPRO).

[d] Among the selection candidates, the maximum z-score value for GY is 2.87 and for [e]GPRO is 4.36.

Table 3.9. Count of genotypes and pedigrees per harvest year. Within group or overall mean inbreeding coefficient (F_i) as estimated by the diagonal elements (g_{ii}) of the realized relationship matrix (G). The mean value of the diagonal elements approximate $1 + F_i$ where the expectation for the inbreeding coefficient equals 1 for fully inbred lines. The off-diagonal elements (g_{ij}) estimate twice the coefficient of coancestry, or the probability of marker alleles being identical by descent, with an expected value of zero for unrelated individuals, 0.25 for half-sibs, and 0.5 for full-sibs.

Harvest year	No. genotypes	No. pedigrees	F_i	g_{ij}
All	775	280	0.55	0
2012	427	218	0.29	0.27
2013	244	27	0.85	0.56
2014	104	35	0.93	0.62

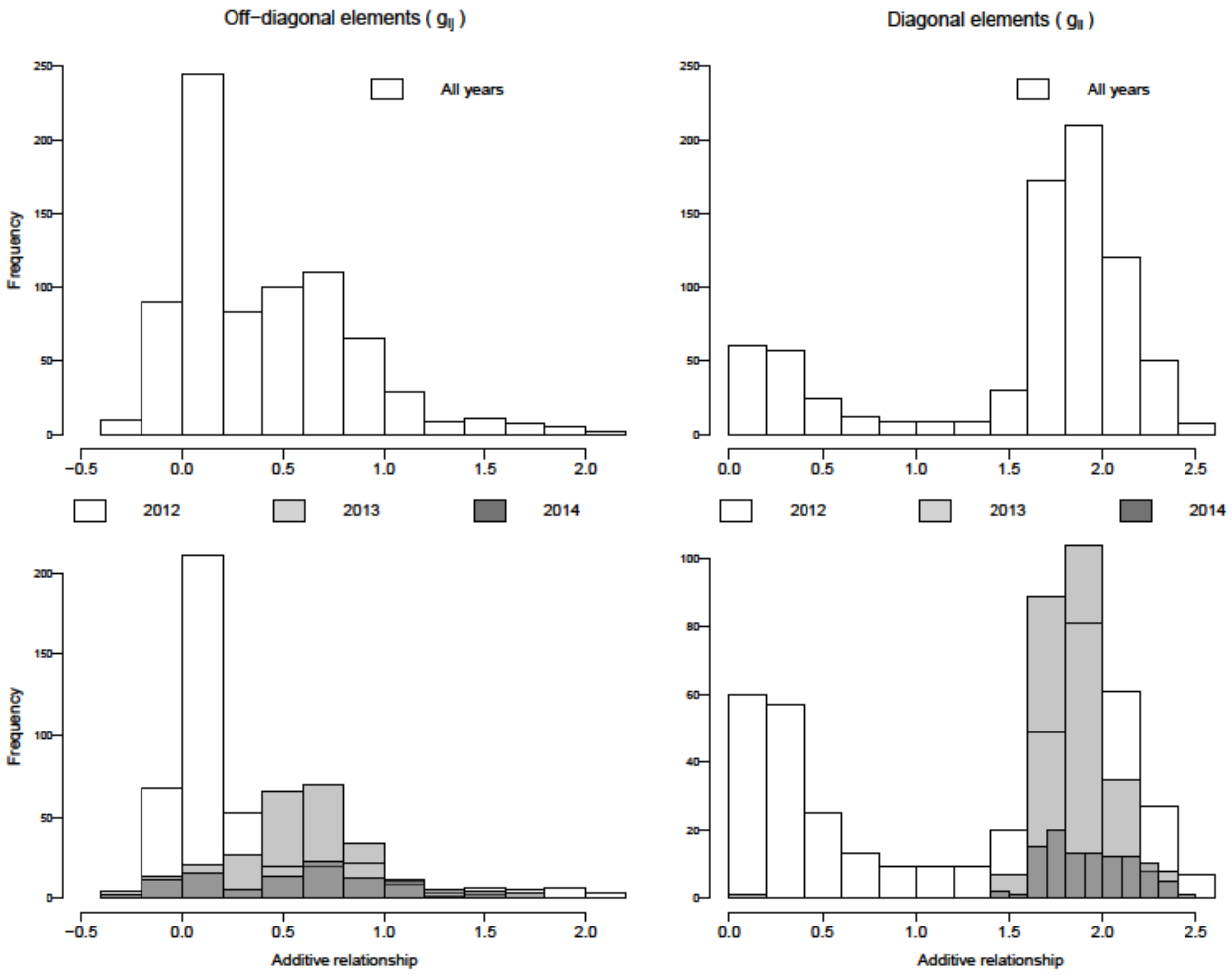


Figure 3.7. Distribution of values for the elements of the realized relationship matrix (G matrix) identified in genotyping-by-sequencing (GBS) genotypes for breeding lines and varieties ($N = 775$). The off-diagonal elements (g_{ij}) estimate twice the coancestry, or the probability of marker alleles being identical by descent, with an expected value of zero for unrelated individuals. The diagonal elements estimate the inbreeding coefficient, or the probability that alleles within an individual are identical by descent, with an expected value of 2 for fully inbred individuals. Upper plots display distributions for all genotypes and the lower plots display overlapping distributions, shaded to distinguish selection cohorts.

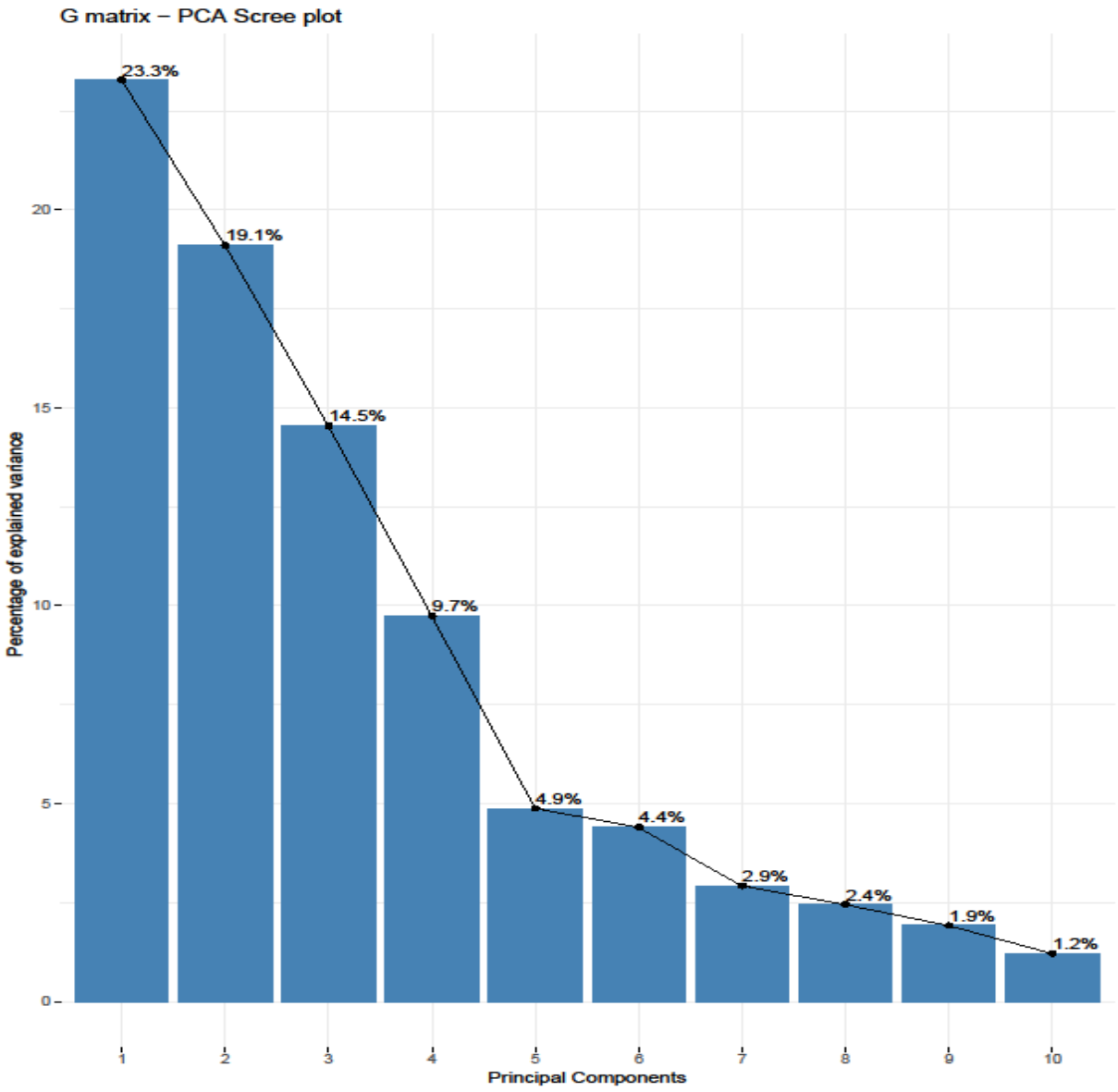


Figure 3.8. Scree plot of the principal components with percentages of variance explained for the realized relationship matrix (G) derived from SNP genotypes for 775 winter wheat varieties and advanced breeding lines that were in the Colorado State University evaluation trials during the 2012-2014 growing seasons.

Figure 3.9. Principal components biplot of PC1 and PC2 estimated from the realized relationship matrix (G) derived from GBS genotypes for 775 winter wheat varieties and advanced breeding lines that were in the Colorado State University evaluation trials during the 2012-2014 growing seasons. Initial year of testing an individual is represented by colored and filled shapes.

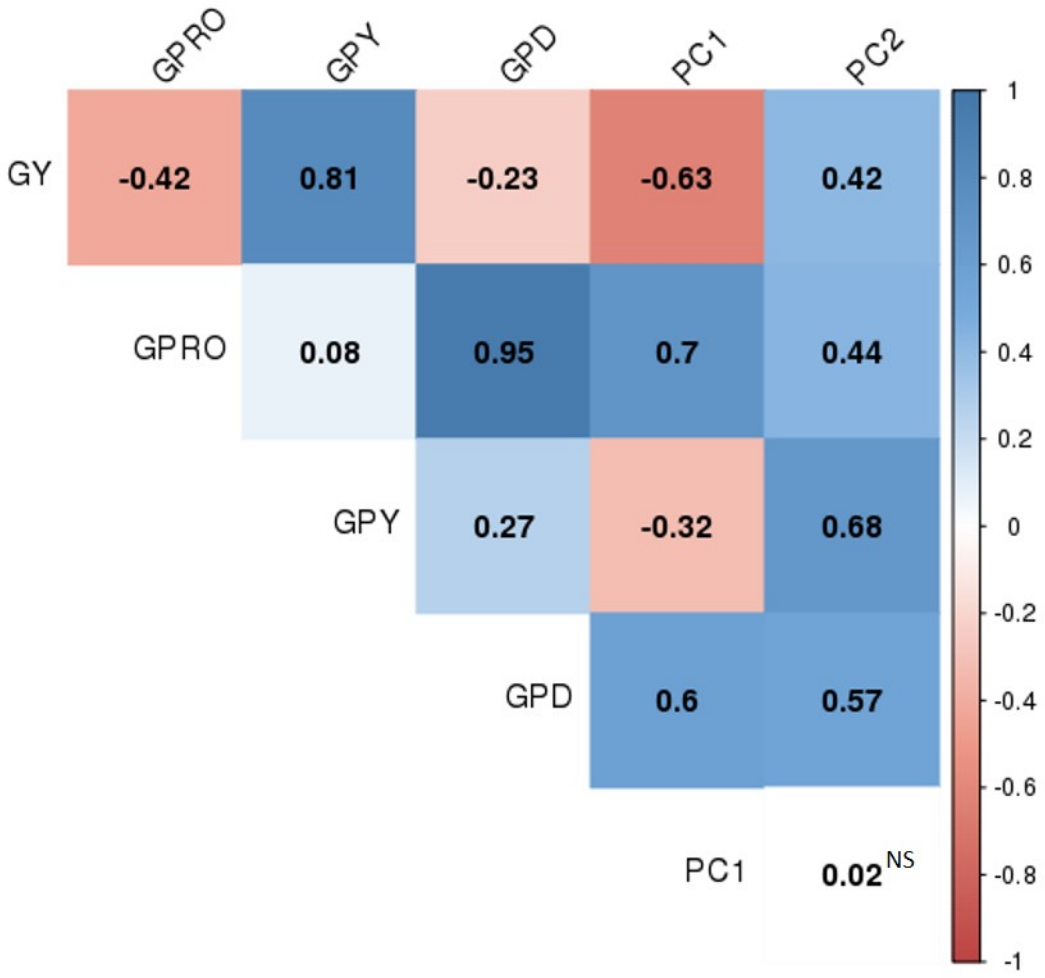


Figure 3.10. Matrix of the Pearson's correlation coefficients between measured phenotypes (wheat grain yield, GY; grain protein concentration, GPRO; grain protein yield, GPY; grain protein deviation, GPD) and the first two principal components (PC) derived from the realized relationship matrix (G matrix). Trait values were best linear unbiased predictors (BLUPS) for 661 breeding lines and varieties from the 2012 and 2013 selection cohorts. All correlations are significant ($p < .05$) except for that which relates PC1 to PC2, which is non-significant (NS). Intensity of color shading scales with the value of the correlation coefficient.

