Evaluating genomewide marker-based breeding methods in traditional and wild relative-derived barley populations

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Tyler Tiede

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Kevin P. Smith

April 2016

Acknowledgements

I thank my advisor, Dr. Kevin P. Smith, for his mentorship throughout my Ph.D. studies. Kevin provided guidance and clarity to my research, while challenging me to expand my scope and be mindful of the next steps. He always supported me as new ideas and directions arose, and provided the expertise, constructive feedback, and resources required to pursue the question.

I also thank my committee members, Dr. Rex Bernardo, Dr. Rebecca W. Doerge, Dr. Gary J. Muehlbauer, and Dr. Brian J. Steffenson. They always made themselves available to provide guidance and lend their expertise to my research. A special thank you to Dr. Doerge for sticking with me beyond my M.S. program – your belief in me has been a constant encouragement.

I am grateful to Karen Beaubien, Ed Schiefelbein, and Guillermo Velasquez for training me in the technical, everyday aspects of running a breeding program. Even more, without them plants would not have been grown, genotyped, nor crossed – thank you for your countless hours of dedicated help. Additionally, I thank Dr. Jochum Wiersma and Carl Duley for their assistance with field trials, and also for being great mentors in the area of agronomy and extension service. I would also like to acknowledge Dr. Shiaoman Chao and Dr. Gina Brown-Guedira for the genotyping services they provided, and for educating me on the finer points of sequencing-based genotyping.

To the lab mates, both past and present – it has been a lot of fun working with you and learning from you: Celeste Falcon, Jeff Neyhart, Alex Ollhoff, Dr. Ana Poets, John Hill Price, Dr. Ahmad Sallam, Dr. Vikas Vikram, and Lu Yin. A special thank you to Dr.

Mohsen Mohammadi for many great conversations and a very fruitful collaboration. I would also like to acknowledge Dr. Liana Nice, not only for her hard work in developing the AB-NAM population, but also for her many excellent suggestions and insights into working with wild barley relatives.

My research assistantship was funded by the USDA-ARS through the U.S. Wheat and Barley Scab Initiative in addition to a year of funding from a UMII MnDRIVE Graduate Research Assistantship. Research activities were funded through the MN Agricultural Experimental Station through the Varietal Development Fund. I also thank Dr. Donald C. Rasmusson for supporting me during my Ph.D. work through the Donald and Carol Rasmusson Fellowship.

The time and energy that graduate school requires often comes at the expense of time spent with family and friends. For their patience and understanding, I thank my wife, Lizzie, my parents, and other close family members and friends.

Dedication

This dissertation is dedicated to my wife, Lizzie, who has been a constant source of love and encouragement, and the perfect example of patience.

Abstract

The genetic improvement of plants for human use is the primary goal of plant breeding. Through repeated processes of population development and selection, breeders have produced highly productive plants that are adapted to a vast array of environments and cultivation practices. The challenges that drove plant breeders in the past, such as increasing production, novel or increasingly prevalent abiotic and biotic stresses, and evolving end-user demands persist today and are compounded with unprecedented population growth. Genomic selection (GS), a genomewide marker-based selection method, has been shown to be an efficient and effective breeding tool. Its general applicability to plant breeding and principles guiding its use have been established by simulation and empirical cross-validation studies. More recently, studies have demonstrated genetic gains over multiple cycles of selection in a variety of crop species.

In the first chapter we provide additional evidence for the effectiveness of GS in an actual breeding program by demonstrating significant gains of 164.74 kg ha^{-1} and -1.41 ppm for grain yield and DON, respectively, two unfavorably correlated quantitative traits, across three cycles of selection in a spring six-row barley breeding population. With its general effectiveness established, the next step is to increase the accuracy of GS and thereby increase genetic gains. For this, we first showed that updating the training population (TP) with phenotyped lines from recent breeding cycles, specifically selected lines, had an overall positive effect on prediction accuracy. Additionally, we investigated four recently-proposed algorithms that seek to optimize the composition of a TP. Overall the optimization algorithms improved prediction accuracy when compared to a randomly

selected TP subset of the same size, but which algorithm performed best was dependent on the trait being predicted and other factors discussed within. This retrospective investigation highlighted the importance of maintaining and optimizing the TP when using GS in real breeding situations to maximize prediction accuracy, thereby maximizing gain from selection and resource utilization.