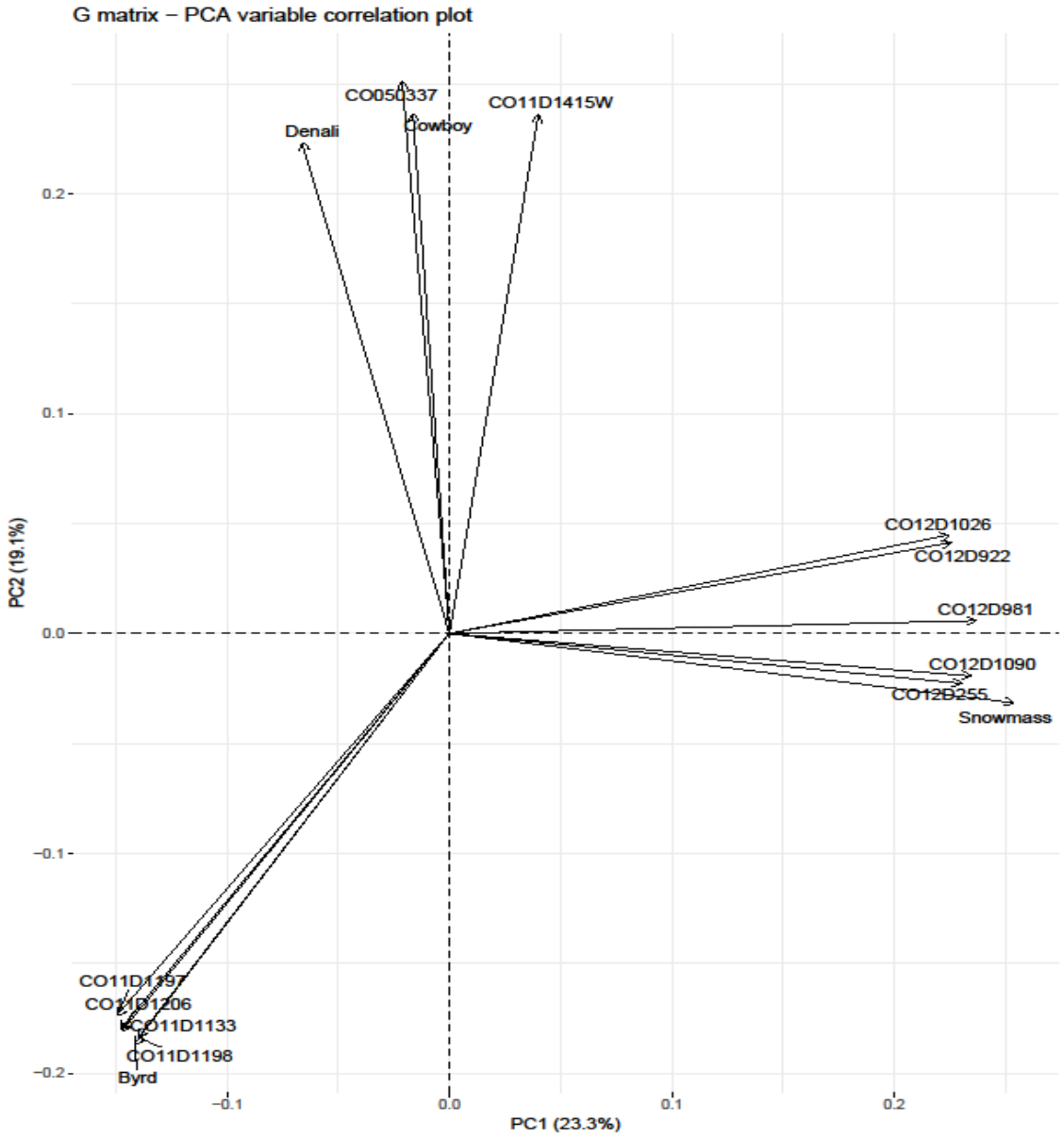


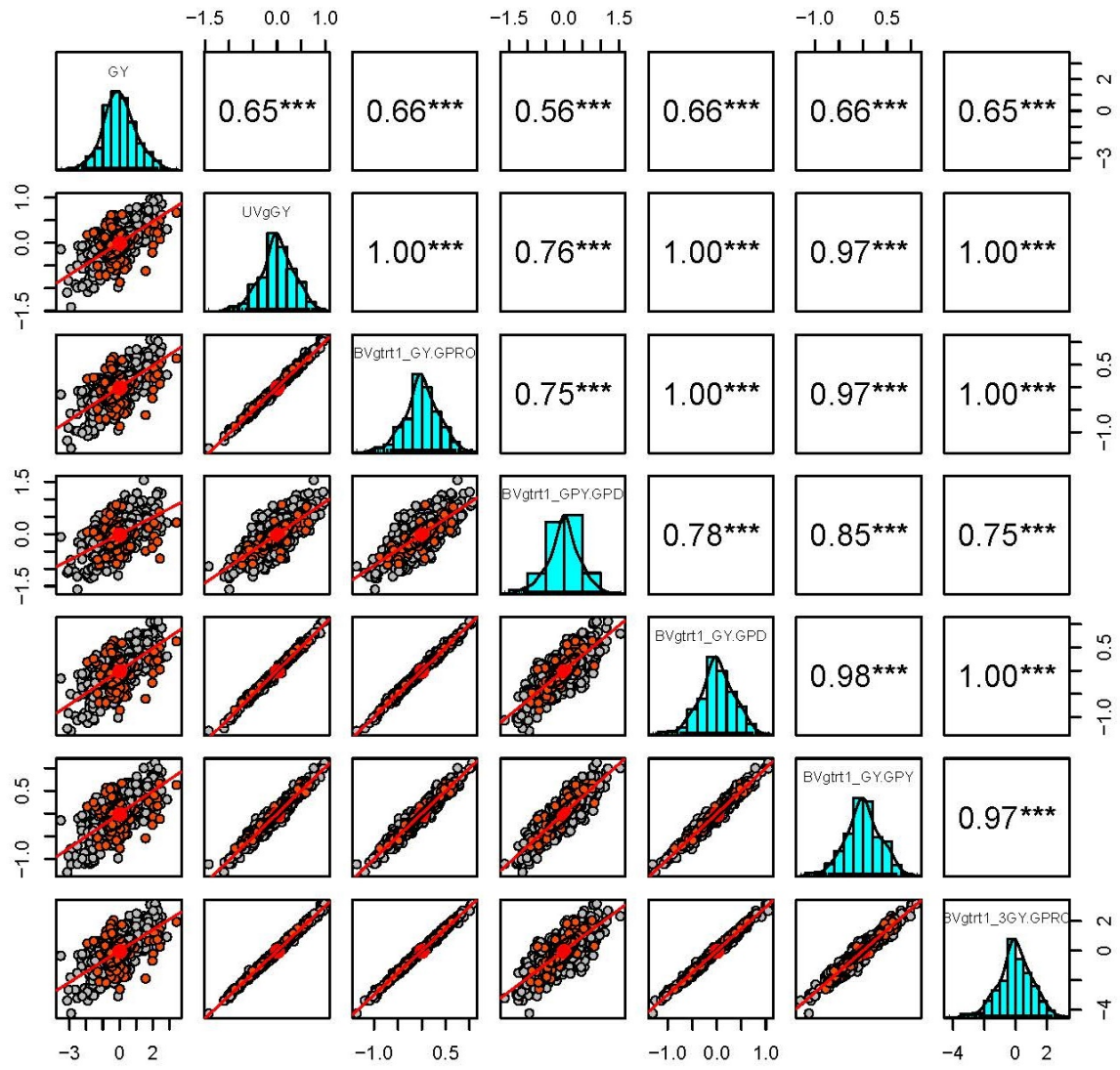
Figure 3.11. Principal components biplot of PC1 and PC2 estimated from the realized relationship matrix (G) for 775 winter wheat varieties and advanced breeding lines in the Colorado State University evaluation trials in the 2012-2014 growing seasons. Vectors for the top 15 squared cosines (cos2) values show the importance of the vector for each observation and the relative contribution of the component to the distance from the center. Each grouping of vectors share half- or full-sib pedigrees.

Table 3.10. Predictive ability in cross-validation (CV) and forward prediction (FP) in univariate genomic selection models for winter wheat varieties and advanced breeding lines in the Colorado State University evaluation trials in the 2012-2015 growing seasons. Selection cohorts were defined by initial year of entry in the trials and categorized model training and validation sets. The 2012 training set included 405 genotypes and the 2012 & 2013 training set included 660 genotypes. The validation sets for 2013 and 2014 contained 255 and 85 genotypes. Narrow sense heritability (h^2) was calculated from variance components estimated from mixed models. Cross-validation (CV) and forward prediction (FP) ability and narrow sense heritability (h^2) are reported as the average over 5 repeats of 5-fold cross-validation.

Univariate Model	*TRAIN*	*VAL*	*CV* $r_{BLUP\|GEBV}$	*sd* (CV)	h^2	$sd\ (h^2)$	*FP* $r_{BLUP\|GEBV}$	*sd* (FP)
Grain yield (GY)	2012	2013	0.703	0.004	0.234	0.009	0.286	0.012
	2012&2013	2014	0.693	0.006	0.216	0.009	0.155	0.008
Grain protein (GPRO)	2012	2013	0.670	0.007	0.198	0.008	0.446	0.013
	2012&2013	2014	0.772	0.003	0.309	0.011	0.191	0.018
Grain protein yield (GPY)	2012	2013	0.788	0.008	0.388	0.013	0.235	0.019
	2012&2013	2014	0.772	0.005	0.342	0.008	-0.072	0.012
Grain protein deviation (GPD)	2012	2013	0.693	0.007	0.226	0.006	0.437	0.012
	2012&2013	2014	0.788	0.004	0.343	0.012	0.076	0.015
Selection index GY + GPRO	2012	2013	0.718	0.005	0.271	0.004	0.266	0.015
	2012&2013	2014	0.730	0.007	0.272	0.007	-0.188	0.020
Selection index 3*GY + GPRO	2012	2013	0.716	0.003	0.258	0.005	0.210	0.022
	2012&2013	2014	0.689	0.007	0.212	0.009	0.001	0.015
Selection index GY + GPD	2012	2013	0.760	0.005	0.340	0.006	0.240	0.022
	2012&2013	2014	0.751	0.008	0.308	0.011	0.034	0.019
Selection index GPY + GPD	2012	2013	0.768	0.005	0.349	0.008	0.324	0.007
	2012&2013	2014	0.785	0.005	0.363	0.011	-0.047	0.016

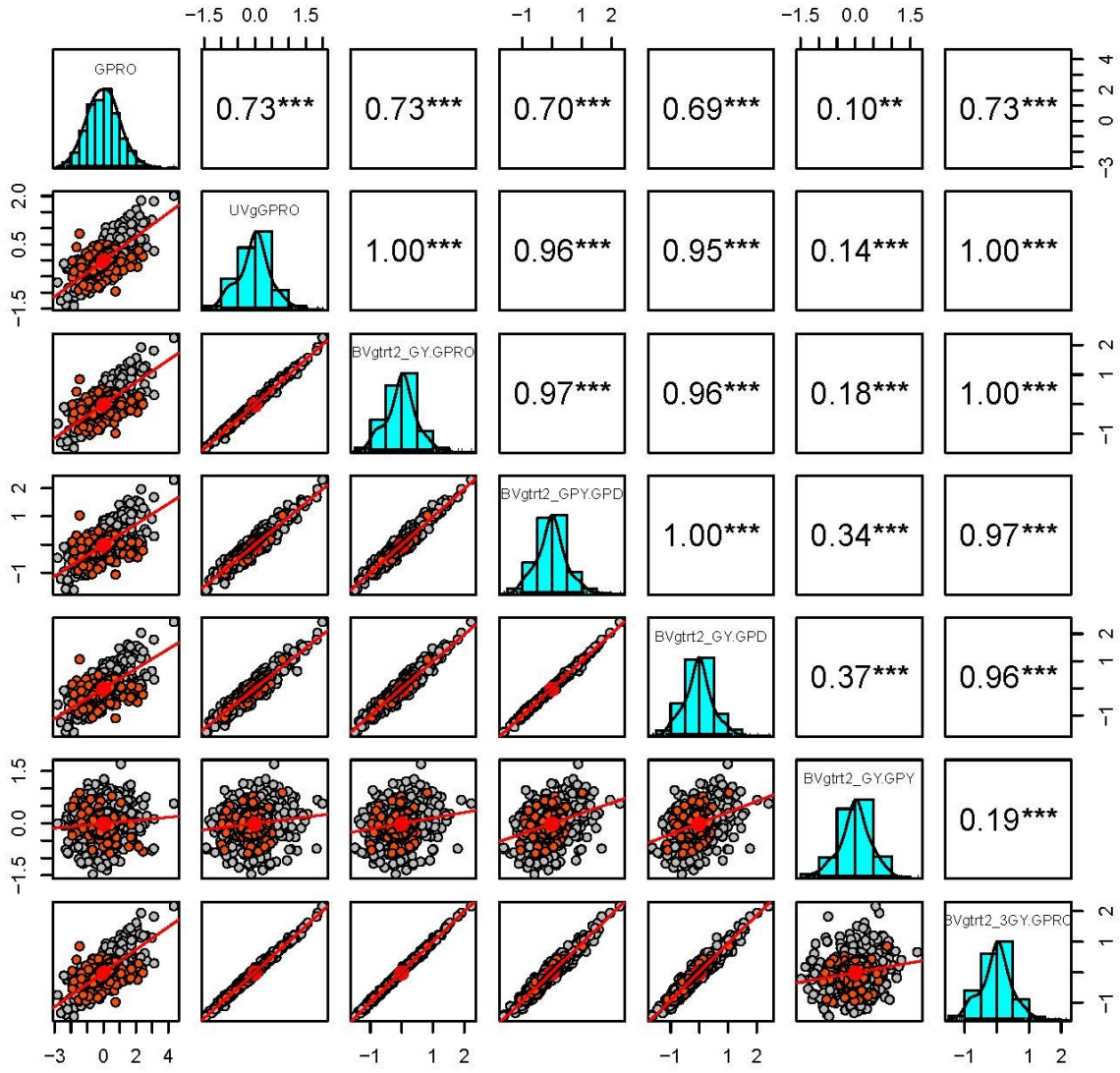
Table 3.11. Predictive ability (*r*) in cross-validation and forward prediction in bivariate genomic selection models for winter wheat varieties and advanced breeding lines in the Colorado State University evaluation trials in the 2012-2015 growing seasons. Selection cohorts were defined by initial year of entry in the trials and categorized model training (TRAIN) and validation (VAL) sets. The 2012 training set included 405 genotypes and the 2012&2013 training set included 660 genotypes. The validation sets for 2013 and 2014 contained 255 and 85 genotypes. Variance components for the training data were estimated from mixed models. Cross-validation (CV) and forward prediction (FP) ability and narrow sense heritability (h^2) are reported as the average over 5 repeats of 5-fold cross-validation for each modeled primary and secondary trait (TRAIT). Genetic and phenotypic covariance of modeled traits were estimated from unstructured covariance matrices. The change in prediction accuracy for bivariate models over univariate models for individual traits is reported as a ratio ('FP ratio' = bivariate *FP r*/ univariate *FP r*).

Bivariate model (1° trait & 2° trait)	*TRAIN*	*VAL*	*TRAIT*	*CV* $r_{BLUP\|GEBV}$	sd (CV)	*covG*	*covP*	h^2	*FP* $r_{BLUP\|GEBV}$	sd (FP)	FP ratio
Grain yield (GY) & grain protein concentration (GPRO)	2012	2013	GY	0.702	0.011	-0.325	-0.453	0.242	0.274	0.007	0.958
			GPRO	0.664	0.006			0.199	0.442	0.018	0.991
	2012&2013	2014	GY	0.693	0.006	-0.390	-0.402	0.220	0.150	0.013	0.968
			GPRO	0.773	0.006			0.317	0.153	0.022	0.801
Grain protein yield (GPY) & Grain protein deviation (GPD)	2012	2013	GPY	0.789	0.006	0.414	0.208	0.397	0.229	0.023	0.974
			GPD	0.682	0.009			0.228	0.426	0.014	0.975
	2012&2013	2014	GPY	0.769	0.009	0.313	0.259	0.341	-0.082	0.014	na
			GPD	0.789	0.003			0.347	0.065	0.008	0.855
Grain yield (GY) & grain protein deviation (GPD)	2012	2013	GY	0.698	0.007	0.043	-0.275	0.244	0.276	0.018	0.965
			GPD	0.684	0.008			0.233	0.420	0.013	0.961
	2012&2013	2014	GY	0.695	0.006	-0.149	-0.213	0.219	0.156	0.012	1.006
			GPD	0.790	0.004			0.364	0.064	0.005	0.842
Grain yield (GY) & grain protein yield (GPY)	2012	2013	GY	0.690	0.007	0.876	0.800	0.247	0.293	0.010	1.024
			GPY	0.784	0.013			0.379	0.240	0.003	1.021
	2012&2013	2014	GY	0.701	0.011	-0.139	-0.222	0.231	0.151	0.009	0.974
			GPY	0.790	0.004			0.353	0.073	0.009	1.013
3*Grain yield (3*GY) & grain protein concentration (GPRO)	2012	2013	3*GY	0.699	0.008	-0.327	-0.449	0.238	0.294	0.012	1.028
			GPRO	0.662	0.006			0.192	0.446	0.015	1.000
	2012&2013	2014	3*GY	0.693	0.007	-0.365	-0.401	0.222	0.137	0.017	0.884
			GPRO	0.768	0.008			0.307	0.141	0.015	0.738



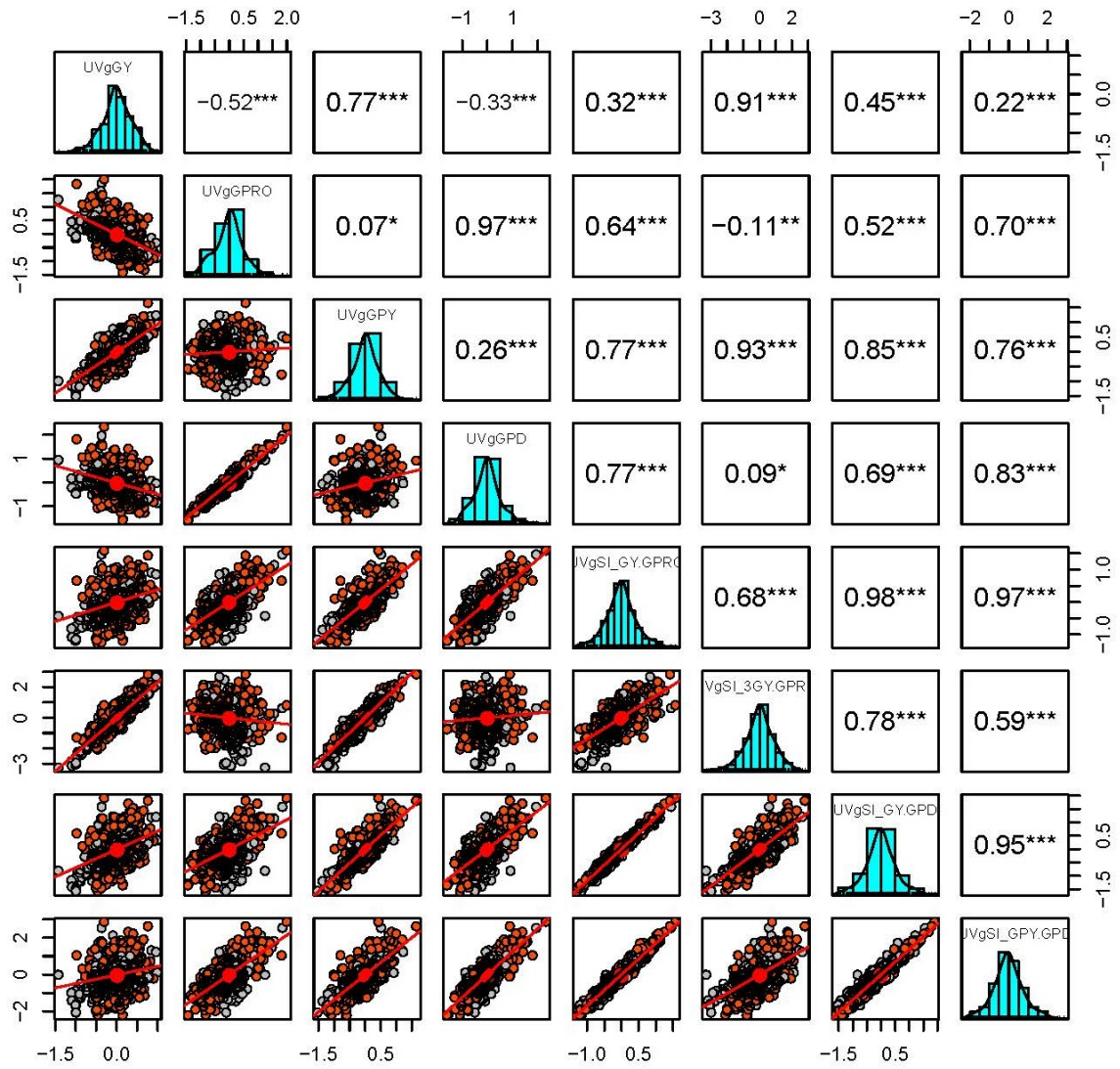
p-value significance levels: 0.01*, 0.001**, 0.0001***

Figure 3.12. Grain yield (GY) related correlation and scatterplot matrix and histograms for GY BLUPs, univariate genomic estimated breeding values (GEBV) for GY, and for primary trait GEBVs derived from bivariate models (BVgtrt1) for GY + GPRO, GPY + GPD, GY + GPD, GY + GPY and 3*GY + GPRO, listed as 'primary.secondary' traits on the histograms. The reported data are from the primary traits. Red circles on the scatterplots represent genotypes in the 2014 validation set and gray circles represent the 2012+2013 training set.



p-value significance levels: 0.01*, 0.001**, 0.0001***

Figure 3.13. Grain protein concentration (GPRO) related correlation and scatterplot matrix and histograms for GPRO BLUPs, univariate genomic estimated breeding values (GEBV) for GPRO, and for secondary trait GEBVs derived from bivariate models (BVgtrt2) for GY + GPRO, GPY + GPD, GY + GPD, GY + GPY and 3*GY + GPRO, listed as 'primary.secondary' traits on the histograms. The reported data are from the secondary traits. Red circles on the scatterplots represent genotypes in the 2014 validation set and gray circles represent the 2012+2013 training set.



p-value significance levels: 0.01*, 0.001**, 0.0001***

Figure 3.14. Correlation and scatterplot matrix and histograms for univariate genomic estimated breeding values (GEBV) for grain yield (GY), grain protein concentration (GPRO), grain protein yield (GPY), grain protein deviation (GPD) and for protein-yield selection indices (UVgSI), GY + GPRO, 3*GY + GPRO, GY + GPD, and GPY + GPD, listed as 'trait1.trait2' on the histograms. Standardized BLUPs for each pair of traits were summed prior to running the prediction model. Red circles on the scatterplots represent genotypes in the 2014 validation set and gray circles represent the 2012+2013 training set.

Table 3.12. Summary statistics for standardized grain yield (zGY) and grain protein concentration (zGPRO) phenotypes for the 2013 validation set (VAL) and for the selected fraction after applying a selection strategy based on genomic estimated breeding values. Dunnett's multiple comparisons were done in contrast to the control group, the selection candidates (VAL).

Selection Strategy[a]	zGY mean	SD	min	max	Dunnett contrast est.	st.err	t val	zGPRO mean	SD	min	max	Dunnett contrast est.	st.err	t val
2013 VAL	-0.02	1.04	-3.01	2.67				-0.19	0.96	-2.54	3.87			
GY	0.20	1.00	-2.28	2.20	0.22	0.15	1.48	-0.63	0.76	-2.09	1.13	-0.44	0.14	-3.04
GPRO	-0.51	1.05	-2.67	2.67	-0.49	0.15	-3.30*	0.55	1.00	-1.80	3.87	0.74	0.14	5.15***
GPY	0.23	0.88	-1.78	2.20	0.25	0.15	1.69	-0.23	0.90	-1.92	1.35	-0.04	0.14	-0.26
GPD	-0.40	1.08	-2.67	2.42	-0.38	0.15	-2.57	0.59	1.00	-1.69	3.87	0.78	0.14	5.45***
UVIC GY + GPRO	0.15	0.96	-2.28	2.08	0.17	0.15	1.15	-0.18	0.85	-2.09	1.35	0.01	0.14	0.09
UVIC GPY + GPD	0.17	0.98	1.79	2.53	0.19	0.15	1.27	0.17	0.83	-1.92	2.19	0.37	0.14	2.55
UVSI_GY + GPRO	0.05	0.95	-1.79	2.42	0.07	0.15	0.47	0.21	1.09	-1.92	3.87	0.40	0.14	2.78
UVSI_3GY + GPRO	0.17	0.88	-1.72	2.20	0.19	0.15	1.25	-0.38	0.90	-2.07	1.35	-0.18	0.14	-1.28
UVSI_GY + GPD	0.07	0.96	-1.79	2.42	0.09	0.15	0.61	0.13	1.06	-1.69	3.87	0.32	0.14	2.23
UVSI_GPY + GPD	-0.35	1.01	-2.67	2.42	-0.33	0.15	-2.20	0.60	1.00	-1.92	3.87	0.79	0.14	5.49***
BVIC GY + GPRO	0.07	0.97	-2.28	2.08	0.09	0.15	0.59	-0.12	0.88	-2.09	1.35	0.07	0.14	0.49
BVIC GPY + GPD	0.11	1.01	-1.79	2.53	0.13	0.15	0.87	0.34	0.96	-1.92	3.87	0.53	0.14	3.67**
BVSI GY + GPRO	0.07	0.94	-1.79	2.42	0.09	0.15	0.58	0.20	1.07	-1.92	3.87	0.40	0.14	2.75
BVSI 3GY + GPRO	0.22	0.91	-1.54	2.20	0.24	0.15	1.62	-0.36	0.94	-2.07	1.35	-0.16	0.14	-1.14
BVSI GY + GPD	0.05	0.96	-1.79	2.42	0.07	0.15	0.44	0.13	1.04	-1.69	3.87	0.33	0.14	2.27
BVSI GPY + GPD	-0.27	1.01	-2.67	2.42	-0.25	0.15	-1.64	0.52	1.04	-1.92	3.87	0.71	0.14	4.96***

*Significance at the .05 probability level. **Significance at the .01 probability level. ***Significance at the .001 probability level.

[a] GY, grain yield; GPRO, grain protein concentration; GPY, grain protein yield; GPD, grain protein deviation; UVIC, independent culling performed on univariate model predictors for trait 1 or trait 2, designated as trait 1 + trait 2; UVSI, selection index calculated by summing BLUPs (trait 1 + trait 2) prior to univariate modeling; BVIC, independent culling performed on bivariate model predictors for trait 1 or trait 2, designated as trait 1 + trait 2; BVSI, selection index calculated by summing bivariate model predictors (trait 1 + trait 2) after bivariate modeling.

Table 3.13. Summary statistics for standardized grain yield (zGY) and grain protein concentration (zGPRO) genomic estimated breeding values for the 2014 validation set and the selected fraction after applying a selection strategy. Dunnett's multiple comparisons were done in contrast to the control group, the selection candidates (VAL). The zGY means were not significantly different from the VAL, therefore contrasts were not warranted.

Selection Strategy[a]	zGY mean	SD	min	max	zGPRO mean	SD	min	max	Dunnett contrast est.	st.err	t val
2014 VAL	-0.16	0.89	-1.84	2.90	0.43	1.09	-1.64	3.12			
GY	0.20	1.23	-0.88	2.90	-0.29	0.92	-1.64	1.65	-0.72	0.26	-2.75
GPRO	-0.13	0.47	-0.85	1.26	0.33	0.98	-1.35	2.84	-0.11	0.26	-0.41
GPY	-0.01	1.10	-0.88	2.90	-0.39	0.87	-1.64	1.65	-0.82	0.26	-3.13
GPD	-0.12	0.54	-0.71	1.62	0.17	0.76	-1.35	1.26	-0.26	0.26	-0.99
UVIC GY + GPRO	0.01	1.14	-1.84	1.99	0.08	1.10	-1.64	2.36	-0.36	0.26	-1.37
UVIC GPY + GPD	-0.19	0.78	-1.15	1.72	0.18	1.13	-1.64	2.14	-0.25	0.26	-0.97
UVSI_GY.GPRO	0.06	1.02	-0.91	1.99	-0.10	0.98	-1.64	1.65	-0.53	0.26	-2.03
UVSI_3GY.GPRO	0.22	1.23	-0.88	2.90	-0.42	0.86	-1.64	1.65	-0.86	0.26	-3.28*
UVSI_GY.GPD	0.08	1.04	-0.91	1.99	-0.16	0.99	-1.64	1.65	-0.60	0.26	-2.29
UVSI_GPY.GPD	-0.02	0.92	-0.91	1.99	0.00	1.02	-1.64	1.65	-0.43	0.26	-1.65
BVIC GY & GPRO	-0.03	1.16	-1.84	1.99	-0.06	1.14	-1.64	2.36	-0.49	0.26	-1.88
BVIC GPY & GPD	-0.16	0.97	-1.84	1.88	0.21	1.14	-1.64	2.36	-0.23	0.26	-0.86
BVSI GY & GPRO	0.07	1.02	-0.91	1.99	-0.04	0.96	-1.64	1.65	-0.47	0.26	-1.80
BVSI 3GY & GPRO	0.22	1.22	-0.88	2.90	-0.33	0.94	-1.64	1.65	-0.76	0.26	-2.91
BVSI GY & GPD	0.10	1.03	-0.91	1.99	-0.12	0.98	-1.64	1.65	-0.55	0.26	-2.11
BVSI GPY & GPD	-0.02	0.92	-0.91	1.99	0.00	1.02	-1.64	1.65	-0.43	0.26	-1.65

*Significance at the .05 probability level. **Significance at the .01 probability level. ***Significance at the .001 probability level.

[a] GY, grain yield; GPRO, grain protein concentration; GPY, grain protein yield; GPD, grain protein deviation; UVIC, independent culling performed on univariate model predictors for trait 1 or trait 2, designated as trait 1 + trait 2; UVSI, selection index calculated by summing BLUPs (trait 1 + trait 2) prior to univariate modeling; BVIC, independent culling performed on bivariate model predictors for trait 1 or trait 2, designated as trait 1 + trait 2; BVSI, selection index calculated by summing bivariate model predictors (trait 1 + trait 2) after bivariate modeling.

Table 3.14. Tukey's procedure for means comparisons between the 2013 validation set (VAL) and the selected fraction obtained after applying a selection strategy for standardized grain yield (zGY) and grain protein concentration (zGPRO) genomic estimated breeding values.