Furthermore, genetic gains depend on genetic variation. Exotic germplasm can be exploited to introduce genetic variability into elite breeding populations to drive genetic gains and address new or changing breeding targets. Wide crosses between elite and exotic germplasm have been widely used to identify large-effect QTL and breed for improved disease or insect resistance. The utility of exotic germplasm to improve quantitative traits, which includes many important agronomic traits, has not been tested as widely. In the second chapter we select parents from an advanced backcross population constructed from 25 wild barley (*Hordeum vulgare* L. spp *spontaneum*) accessions crossed to the common high-yielding malting barley cultivar, Rasmusson. We extended the genomic selection framework to identify parent combinations that, with ideal recombination, should produce progeny with large numbers of exotic introgressions with favorable effects. We compared the marker-based crossing strategy to the traditional methods of crossing the topmost performing parents or the most genetically diverse parents. After one round of marker-based progeny selection from these crosses we identified breeding lines, harboring exotic introgressions, which consistently yielded higher than Rasmusson across five trial locations. While none of these lines were statistically better than Rasmusson, there is compelling evidence that the introgression of

wild alleles contributed to increased yield. The three parent selection strategies were not significantly different for their ability to identify superior progeny.

Table of Contents

List of Tables

Chapter 1

Chapter 2

Lists of Figures

Chapter 1

Chapter 2

Chapter 1: Evaluation and retrospective optimization of genomic selection for yield and disease resistance in spring barley

Introduction

Genomic selection (GS), also called genomewide selection, is becoming a routinely used tool in public and private plant breeding programs. Since its original description (Meuwissen et al., 2001) there has been active research to investigate this methodology and evaluate its effectiveness. Initial support largely came from simulation-based studies (Bernardo and Yu, 2007; Mayor and Bernardo, 2009; Jannink, 2010; reviewed in Jannink et al., 2010 and Lorenz et al., 2011). In 2009, Heffner and Sorrels reviewed the state of GS in plant breeding and identified the need for empirical studies to verify the theory that had been developed to that point (Heffner et al., 2009). Since then, numerous empirical cross-validation (CV) studies have investigated "best prediction practices," such as the impact that marker number, training population (TP) size, statistical models, and other factors have on GS prediction accuracy (Lorenzana and Bernardo, 2009; Crossa et al., 2010; Asoro et al., 2011; Heffner et al., 2011; Heslot et al., 2012; Lorenz et al., 2012; Combs and Bernardo, 2013a; Massman et al., 2013a; Riedelsheimer et al., 2013; Beaulieu et al., 2014; Guo et al., 2014; Grenier et al., 2015; Isidro et al., 2015; Sallam et al., 2015). Prediction accuracy is the Pearson correlation (r) between the predicted genotypic value ($\widehat{G}$) with the respective lines' true genotypic value (G). Outside of simulation true genotypic values are unknown and must be estimated from phenotypic data; therefore prediction accuracy will be abbreviated $r_{\hat{G},G}$ hereafter.

Empirical selection studies collectively demonstrate that GS results in equivalent or greater genetic gains over cycles of selection compared with traditional breeding strategies such as phenotypic selection and pedigree-based best linear unbiased prediction (pedigree BLUP; Combs and Bernardo, 2013b; Asoro et al., 2013; Rutkoski et al., 2015; Beyene et al., 2015) and marker assisted selection (MAS; Massman et al., 2013b; Asoro et al., 2013; Rutkoski et al., 2015).

The breeder's equation $R = ir\sigma_A$ (Falconer and McKay, 1996; Heffner et al., 2009) indicates that the genetic response (R) increases as the selection intensity (i), the square root of the additive genetic variance (σ_A), and accuracy of selection (r) increase. Genomic selection also enables shorter breeding cycles. Genomic selection increases the rate of gain largely through decreasing the length of the breeding cycle (Bernardo and Yu, 2007; Heffner et al., 2009). Rutkoski et al. (2015) were able to perform two cycles of GS in the time it took to perform a single cycle of phenotypic selection and observed that the response to selection on a per unit time basis was similar for GS and phenotypic selection (Rutkoski et al., 2015). Other selection studies showed that GS was equal to or better than the traditional, non-genotype based selection strategies on a per-cycle basis (Combs and Bernardo, 2013a; Asoro et al., 2013; Beyene et al., 2015). Furthermore, Sallam (2014) showed that in five unique populations, for equivalent values of i and σ_A, a single selection cycle of GS and phenotypic selection performed equivalently and resulted in significant genetic gains for disease resistance.

In theory, genetic variance should be preserved under GS better than pedigree BLUP per cycle due to less inbreeding (Daetwyler et al., 2007; Jannink et al., 2010), a

result observed empirically by Asoro et al. (2013). However, plant breeders traditionally make selection based on phenotype, not pedigree BLUP and in reality GS may more rapidly decrease genetic variation and increase inbreeding (reviewed in Jannink et al., 2010 and Rutkoski et al., 2015), which was indeed observed by Rutkoski et al. (2015). Other studies have discussed the possible use of genomewide markers to predict σ_A in bi-parental populations to guide the selection of breeding populations with high mean performance and σ_A (Zhong and Jannink, 2007; Bernardo, 2014; Mohammadi et al., 2015). Initial empirical validations of such an approach are tenuous, but generally supportive (Lian et al., 2015; Tiede et al., 2015).

Prediction accuracy is driven by the genetic relationships between the TP and selection candidates, and the linkage disequilibrium (LD) between markers and QTL in the selection candidates being represented within the TP (Habier et al., 2007). These factors are, in turn, influenced by the number of genetic markers, statistical model used, TP size, and TP composition, as has been demonstrated in the empirical studies cited above. Training population composition, in particular, has received considerable attention recently, and a number of TP optimization algorithms have been proposed (Rincent et al., 2012; Lorenz and Smith, 2015; Isidro et al., 2015; Akdemir et al., 2015).

One way to simultaneously increase the efficiency and effectiveness of GS in plant breeding would be to design smaller TPs that are equally, and perhaps more predictive than their larger counterparts. Smaller TPs would enable breeders to better allocate resources and obtain higher quality phenotypic data, which is known to impact the accuracy of genomic predictions (Combs and Bernardo, 2013b; Lado et al., 2013). To

identify the optimal subset of *n* individuals from a larger set of *N* candidates, Rincent et al. (2012) proposed two procedures. The first seeks to minimize the prediction error variance (PEV) of the contrasts between each on the *n* entries in a TP subsample with the mean of the subsample, whereas the second seeks to maximize the coefficient of determination for the same contrasts (Rincent et al., 2012). Using CV and two empirical datasets, Rincent et al. (2012) demonstrated that maximizing the coefficient of determination (i.e. the CDmean procedure) always resulted in more reliable predictions (reliability was defined as $r\mathrm{G},\hat{\mathrm{G}})^2 \Big/ h^2$, where h^2 is the narrow-sense heritability) than a random sample of *n* individuals, and generally more reliable predictions than their PEV-based method (Rincent et al., 2012). Additional studies further support the use of CDmean for designing more reliable and predictive TPs (Tayeh et al., 2015; Isidro et al., 2015).

Isidro et al. (2015) tested CDmean in structured populations and also introduced a stratified sampling (SS) procedure, which randomly samples individuals proportionally to that subpopulation's respective size within the TP. Under strong population structure SS performed consistently better than CDmean and other procedures tested (Isidro et al., 2015). CDmean and SS each seek to maximize the variance explained by the TP and are convenient insofar that they only require genotypic and phenotypic information from the TP candidates when searching for an optimal TP subset. In contrast, approaches proposed by Lorenz and Smith (2015) and Akdemir et al. (2015) seek an optimal TP subset with respect to a genotyped set of selection candidates.

Lorenz and Smith (2015) outline a process that selects the *n* individuals from the TP with the highest mean relationship to the selection candidates. Starting with the most related TP candidates, $r_{\hat{G},G}$ is increased by the sequential addition of, relatively speaking, less-related candidates to a point, after which the addition of more candidates is detrimental (Lorenz and Smith, 2015). Akdemir et al. (2015) identifies a more predictive TP subset of size *n* by minimizing the ridge regression PEV among the selection candidates, which is related to maximizing the reliability measure of VanRanden (2008).