Linear contrast hypothesis[a]	zGY Estimate	std.error	t value	zGPRO Estimate	std.error	t value
GPY - VAL == 0	0.25	0.15	1.63	-0.04	0.15	-0.25
GPY - BVIC GY & GPRO == 0	0.16	0.20	0.82	-0.11	0.19	-0.56
GPY - BVIC GPY & GPD == 0	0.12	0.20	0.61	-0.56	0.19	-2.98*
GPY - BVSI GPY + GPD == 0	0.50	0.20	2.50	-0.75	0.19	-3.95**
UVSI_GPY + GPD - VAL == 0	-0.33	0.15	-2.13	0.79	0.15	5.37***
UVSI_GPY + GPD - GPY == 0	-0.58	0.20	-2.91*	0.83	0.19	4.35***
UVSI_GPY + GPD - BVIC GY & GPRO == 0	-0.42	0.20	-2.09	0.72	0.19	3.79**
UVSI_GPY + GPD - BVIC GPY & GPD == 0	-0.46	0.20	-2.30	0.26	0.19	1.38
UVSI_GPY + GPD - BVSI GPY + GPD == 0	-0.08	0.20	-0.42	0.08	0.19	0.40
BVIC GY & GPRO - VAL == 0	0.09	0.15	0.58	0.07	0.15	0.48
BVIC GPY & GPD - VAL == 0	0.13	0.15	0.84	0.53	0.15	3.59**
BVSI GPY + GPD - VAL == 0	-0.25	0.15	-1.59	0.71	0.15	4.85***
BVIC GY & GPRO - BVIC GPY & GPD == 0	-0.04	0.20	-0.20	-0.46	0.19	-2.41
BVSI GPY + GPD - BVIC GY & GPRO == 0	-0.34	0.20	-1.68	0.64	0.19	3.39**
BVSI GPY + GPD - BVIC GPY & GPD == 0	-0.38	0.20	-1.88	0.19	0.19	0.98

*Significance at the .05 probability level. **Significance at the .01 probability level. ***Significance at the .001 probability level.

[a]GPY, grain protein yield; VAL, 2013 validation data; BV, bivariate model predictors; IC, independent culling with trait 1 & trait 2; GY, grain yield; GPRO, grain protein concentration; GPD, grain protein deviation; SI, selection index; 'GY + GPRO' or 'GPY + GPD', summed BV predictors; UVSI_GPY + GP D, SI calculated with summed GPY and GPD values and then index values are entered in a univariate model.

References

Alaux, M., Rogers, J., Letellier, T., Flores, R., Alfama, F., Pommier, C., Mohellibi, N., Durand, S., Kimmel, E., Michotey, C., Guerche, C., Loaec, M., Lainé, M., Steinbach, D., Choulet, F., Rimbert, H., Leroy, P., Guilhot, N., Salse, J., …& Quesneville, H. (2018). Linking the International Wheat Genome Sequencing Consortium bread wheat reference genome sequence to wheat genetic and phenomic data [journal article]. *Genome Biology*, *19*(1), 111. https://doi.org/10.1186/s13059-018-1491-4

Asoro, F. G., Newell, M. A., Beavis, W. D., Scott, M. P., Tinker, N. A., & Jannink, J. L. (2013). Genomic, marker-assisted, and pedigree-BLUP selection methods for β-glucan concentration in elite oat. *Crop Science*, *53*(5), 1894-1906.

Baker, D. (1979). Report on cereal foods. *Journal of the Association of Official Analytical Chemists*, *62*(2), 369-370.

Bassi, F. M., Bentley, A. R., Charmet, G., Ortiz, R., & Crossa, J. (2016). Breeding schemes for the implementation of genomic selection in wheat (Triticum spp.). *Plant Science*, *242*, 23-36. https://doi.org/http://dx.doi.org/10.1016/j.plantsci.2015.08.021

Belamkar, V., Guttieri, M. J., Hussain, W., Jarquín, D., El-basyoni, I., Poland, J., Lorenz, A. J., & Baenziger, P. S. (2018). Genomic Selection in Preliminary Yield Trials in a Winter Wheat Breeding Program. *G3: Genes|Genomes|Genetics*, *8*(8), 2735. https://doi.org/10.1534/g3.118.200415

Bernardo, R. (2010). *Breeding for Quantitative Traits in Plants* (2nd ed.). Stemma Press. (2002)

Bogard, M., Allard, V., Brancourt-Hulmel, M., Heumez, E., Machet, J. M., Jeuffroy, M. H., Gate, P., Martre, P., & Le Gouis, J. (2010). Deviation from the grain protein concentration–grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. *Journal of Experimental Botany*, *61*(15), 4303-4312.

Burnham, K. P. (2010). *Model selection and multimodel inference : a practical information-theoretic approach* (2nd ed.. ed.). New York : Springer.

Burnham, K. P., & Anderson, D. R. (2003). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media.

Butler, D. G., Cullis, B. R., Gilmour, A. R., & Gogel, B. J. (2009). *ASReml-R reference manual (Version 3)*. The State of Queensland, Department of Primary Industries and Fisheries. http://www.vsni.co.uk/

Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J., & Thompson, R. (2017). *ASReml-R reference manual (Version 4)*. VSN International Ltd. http://www.vsni.co.uk/

Charmet, G., Tran, L.-G., Auzanneau, J., Rincent, R., & Bouchet, S. (2020). BWGS: AR package for genomic selection and its application to a wheat breeding programme. *PloS one*, *15*(4), e0222733. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7141418/pdf/pone.0222733.pdf

Cooper, J. K., Stromberger, J. A., Morris, C. F., Bai, G., & Haley, S. D. (2016). End-use quality and agronomic characteristics associated with the Glu-B1al high-molecular-weight glutenin allele in US hard winter wheat. *Crop Science*, *56*(5), 2348-2353.

Cormier, F., Faure, S., Dubreuil, P., Heumez, E., Beauchêne, K., Lafarge, S., Praud, S., & Le Gouis, J. (2013). A multi-environmental study of recent breeding progress on nitrogen use efficiency in wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, *126*(12), 3035-3048. https://doi.org/10.1007/s00122-013-2191-9

Crain, J., Mondal, S., Rutkoski, J., Singh, R. P., & Poland, J. (2018). Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding. *The Plant Genome*, *11*(1). https://doi.org/10.3835/plantgenome2017.05.0043

de la Vega, A. J., DeLacy, I. H., & Chapman, S. C. (2007). Progress over 20 years of sunflower breeding in central Argentina. *Field Crops Research*, *100*(1), 61-72.

Doyle, J. J., & Doyle, J. L. (1987). *A rapid DNA isolation procedure for small quantities of fresh leaf tissue*.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PloS one*, *6*(5), e19379. https://doi.org/10.1371/journal.pone.0019379

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

Fehr, W. (1987). Principles of cultivar development: theory and technique. *New York*.

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PloS one*, *9*(2), e90346.

Goesaert, H., Brijs, K., Veraverbeke, W. S., Courtin, C. M., Gebruers, K., & Delcour, J. A. (2005). Wheat flour constituents: how they impact bread quality, and how to impact their functionality. *Trends in Food Science & Technology*, *16*(1), 12-30. https://doi.org/https://doi.org/10.1016/j.tifs.2004.02.011

Graybosch, R., Bockelman, H. E., Garland-Campbell, K. A., Garvin, D. F., & Regassa, T. (2014). Wheat. In J. Specht & B. Carver (Eds.), *Yield Gains in Major U.S. Field Crops* (pp. 459-488). American Society of Agronomy, Inc., Crop Science Society of America,

Inc., and Soil Science Society of America, Inc. https://doi.org/10.2135/cssaspecpub33.c16

Guttieri, M. J., Baenziger, P. S., Frels, K., Carver, B., Arnall, B., & Waters, B. M. (2015). Variation for Grain Mineral Concentration in a Diversity Panel of Current and Historical Great Plains Hard Winter Wheat Germplasm. *Crop Science*, *55*(3), 1035-1052. https://doi.org/10.2135/cropsci2014.07.0506

Guttieri, M. J., Frels, K., Regassa, T., Waters, B. M., & Baenziger, P. S. (2017). Variation for nitrogen use efficiency traits in current and historical great plains hard winter wheat [journal article]. *Euphytica*, *213*(4), 87. https://doi.org/10.1007/s10681-017-1869-5

Haile, J. K., N'Diaye, A., Clarke, F., Clarke, J., Knox, R., Rutkoski, J., Bassi, F. M., & Pozniak, C. J. (2018). Genomic selection for grain yield and quality traits in durum wheat. *Molecular Breeding*, *38*(6), 75. https://doi.org/10.1007/s11032-018-0818-x

Haley, S. D., Johnson, J. J., Peairs, F. B., Quick, J. S., Stromberger, J. A., Clayshulte, S. R., Butler, J. D., Rudolph, J. B., Seabourn, B. W., Bai, G., Jin, Y., & Kolmer, J. (2007). Registration of 'Ripper' Wheat *J. Plant Reg.*, *1*(1), 1-6. https://doi.org/10.3198/jpr2006.10.0689crc

Haley, S. D., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Heaton, E. E., Seifert, S. A., Kottke, R. A., Rudolph, J. B., Martin, T. J., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Seifers, D. L., Chen, M.-S., & Seabourn, B. W. (2011). Registration of 'Snowmass' Wheat *J. Plant Reg.*, *5*(1), 87-90. https://doi.org/10.3198/jpr2010.03.0175crc

Haley, S. D., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Hudson-Arns, E. E., Seifert, S. A., Anderson, V. A., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Chen, M.-S., & Seabourn, B. W. (2017). Registration of 'Sunshine' Hard White Winter Wheat. *Journal of Plant Registrations*, *11*(3), 289-294. https://doi.org/10.3198/jpr2016.12.0075crc

Haley, S. D., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Hudson-Arns, E. E., Seifert, S. A., Valdez, V. A., Kottke, R. A., Rudolph, J. B., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Chen, M.-S., Seabourn, B. W., & Dowell, F. E. (2014). Registration of 'Antero' Wheat. *J. Plant Reg.*, *8*(2), 165-168. https://doi.org/10.3198/jpr2013.12.0072crc

Haley, S. D., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Hudson, E. E., Seifert, S. A., Kottke, R. A., Valdez, V. A., Rudolph, J. B., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Chen, M.-S., & Seabourn, B. W. (2012a). Registration of 'Byrd' Wheat. *J. Plant Reg.*, *6*(3), 302-305. https://doi.org/10.3198/jpr2011.12.0672crc

Haley, S. D., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Hudson, E. E., Seifert, S. A., Kottke, R. A., Valdez, V. A., Rudolph, J. B., Martin, T. J., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Chen, M.-S., & Seabourn, B. W. (2012b). Registration of 'Denali' Wheat. *J. Plant Reg.*, *6*(3), 311-314. https://doi.org/10.3198/jpr2011.12.0675crc

Haley, S. D., Johnson, J. J., Westra, P. H., Peairs, F. B., Stromberger, J. A., Hudson, E. E., Seifert, S. A., Kottke, R. A., Valdez, V. A., Rudolph, J. B., Bai, G., Chen, X., Bowden, R. L., Jin, Y., Kolmer, J. A., Chen, M.-S., & Seabourn, B. W. (2012c). Registration of 'Brawl CL Plus' Wheat. *J. Plant Reg.*, *6*(3), 306-310. https://doi.org/10.3198/jpr2011.12.0673crc

Hayes, B., Panozzo, J., Walker, C., Choy, A., Kant, S., Wong, D., Tibbits, J., Daetwyler, H., Rochfort, S., & Hayden, M. (2017). Accelerating wheat breeding for end-use quality with multi-trait genomic predictions incorporating near infrared and nuclear magnetic resonance-derived phenotypes. *Theoretical and Applied Genetics*, *130*(12), 2505-2519. https://link.springer.com/article/10.1007%2Fs00122-017-2972-7

Hazel, & Lush, J. L. (1942). The efficiency of three methods of selection. *Journal of Heredity*, *33*(11), 393-399.

Hazel, L. N. (1943). The genetic basis for constructing selection indexes. *Genetics*, *28*(6), 476-490. https://www.genetics.org/content/genetics/28/6/476.full.pdf

Heffner, E. L., Sorrells, M. E., & Jannink, J. L. (2009). Genomic Selection for Crop Improvement. *49*, 1-12.

Henderson, C. (1984). Estimation of variances and covariances under multiple trait models. *Journal of Dairy Science*, *67*(7), 1581-1589.

Iqbal, M., Navabi, A., Salmon, D., Yang, R. C., & Spaner, D. (2007). Simultaneous selection for early maturity, increased grain yield and elevated grain protein content in spring wheat. *Plant Breeding*, *126*(3), 244-250.

Jia, Y., & Jannink, J.-L. (2012). Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics*, *192*(4), 1513-1522.

John, J. A., & Williams, E. R. (1995). *Cyclic and computer generated designs.* Chapman and Hall.

Kempthorne, O., & Nordskog, A. W. (1959). Restricted selection indices. *Biometrics*, *15*(1), 10-19.

Kloke, J. D., & McKean, J. W. (2012). Rfit: Rank-based Estimation for Linear Models. *The R Journal*, *4*(2), 57-64.

Koekemoer, F., Labuschagne, M., & Van Deventer, C. (1999). A selection strategy for combining high grain yield and high protein content in South African wheat cultivars. *Cereal Research Communications*, 107-114.

Lado, B., Vázquez, D., Quincke, M., Silva, P., Aguilar, I., & Gutiérrez, L. (2018). Resource allocation optimization with multi-trait genomic prediction for bread wheat (Triticum aestivum L.) baking quality [journal article]. *Theoretical and Applied Genetics*, *131*(12), 2719-2731. https://doi.org/10.1007/s00122-018-3186-3

Lollato, R. P., & Edwards, J. T. (2015). Maximum Attainable Wheat Yield and Resource-Use Efficiency in the Southern Great Plains. *Crop Science*, *55*(6), 2863-2876. https://doi.org/10.2135/cropsci2015.04.0215

Lozada, D. N., & Carter, A. H. (2019). Accuracy of single and multi-trait genomic prediction models for grain yield in US Pacific Northwest winter wheat. *Crop Breeding, Genetics and Genomics*, *1*(1).

Lozada, D. N., Mason, R. E., Sarinelli, J. M., & Brown-Guedira, G. (2019). Accuracy of genomic selection for grain yield and agronomic traits in soft red winter wheat. *BMC Genetics*, *20*(1), 82. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6823964/pdf/12863_2019_Article_785.pdf

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*, *157*(4), 1819-1829. http://www.genetics.org/content/157/4/1819.abstract

Michel, S., Ametz, C., Gungor, H., Epure, D., Grausgruber, H., Löschenberger, F., & Buerstmayr, H. (2016). Genomic selection across multiple breeding cycles in applied bread wheat breeding [journal article]. *Theoretical and Applied Genetics*, 1-11. https://doi.org/10.1007/s00122-016-2694-2

Michel, S., Kummer, C., Gallee, M., Hellinger, J., Ametz, C., Akgöl, B., Epure, D., Löschenberger, F., & Buerstmayr, H. (2018). Improving the baking quality of bread wheat by genomic selection in early generations [journal article]. *Theoretical and Applied Genetics*, *131*(2), 477-493. https://doi.org/10.1007/s00122-017-2998-x

Michel, S., Löschenberger, F., Ametz, C., Pachler, B., Sparry, E., & Bürstmayr, H. (2019a). Combining grain yield, protein content and protein quality by multi-trait genomic selection in bread wheat. *Theoretical and Applied Genetics*, *132*(10), 2767-2780. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6763414/pdf/122_2019_Article_3386.pdf

Michel, S., Löschenberger, F., Ametz, C., Pachler, B., Sparry, E., & Bürstmayr, H. (2019b). Simultaneous selection for grain yield and protein content in genomics-assisted wheat breeding. *Theoretical and Applied Genetics*, *132*(6), 1745-1760. https://doi.org/10.1007/s00122-019-03312-5

Monaghan, J. M., Snape, J. W., Chojecki, A. J. S., & Kettlewell, P. S. (2001). The use of grain protein deviation for identifying wheat cultivars with high grain protein concentration and yield. *Euphytica*, *122*(2), 309-317.