Breeders using GS desire to maximize $r_{\hat{G},G}$, and therefore gain from selection for a defined set of resources. The reality is, however, that breeding decisions are often time-sensitive and based on incomplete information. Our study exemplifies this reality as we consider the first three cycles of selection in a spring, six-row barley breeding population that was initiated in 2010 at the University of Minnesota (UMN) by intercrossing advanced breeding lines from three six-row malting barley breeding programs – UMN, North Dakota State University (NDSU), and Busch Agriculture Resources Inc. (BARI). The first objective of this study was to determine whether genetic gain has been made for increased grain yield (GY) and lower deoxynivalenol (DON) concentration. Grain yield is a classically quantitative trait, and DON accumulation, resulting from infection by the fungal pathogen *Fusarium graminearium* is known to also be a polygenic trait controlled by small effect QTL that are largely influenced by the environment (Ma et al., 2000; Mesfin et al., 2003; Massman et al., 2011). In 2010, the effect of TP size and composition on $r_{\hat{G},G}$ was less understood (Riedelsheimer et al., 2013, Lorenz and Smith, 2015). Consequently, the approach was to include as many relevant entries in the TP as possible. As such, the original TP (ORIG_TP) consisted of 384 breeding lines from each of UMN

and NDSU breeding programs, which had been evaluated as part of the Barley Coordinated Agricultural Project (BCAP, www.barleycap.org). Knowing the importance of TP composition, the second objective of this study was to retroactively investigate the change in $r_{\hat{G},G}$ when i) BARI lines evaluated in BCAP are added to ORIG_TP to more comprehensively represent the composite breeding population, ii) the TP is updated with recently evaluated breeding lines when making predictions, and iii) one of the four recently proposed TP optimization algorithms is implemented.

Materials and Methods

Germplasm Development

In 2009, the University of Minnesota (UMN) spring, six-row barley breeding program initiated a novel breeding population and transitioned to the use of GS. The BCAP provided an essential resource during this transition and is thoroughly explained by Hamblin et al. (2010). Briefly, during the four year span from 2006-09, ten barley breeding programs around the United States contributed up to 96 breeding lines per year for coordinated field testing and genotyping. The lines submitted to the BCAP were evaluated for many agronomic, disease, and malting quality traits; of primary interest in this study are GY and DON. Additionally, the lines submitted to BCAP were genotyped with a common set of SNP markers (Hamblin et al., 2010). The BCAP data can be downloaded from the Triticeae Toolbox website (www.triticeaetoolbox.org).

To establish the new UMN breeding population, advanced breeding lines elite for GY and DON from three breeding programs – seven from UMN, eight from North

Dakota State University (NDSU), and six from Busch Agricultural Resources (BARI) spring six-row programs – were intermated to generate the Cycle 0 (C0) breeding population (Figure 1.1). Recombinant progeny were inbred to the F_3 generation by single seed descent and genotyped. Selection among approximately 1400 F_3 progeny was based on their predicted genotypic values for each trait. Beginning thresholds set to the maximum predicted genotypic value for GY, and minimum predicted genotypic value for DON, were simultaneously relaxed until approximately 100 selection candidates passed the threshold criteria for both traits to generate the Cycle 1 selected population (C1S). A set of malting quality traits were also considered during selection, but only a few extremely inferior lines were removed using these predictions. Concurrently, a similarly sized set of F_3 contemporaries were chosen at random to form the Cycle 1 random population (C1R). The F_3 lines of C1S were intermated to generate the Cycle 2 (C2) selection candidates (Figure 1.1). Additionally, C1S and C1R were advanced to $F_{3:5}$ families for evaluation of GY and DON in the field (Figure 1.1). This breeding cycle was repeated annually to develop an experimental population representing selected and random progeny from the first three cycles.

Datasets:

ORIG_TP: The TP used for the original predictions that were used to make selections (blue box in Figure 1.1). The original TP consisted of 384 UMN and 384 NDSU lines submitted to BCAP. For C3 predictions, a set of 300 random C1 lines generated from crosses between the seven UMN and six NDSU progenitors (Lorenz and Smith, 2015) was also included in the TP.

CAP_TP: An update of ORIG_TP that includes data for BARI lines submitted to BCAP as well (purple box in Figure 1.1). CAP_TP consisted of the adjusted means (i.e. best linear unbiased estimates) of a TP consisting of 374 UMN, 381 NDSU, and 148 BARI lines submitted to BCAP; grain yield data is only available for 114 of the BARI lines. Lines submitted to BCAP were evaluated in a standardized set of regional disease nurseries, but the participants performed yield trials individually. In general, the trials (a specific year-location combination) were grown in a randomized complete block design with two or three replications per trial, and a common check was included in all trials to enable the estimation of trial effects. Ultimately, adjusted means for DON are derived from 17 disease trials across the four BCAP years, and adjusted means for GY are derived from 56 yield trials evaluated during the same time period. All trial data was downloaded from www.triticeaetoolbox.org/barley.

BREED_TP: The adjusted means of C1 and C2 breeding lines tested in field trials during the routine operation of the barley breeding program (red box in Table 1.1). In C1, there were 98 lines selected by GS and 300 random lines from the intermating between the UMN and NDSU progenitor lines, which were also advanced for a separate experiment (Lorenz and Smith, 2015). However, just 72 C1S lines were field tested so C1 in BREED_TP consists of 72 selected lines and 72 lines sampled from the 300 random lines. The full C2 population of BREED_TP consisted of 105 C2S lines and 101 C2R lines. In this study, when it is described that breeding cycles were added to the TP, the values for those breeding lines are derived from BREED_TP.

GS_VP: The fourth dataset consists of approximately 50 selected and 50 random lines from each of the first three breeding cycles, as well as the 21 progenitor lines, that

were evaluated in a balanced set of five yield trials and five disease nurseries (Supplementary Table 1.1). The best linear unbiased predictions (BLUP) calculated from this dataset represent the estimated breeding values of each breeding line in C1 through C3 and, as such, will be used as the values for the validation population (VP) for all analyses, hence GS_VP. The trials were planted in augmented designs based on Lin and Poushinsky's (1985) modified augmented design (type II) consisting of 24 incomplete blocks, each containing 15 subplots. A common check is planted at the center of each incomplete block and three additional check varieties appear in a random subset of incomplete blocks. The design facilitates spatial adjustments (Lin and Poushinsky, 1985); however, spatial covariates calculated with a moving average algorithm or the inclusion of an AR1 x AR1 correlation structure into the mixed model (described in the *spatial corrections* section below) proved to be more effective in reducing spatially-related error variance.

Mixed model for calculating adjusted means and BLUPs

The mixed model used to calculate the adjusted means or BLUPs of the line effects, depending on the dataset, can be generalized to:

$$y_{ljirc} = \mu + trial_j + line_l + rep(trial)_{i(j)} + row(trial)_{r(j)} + column(trial)_{c(j)} + (line * trial)_{lj} + e_{ljirc}, \quad (1.1)$$

where y_{ljrc} is the observed phenotype (GY or DON), μ is the grand mean, $trial_j$ (j = 1,…n_j) is the fixed trial effect, $line_l$ (l = 1,…n_l) is the effect of l (details discussed below), $rep(trial)_{i(j)}$ (i = 1,…n_i) is the random effect of the i^{th} replication in the j^{th} trial