Neuweiler, J. E., Maurer, H. P., & Würschum, T. (2021). Genetic architecture of phenotypic indices for simultaneous improvement of protein content and grain yield in triticale (× triticosecale). *Plant Breeding*.

Nuttall, J. G., O'Leary, G. J., Panozzo, J. F., Walker, C. K., Barlow, K. M., & Fitzgerald, G. J. (2017). Models of grain quality in wheat—A review. *Field Crops Research*, *202*, 136-145. https://doi.org/https://doi.org/10.1016/j.fcr.2015.12.011

Oury, F.-X., & Godin, C. (2007). Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favourable genotypes? *Euphytica*, *157*(1-2), 45-57. https://doi.org/10.1007/s10681-007-9395-5

Peairs, F., & Armenta, R. (2010). Wheat Production and Pest Management for the Great Plains Region. *Colorado State University Extension*.

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J.-L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PloS one*, *7*(2), e32253. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3289635/pdf/pone.0032253.pdf

Poland, J. A., Endelman, J., Dawson, J., Rutkoski, J., Wu, S. Y., Manes, Y., Dreisigacker, S., Crossa, J., Sanchez-Villeda, H., Sorrells, M., & Jannink, J. L. (2012). Genomic selection in wheat breeding using Genotyping-by-Sequencing. *The Plant Genome*, *5*(3), 103-113. https://doi.org/10.3835/plantgenome2012.06.0006

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. In R Foundation for Statistical Computing. https://www.R-project.org/

Rapp, M., Lein, V., Lacoudre, F., Lafferty, J., Müller, E., Vida, G., Bozhanova, V., Ibraliu, A., Thorwarth, P., Piepho, H. P., Leiser, W. L., Würschum, T., & Longin, C. F. H. (2018). Simultaneous improvement of grain yield and protein content in durum wheat by different phenotypic indices and genomic selection. *Theoretical and Applied Genetics*, *131*(6), 1315-1329. https://doi.org/10.1007/s00122-018-3080-z

Rasheed, A., & Xia, X. (2019). From markers to genome-based breeding in wheat. *Theoretical and Applied Genetics*, *132*(3), 767-784. https://link.springer.com/article/10.1007%2Fs00122-019-03286-4

Runcie, D., & Cheng, H. (2019). Pitfalls and Remedies for Cross Validation with Multi-trait Genomic Prediction Methods. *G3: Genes|Genomes|Genetics*, *9*(11), 3727-3741. https://doi.org/10.1534/g3.119.400598

Rutkoski, J., Singh, R. P., Huerta-Espino, J., Bhavani, S., Poland, J., Jannink, J. L., & Sorrells, M. E. (2015). Genetic Gain from Phenotypic and Genomic Selection for Quantitative Resistance to Stem Rust of Wheat. *Plant Genome*, *8*(2). https://doi.org/10.3835/plantgenome2014.10.0074

Schulthess, A. W., Wang, Y., Miedaner, T., Wilde, P., Reif, J. C., & Zhao, Y. (2016). Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes [journal article]. *Theoretical and Applied Genetics*, *129*(2), 273-287. https://doi.org/10.1007/s00122-015-2626-6

Shelton, D., Martin, G., & Peterson, B. (2008). *Wheat and flour testing methods: A guide to understanding wheat and flour quality, Version 2* [technical guide]. Kansas State University Agricultural Experiment Station and Cooperative Extension Service. https://webdoc.agsci.colostate.edu/wheat/linksfiles/WheatFlour.pdf

Simmonds, N. W. (1995). The relation between yield and protein in cereal grain. *Journal of the Science of Food and Agriculture*, *67*(3), 309-315. https://doi.org/10.1002/jsfa.2740670306

Smith, H. F. (1936). A discriminant function for plant selection. *Annals of eugenics*, *7*(3), 240-250.

Thorwarth, P., Liu, G., Ebmeyer, E., Schacht, J., Schachschneider, R., Kazman, E., Reif, J. C., Würschum, T., & Longin, C. F. H. (2019). Dissecting the genetics underlying the relationship between protein content and grain yield in a large hybrid wheat population. *Theoretical and Applied Genetics*, *132*(2), 489-500. https://doi.org/10.1007/s00122-018-3236-x

Thorwarth, P., Piepho, H. P., Zhao, Y., Ebmeyer, E., Schacht, J., Schachschneider, R., Kazman, E., Reif, J. C., Würschum, T., & Longin, C. F. H. (2018). Higher grain yield and higher grain protein deviation underline the potential of hybrid wheat for a sustainable agriculture. *Plant Breeding*, *137*(3), 326-337. https://doi.org/10.1111/pbr.12588

VanRaden, P. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91 11*, 4414-4423.

Ward, B. P., Brown-Guedira, G., Tyagi, P., Kolb, F. L., Van Sanford, D. A., Sneller, C. H., & Griffey, C. A. (2019). Multienvironment and Multitrait Genomic Selection Models in Unbalanced Early-Generation Wheat Yield Trials. *Crop Science*, *59*(2), 491-507. https://doi.org/10.2135/cropsci2018.03.0189

Williams, E., Piepho, H.-P., & Whitaker, D. (2011). Augmented p-rep designs. *Biometrical Journal*, *53*(1), 19-27. https://doi.org/10.1002/bimj.201000102

Williams, E. R. (1986). Row and column designs with contiguous replicates. *Australian Journal of Statistics*, *28*(2), 154-163. https://doi.org/https://doi.org/10.1111/j.1467-842X.1986.tb00594.x

Yao, J., Zhao, D., Chen, X., Zhang, Y., & Wang, J. (2018). Use of genomic selection and breeding simulation in cross prediction for improvement of yield and quality in wheat (*Triticum aestivum* L.). *The Crop Journal*, *6*(4), 353-365. https://doi.org/https://doi.org/10.1016/j.cj.2018.05.003

Zhong, S., & Jannink, J.-L. (2007). Using quantitative trait loci results to discriminate among crosses on the basis of their progeny mean and variance. *Genetics*, *177*(1), 567-576.

Zhou, Y., Vales, M. I., Wang, A., & Zhang, Z. (2016). Systematic bias of correlation coefficient may explain negative accuracy of genomic prediction. *Briefings in bioinformatics*, *18*(5), 744-753. https://doi.org/10.1093/bib/bbw064

APPENDIX

Supplementary table 1. Plot level data for grain yield (GY) and grain protein concentration (GPRO) for breeding lines and commercial varieties grown in 32 evaluation trials during the 2012, 2013, 2014 and 2015 growing seasons. The year that individual genotypes were first entered in yield trials is indicated in the 'GRP_2' column. Environment (ENV) is defined as the combination of year, location, and trial name (open "SupplementaryTable1_Latshaw").

Supplementary table 2. Genotypic values (best linear unbiased estimates), average number of reps, and weights by environment for 676 hard winter wheat breeding lines and varieties grown in 20 environments during the 2012-2014 seasons in Colorado (open "SupplementaryTable2_Latshaw").

Supplementary table 3. Genotypic values (best linear unbiased predictors) for 676 hard winter wheat breeding lines and varieties determined in a two-stage weighted combined analysis across 20 environments during the 2012-2014 seasons in Colorado.

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
Above	0.121	0.169	0.031	0.077
Ankor	-0.127	-0.165	-0.009	-0.068
Antero	0.324	-2.826	0.025	-0.060
Armour	-0.077	2.890	0.003	0.330
Avalanche	-0.148	2.874	-0.005	0.503
Avery	0.364	-0.544	0.037	0.300
BillBrown	0.054	-1.830	0.003	-0.236
BondCL	-0.167	-3.885	-0.026	-0.646
BrawlCLPlus	-0.021	9.059	0.026	1.144
Byrd	0.295	-3.613	0.021	-0.168
CO02W214	-0.616	11.554	-0.023	1.272
CO02W237	-0.447	3.460	-0.031	0.158
CO02W280	-0.378	1.125	-0.048	-0.037
CO03064	-0.102	2.787	-0.010	0.305
CO03064.2	-0.220	3.403	-0.012	0.265
CO03443	-0.282	5.556	-0.022	0.656
CO03W043	-0.292	0.537	-0.036	-0.060
CO03W108	-0.178	1.041	-0.017	-0.061
CO03W127	-0.121	-1.903	0.002	-0.144
CO03W139	-0.142	-3.714	-0.005	-0.535
CO03W146	-0.215	1.708	-0.012	0.193
CO04025	-0.032	2.653	0.010	0.496
CO04039	-0.074	6.144	-0.005	0.734

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO04092	-0.047	2.183	0.008	0.297
CO04111	-0.191	2.640	-0.009	0.242
CO04113	-0.389	6.064	-0.017	0.663
CO04127	-0.306	3.814	-0.001	0.467
CO04227	-0.135	-1.295	-0.033	-0.440
CO04262	-0.204	-0.030	-0.022	-0.094
CO04344	-0.341	-1.482	-0.046	-0.528
CO04393	0.085	2.060	0.012	0.424
CO04447	-0.157	5.347	0.004	0.597
CO04448	-0.256	1.128	-0.024	-0.052
CO04454W	-0.105	0.825	-0.011	0.082
CO04475	-0.108	1.821	-0.005	0.194
CO04499	-0.011	-0.462	0.004	-0.140
CO04544	0.171	-1.876	0.002	-0.022
CO04549	-0.622	2.182	-0.058	0.191
CO04551	-0.124	2.511	-0.019	0.436
CO04553	-0.055	3.397	0.008	0.465
CO04555	0.288	-2.304	0.019	-0.177
CO04574	-0.024	0.342	-0.006	0.042
CO04575	-0.160	-1.373	-0.025	-0.081
CO04W010	-0.123	-0.723	-0.014	0.014
CO04W014	-0.112	-1.463	0.005	-0.144
CO04W028	-0.451	4.192	-0.050	0.158
CO04W029	-0.250	3.845	-0.019	0.258
CO04W038	-0.050	-5.358	-0.018	-0.787
CO04W051	-0.565	1.780	-0.053	-0.156

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO04W061	0.239	0.852	0.008	0.095
CO04W069	0.087	-2.607	0.003	-0.337
CO04W075	0.502	-1.712	0.024	-0.105
CO04W095	0.060	-0.999	0.003	-0.163
CO04W097	0.045	1.974	0.013	0.099
CO04W119	-0.101	-1.958	-0.033	-0.339
CO04W128	-0.271	-0.862	-0.039	-0.185
CO04W135	-0.331	-0.821	-0.025	-0.215
CO04W138	-0.663	1.228	-0.057	-0.132
CO04W164	-0.358	3.466	-0.028	0.176
CO04W179	-0.123	4.017	-0.011	0.536
CO04W188	0.070	-1.650	-0.007	-0.127
CO04W205	-0.006	-4.565	-0.022	-0.739
CO04W210	-0.115	-1.001	-0.026	-0.107
CO04W216	-0.063	0.801	-0.005	-0.051
CO04W281	-0.107	2.585	-0.009	0.366
CO04W299	0.115	4.439	0.017	0.528
CO04W320	-0.407	0.275	-0.032	-0.084
CO04W320.1	-0.649	0.341	-0.061	-0.139
CO04W320.4	-0.030	-1.894	-0.009	-0.250
CO04W323	-0.371	-0.349	-0.043	-0.102
CO04W323.1	-0.334	-1.456	-0.033	-0.327
CO04W369	0.017	0.882	0.008	0.135
CO04W421	-0.205	0.445	-0.028	0.158
CO050133	-0.471	1.999	-0.030	0.186
CO050141	-0.160	4.811	0.008	0.694

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO050165	-0.166	6.268	0.016	0.717
CO050173	0.159	-0.913	0.015	-0.160
CO050175	0.038	-0.611	0.001	-0.086
CO050203	-0.042	-1.775	-0.008	-0.209
CO050217	-0.265	4.157	-0.021	0.598
CO050233	0.238	0.324	0.020	0.230
CO050233.2	0.430	2.247	0.042	0.543
CO050262	0.215	-6.492	0.006	-0.668
CO050270	0.239	-7.157	0.010	-0.634
CO050270.1	0.097	-4.890	0.004	-0.422
CO050303	0.355	-0.695	0.032	0.009
CO050337	0.375	-1.854	0.016	-0.139
CO050337.2	0.295	-3.606	0.015	-0.348
CO050343	0.162	-0.076	0.015	0.075
CO050476	0.118	3.562	0.026	0.535
CO050541	-0.186	2.902	-0.014	0.094
CO05060	-0.089	1.311	0.000	0.153
CO05066	0.008	1.860	-0.004	0.265
CO05066.1	0.138	1.729	0.014	0.246
CO05068	-0.142	-2.004	-0.013	-0.244
CO05079	-0.010	1.118	0.007	0.301
CO05088	-0.207	-3.473	-0.038	-0.552
CO05090	0.069	0.974	0.002	0.006
CO05W001	0.021	-2.572	-0.003	-0.349
CO05W006	-0.311	1.462	-0.018	0.208
CO05W020	0.110	1.447	0.014	0.212

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO05W022	-0.375	4.163	-0.021	0.420
CO05W024	-0.005	2.090	-0.004	0.169
CO05W045	0.243	3.362	0.025	0.583
CO05W056	0.041	4.256	0.034	0.533
CO05W059	-0.029	-0.785	-0.005	-0.044
CO05W062	0.223	-3.352	0.016	-0.201
CO05W064	0.023	1.573	0.010	0.228
CO05W067	-0.216	-1.782	-0.008	-0.107
CO05W101	-0.381	-2.749	-0.046	-0.583
CO05W104	-0.153	-7.812	-0.029	-1.160
CO05W111	-0.077	2.428	0.003	0.209
CO05W112	0.052	-1.353	0.001	0.034
CO05W115	-0.127	0.304	-0.010	0.020
CO05W130	-0.016	2.498	0.015	0.281
CO05W150	0.005	-0.092	0.014	-0.041
CO05W153	-0.074	3.378	0.003	0.450
CO05W156	-0.054	-1.137	-0.003	-0.161
CO05W165	-0.387	4.379	-0.025	0.550
CO05W171	-0.118	4.623	0.006	0.622
CO05W176	-0.260	10.104	0.024	1.237
CO05W180	-0.338	6.645	-0.010	0.664
CO05W194	0.132	-2.297	0.010	-0.111
CO05W250	-0.092	1.162	-0.011	0.159
CO06024	-0.217	8.117	-0.014	0.923
CO06041	-0.158	3.567	0.001	0.498
CO06044	-0.154	5.805	-0.005	0.676