when analyzing an RCBD, with $rep(trial)_{i(j)} \sim \mathcal{N}(0, \mathbf{I}_{nj} \otimes \mathbf{I}_{ni} \otimes \sigma^2_{rep})$ where σ^2_{rep} is the variance of the random replication effect, $row(trial)_{r(j)}$ ($r = 1,\ldots n_r$) is the random effect of the r^{th} row in the j^{th} trial with $row(trial)_{r(j)} \sim \mathcal{N}(0, \mathbf{I}_{nj} \otimes \mathbf{I}_{nr} \otimes \sigma^2_{row})$ where σ^2_{row} is variance of the random row effect, and $column(trial)_{c(j)}$ ($c = 1,\ldots n_c$) is the effect of the c^{th} column in the j^{th} trial with $column(trial)_{c(j)} \sim \mathcal{N}(0, \mathbf{I}_{nj} \otimes \mathbf{I}_{nc} \otimes \sigma^2_{column})$ where σ^2_{column} is the variance of the random column effect, $(line * trial)_{lj}$ is the random interaction of the l^{th} line and the j^{th} trial with $(line * trial)_{lj} \sim \mathcal{N}(0, \mathbf{G}_{nl} \otimes \mathbf{I}_{nj} \otimes \sigma^2_{line\ x\ trial})$ where $\sigma^2_{line\ x\ trial}$ is the variance of the random *line* x *trial* effect, and e_{ljirc} is the residual term with zero mean and variance σ^2_e. Additionally, $\mathbf{I}_x$ is an identity matrix of order x, $\otimes$ is the Kronecker product, and **G** is the *nl x nl* genomic relationship matrix between the *l* lines calculated from genomewide markers using method 1 of VanRaden (2008). When calculating adjusted means (i.e. best linear unbiased estimates) for CAP_TP and BREED_TP, $line_l$ is fit as a fixed effect, but when calculating BLUPs for GS_VP then $line_l \sim \mathcal{N}(\mathbf{G}_{nl} \otimes \sigma^2_{line})$ where σ^2_{line} is the variance of the random *line* effect. To avoid over fitting the model, only significant *rep*, *row*, and *column* effects within trials were fit in multi-environment models.

Spatial corrections

When row and column coordinates were available for a trial, spatial error was modeled in one of two ways. The first was to add a fixed covariate to Eq. 1.1 accounting for local field trends at row *r* and column *c* in trial *j* calculated using the moving average function in the R package (R Core Team, 2015) mvngGrAd (Technow, 2015) as

described in Mohammadi et al. (2015). The second option involved fitting an additional random term $(column * row)[trial]_{rc(j)}$ in Eq. 1.1 with $(column * row)[trial]_{rc(j)} \sim \mathcal{N}(\mathbf{AR1}(\rho_{c(j)}) \otimes \mathbf{AR1}(\rho_{r(j)}) \otimes \mathbf{I}_{nj} \otimes \sigma^2_{column\ x\ row})$ where $\mathbf{AR1}(\rho_{c(j)})$ and $\mathbf{AR1}(\rho_{r(j)})$ are first order autoregressive processes for columns and rows within a trial, respectively. The correlation parameters ρ_c and ρ_r are bound between zero and one and jointly selected to maximize the likelihood function of models fit for individual trials within a multi-trial dataset. The use of **AR1** $\otimes$ **AR1** correlation structures is often used to model **R**, the covariance structure of the residual term (Cullis and Gleeson, 1991; Cullis et al., 1998; Cullis et al., 2006), whereas we fit the **AR1** $\otimes$ **AR1** correlation structure as the covariance matrix for the random interaction between row and columns, as in Robbins et al. (2012).

Phenotyping

The protocol for measuring FHB disease severity in misted and inoculated disease nurseries is described in Massman et al. (2011). Briefly, 3m single-plot rows are grown 0.3m apart in a mist-irrigated field and misted for 10 minutes once per hour, for 10 hours per day. In the Saint Paul, MN disease nursery, plants are inoculated shortly after heading with macrconidia produced from a collection of over 50 local isolates of *Fusarium graminearum* maintained by R. Dill-Macky (Steffenson, 2003). All other disease nurseries are inoculated using the grain spawn method previously described by Horsley et al. (2006). All of the spikes from the center 2.5m of each row are hand-harvested, bulked,

and threshed, dried, and then, the concentration of DON in each sample is measured using combination gas chromatography/mass spectrometry (Mirocha et al., 1998).

Yield trials were planted at a density of 1.2 million seeds per acre. Fertility and weeds were managed using standard practices for barley. Plot sizes were 1.9 m^2 in Saint Paul, 3 m^2 in Crookston, and 3.8 m^2 in Morris, MN, and were mechanically harvested at approximately 13.5% moisture.

Genotyping

All lines submitted to the CAP were genotyped with a set of barley oligonucleotide pool assay (BOPA) single nucleotide polymorphism (SNP) markers (Close et al., 2009; Hamblin et al., 2010). The breeding program selected a subset of 384 BOPA SNPs that were informative for the TP and evenly spaced throughout the genome and created a custom Illumina Veracode (Illumina, Inc., San Diego, CA) assay. This Veracode assay was used to genotype all breeding lines (i.e. those in BREED_TP and GS_VP).

All genotypic data were downloaded from the T3 database (www.triticeaetoolbox.org). Filters for missing marker and entry data were applied to the unique complement of entries included in the datasets to ensure that an identical set of markers is used for each analysis. Entries missing more than 20% marker data, and then markers missing more than 20% data or with a minor allele frequency less than 5% were removed. High correlations ($r > 0.975$) between entries or marker were resolved by retaining the observation with the lowest missing data and removing the others. Ultimately, 335 SNP markers were used for all analyses.

Genomic prediction

Genomic predictions were based on the following model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}\theta + \mathbf{e}, \quad (1.2)$$

where **y** is a vector of adjusted means of length *nTP*, the number of entries in the TP, **X** is an *nTP* x *nb* design matrix for **b**, which is a vector of fixed effect coefficients for the *nb* datasets that contribute entries to the TP, **Z** is the *nTP* x *N* design matrix for the random effects, where $N = nTP + nVP$, with *nVP* being the number of entries in the VP, and **e** is the residual error term with $\mathbf{e} \sim \mathcal{N}(0, \mathbf{I}_{nTP} \otimes \sigma_e^2)$. Replacing θ with **g**, such that $\mathbf{g} \sim \mathcal{N}(0, \mathbf{G}_N \otimes \sigma_g^2)$, where $\mathbf{G}_N$ is an *N x N* additive relationship matrix and σ_g^2 is the additive genetic variance results in prediction by genomic BLUP (GBLUP; Bernardo, 1994). Alternatively, replacing θ with $\boldsymbol{\alpha}$, such that $\boldsymbol{\alpha} \sim \mathcal{N}(0, \mathbf{K}_N \otimes \sigma_\alpha^2)$ where $\mathbf{K}_N$ is an *N x N* correlation structure defined by the reproducing kernel Hilbert spaces (RKHS) procedure (Gianola and van Kaam, 2008). Since the kernel of RKHS is not necessarily linear it may better capture non-additive effects, and therefore σ_α^2 could explain genetic variance beyond σ_g^2. The RKHS procedure herein is based on previous work (Heslot et al., 2012; de los Campos and Pérez, 2015) such that $\mathbf{K}_N = e^{-h * d_{ij}}$, where d_{ij} is the Euclidean distance between entries calculated from their marker genotypes, and *h* is the smoothing parameter of the kernel that controls the decay rate of the correlation coefficients representing the pairwise similarity among lines. To select a reasonable value of *h* we search $h \subset [0.1, 2.5]$ (a range offered in the vignette of de los Campos and Pérez, 2015) for the value of *h* that maximizes the ten-fold CV accuracy of the TP.

After solving the mixed model a sub-vector of $\mathbf{g}$ or $\boldsymbol{\alpha}$ coinciding with the selection candidates are taken as the predicted genotypic value or total genetic value, respectively. All analyses were conducted in R (R Core Team, 2015), and all mixed models in this study were solved using package 'EMMREML' (Akdemir and Godfrey, 2015).

Measuring gain from selection across breeding cycles

Original selections were based on the predicted genotypic values for GY and DON, so the simultaneous response of GY and DON was evaluated from an index calculated as:

$$\mathrm{GY.DON} = (100 + \mathrm{GY}^{*}) - \mathrm{DON}^{*},$$

where GY^{*} and DON^{*} are GY and DON rescaled to the interval [0, 100]. The GY.DON index rewards higher values to lines with higher GY or lower DON.