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO06049	-0.192	2.080	-0.011	0.256
CO06059	-0.135	4.609	-0.008	0.565
CO06065	0.055	-1.272	0.005	-0.147
CO06072	0.006	-3.364	0.001	-0.384
CO06093	-0.016	1.919	0.005	0.310
CO06107	-0.027	1.759	0.008	0.343
CO06129	0.099	2.494	0.016	0.394
CO06138	-0.041	-2.863	-0.023	-0.493
CO06277	-0.181	3.590	-0.012	0.362
CO06424.F10	0.277	-3.496	0.015	-0.131
CO06424.F2	0.579	-3.000	0.046	-0.171
CO06530	0.236	-0.272	0.013	0.109
CO06531	-0.080	-3.067	0.000	-0.415
CO06533	0.343	-4.856	0.022	-0.543
CO06534	0.279	-4.753	0.013	-0.412
CO06535	0.046	-6.795	-0.015	-0.774
CO06539	0.079	-0.129	0.014	0.332
CO06540	0.168	-2.997	0.011	-0.245
CO06542	0.022	-2.217	-0.002	-0.305
CO06M240	0.156	-3.954	-0.004	-0.521
CO06M242	-0.244	-4.669	-0.042	-0.729
CO06M243	-0.054	-4.420	-0.037	-0.608
CO06W002	-0.251	9.007	0.002	1.196
CO06W058	-0.195	5.415	-0.001	0.639
CO06W091	0.083	1.614	0.005	0.303
CO06W096	-0.138	1.055	-0.013	0.082

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO06W153	-0.137	-1.626	-0.020	-0.330
CO06W183	0.128	3.324	0.028	0.624
CO06W216	-0.016	-1.799	-0.009	-0.171
CO06W217	-0.063	0.650	0.000	0.115
CO07008	-0.017	3.906	0.016	0.537
CO07033	0.133	1.946	0.018	0.318
CO07078	-0.173	1.983	-0.002	0.095
CO07092	-0.110	7.927	0.005	1.048
CO07202	0.260	-5.565	0.022	-0.419
CO07203	0.146	-2.588	0.016	-0.236
CO07205	0.041	-1.347	-0.003	-0.054
CO07208	0.279	-4.871	0.019	-0.526
CO07247	-0.045	2.311	0.008	0.236
CO07253	-0.043	0.109	-0.012	0.092
CO07274	0.034	2.043	0.007	0.258
CO07279	0.176	6.508	0.033	1.086
CO07279.F1	-0.093	4.449	0.012	0.600
CO07282	-0.009	5.897	0.010	0.668
CO07282.F1	0.114	-0.641	0.000	-0.064
CO07288	-0.076	2.143	0.004	0.335
CO07290	0.342	2.889	0.040	0.622
CO07292	0.087	1.542	0.007	0.341
CO07292.F3	0.185	-3.498	0.016	-0.428
CO07293	-0.179	3.957	-0.006	0.503
CO07M101	-0.051	-2.648	-0.006	-0.269
CO07M102	0.104	-1.854	0.013	-0.314

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO07M131	-0.100	2.686	-0.002	0.203
CO07MAS114	0.309	-0.138	0.025	-0.015
CO07MAS120	0.146	-2.656	0.014	-0.350
CO07MAS121	0.443	-5.802	0.023	-0.646
CO07MAS151	0.370	-3.944	0.031	-0.390
CO07MAS157	0.431	-2.936	0.045	-0.372
CO07RWA15	-0.261	3.424	-0.006	0.431
CO07RWA2	-0.038	1.752	0.001	0.157
CO07W151	-0.202	3.277	-0.017	0.281
CO07W245.F2	0.088	-4.233	0.000	-0.453
CO07W246	0.197	-1.025	0.025	-0.083
CO07W247	0.223	-8.119	-0.005	-0.917
CO07W247.F1	0.373	-7.546	0.014	-0.897
CO07W252	0.231	-1.691	0.023	-0.074
CO07W252.F3	0.329	-2.351	0.034	-0.073
CO07W252.F5	0.273	0.928	0.030	0.360
CO07W322	0.187	4.524	0.031	0.650
CO07W380	0.425	2.322	0.056	0.617
CO07W452	0.010	1.504	0.002	0.176
CO07W607	-0.026	-2.705	-0.002	-0.370
CO07W614	0.115	-2.666	0.019	-0.366
CO07W620	-0.366	2.850	-0.025	0.411
CO07W679	0.085	-0.011	-0.005	0.147
CO07W683	-0.103	2.839	0.001	0.405
CO07W718	0.322	-3.540	0.013	-0.233
CO07W722	0.514	-1.568	0.054	-0.008

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO07W722.F5	0.469	-3.267	0.043	-0.044
CO08127	0.046	-1.433	0.003	-0.095
CO08128	0.216	-3.820	0.011	-0.333
CO08136	0.413	-4.441	0.024	-0.471
CO08185	-0.197	1.138	-0.019	0.061
CO08203	-0.134	-1.766	-0.015	-0.278
CO08219	0.088	-0.567	0.018	-0.057
CO08224	-0.090	-2.947	-0.012	-0.395
CO08253	0.056	-0.439	-0.010	0.006
CO08259	-0.289	-1.827	-0.025	-0.067
CO08263	0.336	-4.108	0.008	-0.249
CO08302	0.073	5.807	0.033	0.975
CO08323	0.220	-2.262	0.010	-0.202
CO08327	-0.214	0.578	-0.021	0.188
CO08329	-0.035	0.204	-0.003	-0.018
CO08340	-0.148	3.271	0.005	0.557
CO08344	0.019	-3.095	-0.009	-0.359
CO08346	-0.012	2.077	0.000	0.296
CO08354	-0.036	-0.264	-0.011	0.092
CO08395	0.035	2.642	0.008	0.308
CO08412	-0.006	6.516	0.022	0.959
CO08454	0.121	1.069	0.009	0.123
CO08522	0.076	-3.423	-0.017	-0.481
CO08523	0.141	-0.910	-0.002	0.052
CO08530	-0.284	0.621	-0.030	0.113
CO08M011	-0.002	-4.136	-0.023	-0.614

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO08M045	-0.018	-2.063	-0.017	-0.201
CO08M045.F3	-0.129	1.380	-0.003	0.027
CO08RWA050	-0.174	-0.077	-0.017	-0.082
CO08RWA060	-0.193	-1.133	-0.023	-0.330
CO08W079	0.066	-1.555	0.001	-0.107
CO08W119	0.130	0.658	0.008	0.047
CO08W211	-0.067	4.558	0.011	0.796
CO08W218	0.337	-5.447	0.020	-0.451
CO08W218.F2	0.106	-2.138	0.018	0.076
CO08W232	-0.059	-2.176	-0.014	-0.413
CO08W328	-0.092	2.910	-0.012	0.363
CO08W328.F3	-0.102	4.353	-0.027	0.134
CO08W328.F5	-0.093	0.483	-0.029	-0.229
CO08W393	0.106	-7.830	-0.023	-1.316
CO08W412	-0.118	-0.253	-0.025	0.056
CO08W432	0.001	1.477	0.015	0.253
CO08W454	0.078	1.669	0.005	0.383
CO09005	0.019	4.285	0.008	0.480
CO09007	0.088	7.499	0.027	0.938
CO09040	-0.015	-1.979	-0.006	-0.292
CO09050	-0.034	-2.274	-0.006	-0.373
CO09059	-0.008	-2.006	-0.012	-0.260
CO09060	0.199	-0.652	0.023	-0.045
CO09075	0.118	0.683	0.007	0.085
CO09077	0.007	-2.061	-0.003	-0.211
CO09081	0.253	-1.518	0.021	-0.197

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO09112	0.129	1.496	0.001	0.335
CO09148	-0.745	4.604	-0.041	0.480
CO09149	-0.054	-1.948	-0.015	-0.181
CO09153	0.382	-7.582	0.020	-0.469
CO09157	-0.155	-2.353	-0.015	-0.271
CO09159	0.224	1.388	0.024	0.172
CO09183	-0.252	4.131	-0.007	0.472
CO09185	0.023	-1.539	0.013	-0.108
CO09193	0.028	0.558	0.000	0.183
CO09226	0.151	-0.294	0.005	-0.104
CO09227	0.010	-0.503	-0.012	-0.067
CO09231	0.291	-4.435	0.002	-0.469
CO09237	0.189	-5.141	0.011	-0.569
CO09241	0.187	-2.944	0.007	-0.386
CO09272	-0.223	-0.766	-0.038	-0.447
CO09279	0.254	-2.910	0.003	-0.381
CO09284	0.193	1.540	0.014	0.411
CO09287	0.094	-0.304	0.002	0.078
CO09292	0.088	4.410	0.018	0.677
CO09296	-0.173	6.817	-0.003	0.843
CO09300	0.155	-1.994	0.006	-0.053
CO09301	0.196	-2.843	0.013	-0.164
CO09306	0.200	-4.205	0.008	-0.236
CO09309	0.448	-4.003	0.019	-0.480
CO09316	-0.002	-1.535	-0.001	-0.192
CO09317	0.171	0.947	0.016	-0.009

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO09325	-0.153	7.119	0.007	0.932
CO09348	-0.125	-0.182	-0.004	-0.083
CO09369	-0.016	-2.946	-0.014	-0.245
CO09370	-0.424	2.636	-0.023	0.269
CO09371	-0.223	0.824	-0.015	0.025
CO09382	0.017	-2.654	-0.010	-0.325
CO09384	-0.439	9.901	-0.011	1.007
CO09385	-0.254	8.508	-0.001	1.070
CO09393	-0.414	1.114	-0.048	-0.139
CO09394	-0.070	3.422	0.003	0.360
CO09M0011	0.127	3.120	0.016	0.586
CO09M0022	-0.243	-1.763	-0.026	-0.220
CO09M0023	-0.237	2.416	-0.026	0.307
CO09M008	0.177	0.425	0.029	0.222
CO09W009	0.173	-3.342	0.014	-0.214
CO09W011	-0.201	3.545	-0.021	0.434
CO09W024	-0.273	1.295	-0.031	0.019
CO09W027	-0.101	9.287	0.013	0.998
CO09W028	-0.054	8.808	0.024	1.169
CO09W031	-0.204	5.299	-0.013	0.612
CO09W040	0.315	-8.784	0.011	-0.729
CO09W040.F1	-0.023	-2.794	-0.009	-0.254
CO09W052	-0.071	1.644	-0.007	0.139
CO09W061	-0.101	-0.849	-0.028	-0.192
CO09W091	-0.113	3.664	-0.009	0.314
CO09W106	0.027	4.559	0.012	0.484

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO09W107	-0.527	-0.641	-0.052	-0.408
CO09W109	0.182	-1.876	0.008	-0.296
CO09W110	0.031	-0.950	-0.012	-0.264
CO09W118	0.204	2.430	0.020	0.265
CO09W123	-0.054	2.385	-0.006	0.233
CO09W141	-0.388	4.638	-0.012	0.521
CO09W143	-0.026	-0.061	-0.008	-0.150
CO09W153	-0.117	0.543	-0.017	-0.064
CO09W154	-0.200	2.829	-0.003	0.211
CO09W165	-0.067	1.680	-0.004	0.155
CO09W169	-0.005	0.447	0.016	-0.008
CO09W172	0.134	-5.300	-0.007	-0.690
CO09W180	-0.020	1.955	0.007	0.371
CO09W181	-0.242	4.373	-0.007	0.537
CO09W190	-0.020	6.212	0.010	0.747
CO09W191	-0.083	2.938	0.000	0.291
CO09W202	-0.072	2.030	0.007	0.152
CO09W229	-0.488	-0.528	-0.045	-0.198
CO09W246	-0.157	1.517	-0.017	-0.006
CO09W248	-0.395	-0.615	-0.041	-0.223
CO09W284	-0.032	2.454	0.024	0.280
CO09W289	0.244	-2.052	0.023	-0.129
CO09W291	0.230	-3.585	0.022	-0.398
CO09W302	0.023	7.128	0.018	0.915
CO09W304	-0.250	1.177	-0.023	0.061
CO09W308	-0.220	4.405	-0.004	0.680

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO09W312	-0.268	5.502	-0.009	0.656
CO09W322	-0.211	7.107	0.010	0.854
CO09W323	-0.165	-1.542	-0.012	-0.344
CO09W330	-0.074	-0.035	-0.013	-0.033
CO09W332	-0.015	-1.327	-0.001	-0.178
CO09W333	-0.068	-1.335	-0.021	-0.249
CO09W334	-0.035	2.037	-0.017	0.151
CO09W342	-0.024	4.249	0.011	0.496
CO09W356	-0.196	2.255	-0.019	0.258
CO09W370	-0.184	3.571	-0.012	0.510
CO09W376	0.019	0.525	-0.016	0.042
CO09W379	0.094	2.830	-0.001	0.280
CO09W382	0.145	-0.386	0.009	0.032
CO09W389	-0.025	-2.775	-0.013	-0.323
CO09W391	0.206	0.801	0.019	0.187
CO09W399	-0.203	2.058	-0.004	0.304
CO09W412	-0.057	2.578	-0.013	0.231
CO09W418	-0.183	1.240	-0.028	-0.073
CO09W420	0.009	-1.217	-0.015	-0.279
CO09W428	-0.484	5.629	-0.032	0.486
CO09W434	-0.459	6.127	-0.030	0.633
CO09W435	-0.430	2.013	-0.022	0.238
CO09W448	-0.095	4.014	-0.001	0.561
CO09W451	0.071	-2.251	-0.001	-0.442
CO09W454	-0.140	-0.253	-0.014	-0.250
CO09W472	-0.159	-1.679	-0.033	-0.437

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO09W476	-0.024	4.912	-0.013	0.855
CO09W477	-0.304	-3.869	-0.043	-0.612
CO09W478	-0.189	10.381	0.021	1.141
CO09W479	0.131	1.291	0.011	0.277
CO09W480	-0.204	0.640	-0.015	-0.018
CO09W481	0.232	1.145	0.013	0.149
CO09W483	-0.010	-2.680	0.007	-0.068
CO09W487	-0.162	3.102	-0.009	0.324
CO09W504	0.012	3.386	0.014	0.510
CO09W542	-0.113	3.507	-0.009	0.297
CO09W546	-0.034	3.023	-0.009	0.305
CO09W549	0.038	-4.565	-0.007	-0.515
CO09W562	-0.438	8.146	-0.022	0.977
CO09W565	-0.478	6.949	-0.032	0.804
CO09W574	-0.234	1.919	-0.010	0.251
CO09W575	0.049	-3.087	-0.013	-0.247
CO09W584	-0.537	-2.101	-0.068	-0.590
CO09W590	-0.031	-1.193	-0.003	-0.225
CO09W591	0.305	-0.824	0.039	-0.173
CO09W595	-0.034	-2.083	0.005	-0.092
CO09W597	-0.171	0.480	-0.007	0.017
CO09W602	0.082	0.239	0.010	0.142
CO09W605	0.205	-2.258	0.005	-0.192
CO09W607	-0.108	4.710	-0.010	0.457
CO09W608	-0.320	12.079	0.007	1.522
CO09W617	-0.208	6.244	-0.004	0.684

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO10168	0.099	0.113	0.016	0.234
CO10W107	-0.044	-2.098	-0.009	-0.224
CO10W133	-0.205	-2.352	-0.013	-0.440
CO10W171	-0.134	0.037	-0.002	-0.136
CO10W172	-0.592	4.508	-0.057	-0.242
CO10W181	0.017	1.040	-0.012	-0.063
CO10W183	0.014	-0.895	-0.002	-0.198
CO10W302	0.082	3.209	0.032	0.681
CO10W314	0.098	1.684	0.023	0.478
CO10W444	-0.295	1.094	-0.023	-0.215
CO10W446	-0.363	0.379	-0.026	-0.048
CO10W465	-0.087	-3.042	-0.010	-0.314
CO11010	-0.348	8.970	-0.004	0.963
CO11274	-0.171	4.763	0.009	0.644
CO11296	-0.222	4.961	-0.006	0.660
CO11353	-0.086	-2.364	-0.014	-0.170
CO11442	-0.166	3.778	-0.002	0.442
CO11604	-0.393	5.310	-0.019	0.431
CO11D043	0.026	-4.516	-0.012	-0.578
CO11D053	-0.067	-3.556	-0.022	-0.556
CO11D069	-0.378	-4.347	-0.057	-0.923
CO11D100	-0.032	-0.511	0.004	0.231
CO11D1104	-0.147	-0.110	-0.015	0.118
CO11D1105	0.299	1.811	0.038	0.494
CO11D1106	0.111	6.075	0.035	0.991
CO11D1108	0.010	4.767	0.015	0.630

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D1111	0.280	-3.171	0.037	0.009
CO11D1118	0.059	5.332	0.022	0.930
CO11D1119	-0.208	6.084	-0.008	0.686
CO11D1121	0.043	1.540	0.006	0.252
CO11D1122	0.167	2.483	0.024	0.157
CO11D1125	0.122	2.396	0.025	0.386
CO11D1127	0.234	0.425	0.031	0.170
CO11D1132	0.098	-0.095	0.017	0.151
CO11D1133	-0.028	0.283	0.004	0.038
CO11D1134	-0.034	3.853	0.006	0.552
CO11D1140	0.387	-3.368	0.037	-0.178
CO11D1155	-0.137	-5.220	-0.025	-0.998
CO11D1156	-0.055	-1.692	-0.008	-0.510
CO11D1158	0.103	-4.693	-0.002	-0.287
CO11D1159	0.319	2.185	0.044	0.566
CO11D1162	0.117	-4.145	0.004	-0.320
CO11D1165	0.179	-1.227	0.019	-0.048
CO11D1168	0.058	1.724	0.004	0.118
CO11D1174	0.299	-6.748	0.013	-0.695
CO11D1182	-0.193	1.999	-0.022	0.056
CO11D1184	-0.092	0.682	-0.009	0.173
CO11D1186	-0.067	2.000	0.002	0.395
CO11D1190	0.061	-6.534	-0.005	-0.886
CO11D1193	0.054	0.076	0.006	0.061
CO11D1197	0.137	-1.386	0.015	0.093
CO11D1198	-0.168	-2.148	-0.021	-0.297

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D1206	-0.088	2.512	-0.003	0.348
CO11D1207	0.092	1.260	0.018	0.301
CO11D1208	0.077	1.907	0.018	0.265
CO11D1210	-0.073	2.550	-0.010	0.329
CO11D1213	0.112	-5.630	-0.008	-0.644
CO11D1216	0.002	2.626	0.011	0.560
CO11D1219	-0.012	1.383	-0.002	0.271
CO11D1221	0.449	1.105	0.058	0.380
CO11D1223	-0.018	2.898	0.000	0.358
CO11D1225	0.066	4.022	0.017	0.902
CO11D1229	-0.051	6.593	0.013	0.955
CO11D1231	-0.392	1.837	-0.044	0.018
CO11D1232	0.183	3.528	0.022	0.511
CO11D1234	0.187	3.626	0.037	0.548
CO11D1235	0.091	0.339	0.013	0.306
CO11D1236	0.115	-8.424	-0.005	-0.908
CO11D1240	0.351	-6.859	0.024	-0.472
CO11D1242	0.098	0.166	0.020	-0.050
CO11D1243	0.113	3.996	0.029	0.521
CO11D1246	-0.211	-2.626	-0.033	-0.235
CO11D1247	-0.092	-4.266	-0.023	-0.567
CO11D1248	-0.539	-0.125	-0.058	-0.359
CO11D125	0.054	0.496	0.027	0.416
CO11D1252	0.083	-5.649	-0.007	-0.744
CO11D1261	-0.141	-4.838	-0.032	-0.569
CO11D1267	0.146	-3.190	0.014	-0.355

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D1268	-0.189	2.393	-0.022	0.112
CO11D1270	0.127	-3.103	0.008	-0.368
CO11D1271	-0.406	-3.379	-0.057	-0.699
CO11D1272	-0.079	-0.673	-0.011	0.020
CO11D1282	0.232	-5.146	0.010	-0.613
CO11D1287	-0.020	1.739	-0.004	0.285
CO11D1289W	0.016	-1.318	-0.005	-0.268
CO11D1290W	-0.198	-4.651	-0.027	-0.425
CO11D1292	-0.169	2.822	-0.018	0.337
CO11D1293	-0.041	-2.807	-0.001	-0.336
CO11D1294	0.358	-0.346	0.041	0.148
CO11D1296W	0.434	-1.575	0.053	-0.005
CO11D1298	0.251	-8.318	0.004	-0.934
CO11D1300W	0.469	-2.443	0.048	-0.235
CO11D1302W	-0.149	-0.153	-0.022	-0.068
CO11D1305	0.084	-2.633	-0.004	-0.216
CO11D1306W	0.432	-2.200	0.043	-0.166
CO11D1309W	0.077	-4.841	-0.005	-0.847
CO11D1311	-0.063	-4.380	-0.018	-0.720
CO11D1312	0.326	-4.240	0.025	-0.444
CO11D1315W	-0.067	-0.171	-0.004	0.187
CO11D1316W	0.408	-5.726	0.026	-0.580
CO11D1317	-0.038	-0.091	-0.007	-0.019
CO11D1322	-0.097	0.699	-0.018	0.211
CO11D1325	-0.011	-2.190	-0.011	-0.299
CO11D1332	-0.056	-3.923	-0.019	-0.502

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D1333	0.218	-0.167	0.028	0.275
CO11D1334	0.025	1.276	0.007	0.451
CO11D1335W	-0.001	-0.257	-0.008	-0.173
CO11D134	-0.267	0.669	-0.015	0.077
CO11D1341	0.135	2.090	0.023	0.443
CO11D1343	-0.152	-2.188	-0.030	-0.348
CO11D1345	-0.021	-0.573	-0.010	0.007
CO11D1351	0.009	-3.443	-0.013	-0.361
CO11D1352	-0.116	-0.169	-0.014	0.064
CO11D1353	0.208	-3.733	0.015	-0.334
CO11D1355	-0.094	-3.748	-0.017	-0.447
CO11D1356	-0.052	3.319	0.003	0.526
CO11D1360	-0.084	-1.272	-0.019	-0.212
CO11D1361	0.020	-0.425	-0.002	0.081
CO11D1364	0.059	-4.407	-0.015	-0.490
CO11D1367	-0.050	-3.673	-0.013	-0.398
CO11D1371	-0.194	0.414	-0.024	-0.004
CO11D1374	0.079	-8.359	-0.017	-1.142
CO11D1376	0.075	0.691	0.008	0.314
CO11D1377	0.048	-1.858	0.004	-0.176
CO11D1380	-0.213	1.268	-0.022	0.194
CO11D1382	0.007	-1.072	-0.013	-0.168
CO11D1383	-0.360	-5.322	-0.058	-0.739
CO11D1385	-0.106	-1.004	-0.019	-0.347
CO11D1390	0.169	-8.778	0.002	-1.062
CO11D1392	0.024	-6.166	-0.006	-0.870

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D1393	-0.186	-3.613	-0.022	-0.499
CO11D1397	0.199	-7.034	-0.001	-0.688
CO11D1401	0.119	-6.139	0.004	-0.924
CO11D1406	0.066	-4.877	0.003	-0.189
CO11D1407	0.132	-0.442	0.016	0.136
CO11D1409	0.120	-2.014	0.006	-0.275
CO11D1412	-0.021	-0.846	-0.006	-0.061
CO11D1414	-0.175	-1.254	-0.027	-0.153
CO11D1415W	-0.085	-0.970	-0.010	-0.259
CO11D1416	0.045	4.510	0.014	0.801
CO11D1418	-0.083	2.748	-0.006	0.250
CO11D1421W	0.116	-0.856	0.013	-0.045
CO11D1422	0.071	-2.022	0.003	-0.119
CO11D1424	-0.110	-1.804	-0.032	-0.174
CO11D1428W	-0.012	2.851	0.010	0.589
CO11D1431	0.017	6.055	0.011	0.762
CO11D1528	0.096	-4.857	0.004	-0.546
CO11D1534	-0.190	-5.716	-0.037	-0.890
CO11D1535	0.131	0.575	0.031	0.351
CO11D1536	0.065	-4.234	-0.003	-0.350
CO11D1539	0.212	-4.375	0.016	-0.329
CO11D1542	0.226	-3.525	0.012	-0.297
CO11D1543	-0.293	1.760	-0.034	0.027
CO11D1545	-0.165	-2.780	-0.032	-0.608
CO11D1546	0.131	-5.120	0.004	-0.592
CO11D1547	-0.051	-5.573	-0.017	-0.505

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D1551	-0.202	4.893	-0.012	0.820
CO11D1553W	-0.144	5.987	0.007	1.020
CO11D1555W	0.048	2.872	0.012	0.545
CO11D1557W	-0.137	5.675	0.001	0.820
CO11D1559	-0.197	8.738	-0.007	1.249
CO11D1561	-0.089	0.928	-0.011	0.021
CO11D1564	-0.345	2.797	-0.053	0.182
CO11D1570	-0.251	2.687	-0.024	0.482
CO11D1577	-0.178	1.356	-0.010	0.210
CO11D1579	-0.164	4.007	-0.008	0.393
CO11D1581W	-0.200	3.886	-0.014	0.427
CO11D1582W	-0.015	16.622	0.036	2.455
CO11D1584	-0.477	12.015	-0.034	1.446
CO11D1586W	-0.059	6.557	0.008	0.941
CO11D1588	-0.078	4.069	-0.001	0.637
CO11D1592	-0.057	4.866	0.007	0.652
CO11D1594	-0.292	4.059	-0.031	0.365
CO11D1600W	0.051	0.793	0.010	0.200
CO11D1603W	-0.180	2.273	-0.021	0.135
CO11D1606	-0.346	4.957	-0.036	0.637
CO11D1609	0.087	-0.529	0.020	-0.031
CO11D1613W	0.107	5.985	0.033	1.119
CO11D1615	-0.316	9.575	-0.019	1.212
CO11D1618	-0.110	3.621	-0.003	0.683
CO11D1620	0.123	1.591	0.020	0.293
CO11D1623	0.055	-0.027	0.006	0.057

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D1624	0.163	-0.434	0.024	0.350
CO11D1627	0.159	1.434	0.032	0.624
CO11D1629	0.204	0.461	0.038	0.339
CO11D1631	0.290	4.367	0.059	0.836
CO11D1633	-0.027	2.788	0.005	0.527
CO11D1636	-0.131	2.279	-0.011	0.407
CO11D1642	0.011	-0.078	0.010	0.456
CO11D1644	-0.064	6.011	0.016	0.868
CO11D1645	0.006	3.963	0.015	0.611
CO11D1647	0.114	2.726	0.024	0.483
CO11D1654	0.158	-2.198	0.009	-0.345
CO11D1656	-0.315	4.604	-0.024	0.611
CO11D1657	-0.074	1.977	0.001	0.370
CO11D1659	-0.098	1.911	-0.006	0.489
CO11D1665	-0.163	5.326	-0.002	0.831
CO11D1666	-0.137	0.592	-0.004	0.414
CO11D1672	0.089	-1.198	0.003	-0.261
CO11D1677W	0.006	-6.494	-0.019	-0.889
CO11D1680	-0.014	0.805	0.000	0.023
CO11D1683	-0.029	-0.377	-0.003	0.030
CO11D1685W	0.029	-2.823	0.003	-0.170
CO11D1686	0.495	-7.214	0.035	-0.692
CO11D1689W	-0.045	-3.731	-0.025	-0.487
CO11D1692	0.248	-2.874	0.016	-0.165
CO11D1694	0.091	0.660	0.010	0.073
CO11D1697	-0.238	-4.398	-0.041	-0.752

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D1698	-0.132	-2.182	-0.023	-0.413
CO11D1704	0.042	-6.179	-0.017	-0.788
CO11D1706	0.201	-2.527	0.018	-0.377
CO11D1710	0.191	-0.988	0.026	-0.153
CO11D1712	-0.296	-3.144	-0.041	-0.434
CO11D1715	0.355	-5.594	0.020	-0.510
CO11D1719	0.097	-4.816	0.006	-0.460
CO11D1723	0.108	-2.856	0.005	-0.290
CO11D1725	0.031	-0.661	0.007	0.105
CO11D1728	0.013	-1.008	0.003	-0.037
CO11D1731	-0.142	0.009	-0.006	0.098
CO11D1733	-0.107	3.661	-0.007	0.739
CO11D1734	0.322	-5.012	0.024	-0.514
CO11D1739	0.210	-7.693	0.001	-0.893
CO11D1740	-0.140	-5.520	-0.030	-0.716
CO11D1742	0.324	-4.142	0.023	-0.571
CO11D1743W	0.026	-1.115	0.002	-0.070
CO11D1744	0.066	-5.098	-0.003	-0.415
CO11D1746	-0.297	-1.978	-0.041	-0.283
CO11D1748	0.255	-9.488	-0.003	-1.057
CO11D1749	-0.172	-4.844	-0.038	-0.627
CO11D1751	0.168	1.191	0.022	0.314
CO11D1752	-0.272	1.176	-0.026	-0.166
CO11D1754	-0.108	5.298	-0.003	0.883
CO11D1758W	-0.318	3.454	-0.035	0.298
CO11D1759	-0.175	0.634	-0.019	0.042

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D1767	0.316	-4.633	0.020	-0.497
CO11D1769	0.160	2.073	0.019	0.305
CO11D1772	-0.011	4.821	0.012	0.446
CO11D1783	0.040	2.583	0.009	0.460
CO11D1785	0.078	-1.832	0.011	-0.035
CO11D1787	-0.152	-1.654	-0.024	-0.296
CO11D1788	-0.210	-7.959	-0.038	-1.136
CO11D1789	-0.185	3.478	-0.019	0.406
CO11D1790	0.064	-0.597	0.004	0.050
CO11D1792	0.084	0.643	0.010	0.168
CO11D1794	0.145	6.641	0.050	1.160
CO11D1796	-0.204	1.166	-0.023	0.335
CO11D1798	0.035	-2.880	-0.015	-0.554
CO11D1799	0.072	-6.350	-0.014	-0.940
CO11D1800W	0.113	-2.585	0.010	-0.370
CO11D1802	-0.201	-5.466	-0.033	-0.729
CO11D1804	-0.099	-0.828	-0.018	-0.292
CO11D1805	-0.271	3.484	-0.023	0.486
CO11D1808W	0.336	-0.851	0.031	0.057
CO11D1809	0.216	2.172	0.028	0.294
CO11D1810	0.106	-0.042	-0.002	0.192
CO11D188	-0.179	5.128	0.010	0.714
CO11D243	0.341	-6.406	0.020	-0.481
CO11D323	0.315	-10.286	0.000	-1.047
CO11D346	0.350	-6.193	0.014	-0.478
CO11D378	0.202	-2.669	0.010	-0.228

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
CO11D421	0.332	-3.203	0.023	-0.141
CO11D424	-0.303	0.294	-0.021	-0.242
CO11D428	-0.236	-0.272	-0.018	-0.149
CO11D444	-0.041	2.324	0.005	0.211
CO11D461	0.014	-1.793	-0.003	-0.075
CO11M045	0.020	-4.762	-0.015	-0.625
CO11M106	-0.058	-0.672	-0.007	0.087
CO11W381	0.006	-2.811	-0.009	-0.389
CO11W393	-0.093	3.180	0.009	0.466
Cowboy	0.511	-3.035	0.027	-0.232
Danby	-0.107	1.988	-0.005	0.257
Denali	0.046	-2.829	-0.002	-0.256
Hatcher	0.201	-0.994	0.014	-0.004
Jagalene	-0.265	3.556	-0.021	0.518
Keota	-0.123	6.750	0.023	0.849
Langin	0.193	-1.702	0.021	0.029
LCSMint	0.012	0.330	0.011	0.158
Longhorn	-0.338	4.150	-0.015	0.207
NuDakota	0.067	2.873	0.020	0.494
NuFrontier	-0.150	1.046	-0.012	-0.009
Ripper	0.219	5.120	0.042	0.774
SettlerCL	0.001	0.393	0.022	0.113
Snowmass	-0.085	-0.689	-0.017	-0.175
Sunshine	0.140	3.905	0.026	0.560
TAM111	0.220	1.205	0.019	0.134
TAM112	0.107	3.717	0.018	0.403

Trait	GY BLUP	GPRO BLUP	GPY BLUP	GPD BLUP
Genotype	Mg ha^{-1}	g kg^{-1}	Mg ha^{-1}	g kg^{-1}
ThunderCL	-0.109	-1.073	-0.012	-0.338
WB.Grainfield	-0.219	6.811	0.005	0.707
Winterhawk	-0.112	1.478	-0.001	0.029

Supplementary table 4. Realized relationship matrix (G) (open "SupplementaryTable3_Latshaw").

Supplementary table 5. Pedigrees included in the multi-environment trials 2012-2014.

Pedigree	Count of ID
03A-KCB-8/Hatcher//Hatcher	3
03WSR-197/Ankor//Danby	2
03WSR-203/Yuma//Danby	1
12SAWYT-22/Hatcher//Hatcher	1
2002 Altus-034/Bond CL//TAM 111	2
2414-11/5*CO00554	2
474S10-1/X87807-26//HBK0736-3	1
89-27/CO970547//2*CO970547-7/CO970547-7	1
89-27/CO970547//3*CO970547-7	1
94M370/6*Yuma	2
96X0799-11W/Avalanche//CO970547-7	1
96X0799-11W/Stanton//CO970547-7	1
96X0856-01W/Trego//CO970547-7	1
98HW423(JGR/93HW242)/96HW94	1
98HW519(93HW91/93HW255)/96HW94	1
98HW521(93HW91/93HW255)/98HW165(ARL/WGRC15)	1
A97201S-B-34/Avalanche//CO99W329	2
Abilene/Jagger	1
Above//OK95616-14C/G980103W	1
Above/Stanton	1
Akron/Halt//4*Akron	1
Ankor/96X0856-01W//KS01HW168-4	1
Antero/CO050233-2	10
Antero/HV9W07-482W//CO05W111	1
Avalanche/CO980630	1
Avalanche/CO99148	1
Avalanche/CO99W075	1
Avalanche/KS920946-B-15-1	1
Avalanche/KS970392-1-1-2//CO01W172	4
Avalanche/NuFrontier	2
Avalanche/NW97S343	1
Avalanche/TAM 112//OK Rising	1
Avalanche/W97-189	1
B1551-WH/KS94U326	1
Baker's White/96X0799-15W//CO970547-7	2
Bill Brown/Byrd	3
Bill Brown/Garrison	1

Pedigree	Count of ID
Bill Brown/KS05HW14-3	2
Bill Brown/NuDakota//Bill Brown	5
Bill Brown/OK Rising	1
Bill Brown/Thunder CL	1
Brawl CL Plus/CO05W111	20
Burchett/CO960293-2//Stanton	1
Byrd/Antero	13
Byrd/CO07W247//CO050337-2	1
CIMMYT01-59/CO970547//3*CO970547-7	1
CO00580/TAM 111	2
CO00739/KS01HW152-6//OK Rising	1
CO00739/RonL//OK02518W	1
CO01385/Danby	2
CO01385/G001172W	4
CO01385/KS01HW163-4	1
CO01W171/CO02W040	1
CO01W171/CO02W283	1
CO01W172/KS02HW90-5	1
CO01W173/KS01HW152-6	1
CO01W173-A3/CO02W021	1
CO01W189-A1/KS03HW97-1	1
CO01W191/CO02W010	2
CO01W191/CO02W180	1
CO01W191/HV9W02-110W	2
CO01W191/KS03HW38-2	1
CO01W191/TX00V1117	2
CO02W040/KS01HW152-6	1
CO02W183/CO01W191	1
CO02W237/CO01W173-A3	1
CO03W054/CO940610	4
CO03W054/Hatcher//CO03W054	1
CO03W054/OK05723W	2
CO050173/Antero//Byrd	3
CO050173/Cowboy	3
CO050233-2/Byrd	13
CO050233-2/CO050337-2	6
CO050233-2/Cowboy	14
CO050270/Byrd	6
CO050270/Hatcher	1

Pedigree	Count of ID
CO050337-2/Antero	14
CO050337-2/Byrd	26
CO05W111/Antero	1
CO06072/4*Byrd	4
CO06072/4*CO050337-2	1
CO07MAS114/CO050173	5
CO07MAS114/CO050233-2	7
CO07MAS114/CO05W111	1
CO07MAS114/Cowboy	27
CO07MAS114/Denali	32
CO940606/TAM107R-2	1
CO940610/CO960293//CO99W189	1
CO950043//CO940610/KS99HW24	1
CO950043/Above//CO970547	2
CO950043/CO99141	1
CO950043/TX90A9528	1
CO950635/CO99W1126	5
CO960691/CO970655	1
CO970498/CO950043//CO970547	4
CO970498/CO970940//CO980376	3
CO970498/KS98HW220-5	3
CO970498/Stanton	1
CO970547/Prowers 99	2
CO970547-7/KS01HW152-6	2
CO980352/CO970235	1
CO980376/CO99W254	2
CO980630/CO99W076	4
CO980829/CO950043	1
CO980829/TAM 111	8
CO980862/Lakin	1
CO99314/CO00580	1
CO99314/W96x1080-21//Jagalene	1
CO99W076/TAM 111	1
CO99W1126/CO980352	2
CO99W1126/KS96HW10-3	6
CO99W1126/KS99HW24	1
CO99W165/CO99526	1
CO99W165/G97252	1
CO99W165/OK98G502W	1

Pedigree	Count of ID
CO99W183/G980091	1
CO99W189/Lakin	2
CO99W254/CO99148	3
CO99W254/KS00HW114	4
Cowboy/Antero	14
Cowboy/Antero//Byrd	4
Custer/Jagger	1
Danby(TREGO/JGR 8W)/BC97ROM-41W	2
Danby/CO02W010	2
Danby/CO02W214	2
Danby/CO970547-7	10
Danby/Everest//CO03W054	1
Danby/TX00V1117	6
Denali/Antero	16
Denali/Antero//Byrd	2
Denali/Antero//Snowmass	6
Denali/Byrd	17
Denali/CO050233-2	11
Denali/CO07W322//Byrd	1
Denali/HV9W07-482W//Antero	5
Denali/KS06HW46-3//Byrd	3
G001172W/CO01212	1
G970209W/CO970547	5
G970209W/CO980829	1
G97343/CO99W165	1
G982231/G982159//KS920709W	1
Hallam/CO99502	1
Hatcher/HV9W02-267W//KS02HW35-5	1
Hatcher/KS01HW152-6//OK Rising	1
Hatcher/KS05HW120//CO03W054	1
Hatcher/NuGrain//KS02HW35-5	1
Hatcher/NW97S295	11
Hatcher/OK02518W//Danby	3
Hatcher/OK03716W//KS02HW35-5	2
HV9W02-267W/Danby	6
HV9W03-280W-2/CO04W323//CO06062	2
Jagalene/KS01HW168-1	1
JAGALENE/KS03HW122(LAKIN/TGO//96HW71)//03-6149(TREGO/CO960293)	1

Pedigree	Count of ID
JAGALENE/KS03HW122(LAKIN/TGO//96HW71)//KS01HW152-1-2(TREGO/BTY SIB)	1
JAGALENE/KS03HW149-1(TREGO/CO960293)//03-6149(TREGO/CO960293)	1
Jagalene/TAM 112	4
Jagger/Romanian	1
KS00HW115/CO99W078	1
KS00HW151-4/CO00580	5
KS00HW177/CO980719	1
KS00HW183/CO99W188	1
KS01-5539/CO99W165	1
KS01HW152-1/TAM 111	6
KS01HW152-6/CO99141	5
KS01HW152-6/CO99314	5
KS01HW152-6/G001011W	2
KS01HW152-6/HV9W02-267W	4
KS01HW152-6/NuFrontier	1
KS01HW168-1/TAM 111//CO02W214	2
KS01HW168-1/TAM 111//Platte	4
KS01HW168-4/TX00V1117//CO02W214	3
KS02HW112/TAM 111	1
KS02HW30/TAM 111	1
KS02HW30/W98-363$//CO99314	1
KS02HW35-5/CO02W237	3
KS02HW35-5/OK00611W	3
KS02HW35-5/OK02518W	2
KS02HW89/CO00739//CO01W172	1
KS02HW89/TX00V1117//CO01W172	2
KS02HW89-1(TREGO*2/JGR8W)/BC9503565-6	1
KS02HW89-1(TREGO*2/JGR8W)/BC97ROM-41W	2
KS02HW90/CO970547-7//CO02W040	3
KS02HW90/TAM 112//CO01W172	3
KS02HW90/TX00V1117//CO02W040	1
KS02HW91/CO00554//Danby	4
KS02HW91/Endurance//CO02W214	1
KS02HW91/Ripper//Platte	2
KS02HW91/TAM 112//CO02W214	1
KS02HW91-6/CO01W172	2
KS02HW91-6/G001011W	7
KS03HW38-2/CO01W189-A1	4

Pedigree	Count of ID
KS04HW47-3/CO03W054	4
KS04HW47-3-4/NuDakota//Hatcher	1
KS05HW120/HV9W02-243W//CO03W054	2
KS05HW121-1/Bill Brown//KS05HW14-3	1
KS05HW121-2/Hatcher	1
KS05HW15-2/CO03W054	1
KS87H325/Rio Blanco	1
KS96HW94//Trego/CO960293	1
KS96HW94/CO980352	1
KS98HW151-5/CO99W075	2
KS98HW452/CO960293//Lakin	1
KS99HW24/NuHorizon	2
KS99HW36/Avalanche	2
Lakin/CO950635	1
Lakin/CO980352	2
N95L164/3/MILLENNIUM SIB//TXGH125888-120*4/FS2	1
NS2630-1/Thunderbird	1
NuDakota/KS05HW122-5//Bill Brown	1
NuFrontier/CO01385	2
NuFrontier/Danby	1
NuHills/Baker's White//CO99141-A5	1
NuHills/CO980630	1
NW97S343/Akron	3
OK Rising/Bill Brown//Hatcher	2
OK Rising/CO04W323//CO06072	1
OK Rising/Danby	1
OK02518W/Aspen	2
OK02518W/CO01W189-A1	1
OK02518W/KS05HW42	1
OK03716W/HV9W97-2112W-1	1
Overley/CO980829	4
Pioneer bulk selection (HBK0927)	1
Ripper/CO050173	3
Ripper/SRS2-31//3*Ripper	7
Ripper/Thunder CL//CO06062	1
RT01-10/Ankor//Ankor	1
Snowmass/Antero	9
Snowmass/Byrd	1
Snowmass/CO07MAS114//Snowmass	1

Pedigree	Count of ID
Snowmass/CO07W322//Snowmass	2
Snowmass/CO08RWA060//Antero	1
Snowmass/CO08RWA060//CO05W111	1
Snowmass/CO08W454	9
SRS2-31/4*Hatcher	4
Stanton/CO950043	1
TAM 110*4/FS2	1
TAM 111/CO99526	2
TAM 111/KS02HW89//CO01W171	3
TAM 111/Trego//CO99W329	5
TAM 112/Byrd	2
TAM 112/CO970547-7	3
TAM 112/OK02518W//HV9W02-243W	1
TAM-107//TX78V3630/CTK78/3/TX87V1233	1
Teal 11A/2*BondCL//Ripper	1
Teal 11A/2*Protection//Hatcher	2
Teal 11A/3*Bond CL	3
Teal 11A/4*Bond CL	3
Teal 11A/Above//Bond CL	2
Teal 11A/Above//CO99314	7
Teal 11A/Above//KS01HW163-4	2
Teal 11A/Bond CL//CO980684-1	1
Teal 11A/CO991350//Jagalene	1
Teal 11A/CO991350//Stanton	1
Teal 11A/KS01-5539//CO99W183	1
Teal 11A/KS01-5539//TAM 111	1
Teal 11A/Protection//CO99141	1
Trego/CO99148	1
TREGO/JGR 8W	1
TX00V1117/CO01W189	3
TX97V2838/NuHills//CO970547-7	1
TX97V2838/NuHills//KS01HW168-4	3
TX97V2839/CO99W182	2
TX98VR8426/Lakin	1
U1254-7-9-2-1/TXGH10440	1
W96x1080-21/CO99W183//KS01HW168-4	1
W98-363$/Ankor//CO99W254	3
W98-363$/CO99W183	1
W98-363$/CO99W277	1

Pedigree	Count of ID
W98-363$/TAM 111//CO970547-7	6
W98-363$/Trego//KS01HW168-4	1
WB411W/TAM 111//CO970547-7	3
WB411W/TAM 111//KS01HW168-1	9
Winterhawk/CO050233-2	2
Winterhawk/Danby//CO03W054	2
Yuma/Cutter//CO980376	1
Yuma/Hatcher	1
Yuma/PI 372129//TAM 200/3/4*Yuma/4/KS91H184/Vista	1
Yumar//TXGH12588-120*4/FS2	1
Yumar/Arlin	1