The selected lines in each breeding cycle are compared to i) the mean of the randomly sampled lines in the respective breeding cycle, and ii) the C1R lines. Ideally (ii) would be a comparison to the C0 population, but the C0 population was not explicitly evaluated as part of GS_VP. Instead, (ii) considers a comparison to C1R since, we argue, it is the best representation of the C0 population (Figure 1.1). We believe this to be true despite the fact that C1R is derived from only crosses between UMN and NDSU progenitors. In fact, BARI lines are related to both the UMN and NDSU breeding populations as demonstrated by previous population structure analyses (Hamblin et al., 2010; Lorenz et al., 2012). Finally, all sample means and variances are calculated from lines' BLUPs from GS_VP. Differences detected by comparisons (i) and (ii) exceeding a

least significant difference (LSD) value calculated at $\alpha = 0.05$ were considered significantly different.

Genetic Variance

Genetic variances specific to breeding cycles and to either random or selected lines within the respective breeding cycle were calculated using a two-step procedure (Rutkoski et al., 2015). First, the mixed model described by Eq. 1.1 is fit without the random *line* term, but modified to include a factor indicating the subpopulation of $line_i$. The residuals (y') of that model fit (i.e. the adjusted phenotypic values) are fit in a subsequent cycle-specific model:

$$y_i' = \mu + line_i + e_i. \tag{1.3}$$

The variance component of *line* when fit as a random effect as described for Eq. 1.1 results in an estimate of the additive genetic variance, whereas modeling $line_i$ as $line_i \sim \mathcal{N}(0, \mathbf{I}_{nl} \otimes \sigma^2_{line})$ results in an estimate of the genetic variance, where $\mathbf{G}_{nl}$ and $\mathbf{I}_{nl}$ relate to the lines specific to the breeding cycle under consideration but are otherwise as described for Eq. 1.1. The square roots of the genetic and error variances are reported.

Updating TP

We employ a series of strategies to retrospectively investigate how practices, if implemented at various stages during the progression of the UMN barley breeding program, would have influenced $r_{\hat{G},G}$ and consequently genetic gain. Recall that the original predictions using ORIG_TP did not include BCAP entries from BARI (Figure

1.1). Accordingly, we first compare the original prediction accuracies to prediction accuracies using CAP_TP.

The UMN barley program routinely evaluates selected progeny in each breeding cycle at the $F_{3:5}$ generation during the following growing season. For the first three cycles, the program also advanced a randomly selected set of progeny (Figure 1.1). The second updating strategy involves adding phenotyped breeding lines to the TP when predicting subsequent cycles. For example, we test the change in $r_{\hat{G},G}$ when C1S lines from BREED_TP are added to CAP_TP to predict C2 lines, which is also compared to the addition of all C1 lines to the TP to evaluate the benefit of advancing a random set of lines. Continuing the example, we also evaluate the change in $r_{\hat{G},G}$ of C3 lines when the accumulation of C1S and C2S lines are added to CAP_TP, and again, when the accumulation of C1 and C2 lines are added to the TP.

Ultimately the effect of updating the TP is tested across nine combinations of TP and VP. These nine combinations will be referred to as 'cases' hereafter. When the VP is comprised of progeny derived from the TP we refer to $r_{\hat{G},G}$ as progeny-based as opposed to cross-validation based on subsampling within the TP. A 95% confidence interval was constructed for the progeny-based $r_{\hat{G},G}$ of each case from 2,000 bootstrapped samples (Efron, 1979). Non-overlapping confidence intervals suggest that the prediction accuracies are significantly different at $\alpha = 0.05$.

TP optimization procedures

For brevity we refer the reader to the original publications of each TP optimization procedure for specific details, and provide just the details and parameters pertinent to our study here:

SS (Stratified Sampling; Isidro et al., 2015): Population structure is known to exist within the ORIG_TP and CAP_TP (Hamblin et al., 2010; Lorenz et al., 2012). Training population candidates were assigned to one of three subpopulations defined by passing a Euclidean distance matrix calculated from the marker genotypes of the TP candidates to a hierarchical clustering algorithm in R (R Core Team, 2015); three clusters were chosen from knowledge of the breeding population's development. Training populations of size n were then built by combining n_c individuals from each of the *c* clusters, where $n_c \propto {nTP_c}/{nTP}$, with nTP_C being the number of individuals in the TP assigned to cluster *c*, such that $n = \sum_1^c n_c$.

G_mean (Lorenz and Smith, 2015): Marker-based relationships between the *N* lines in the TP and VP were calculated using method 1 of VanRaden (2008), resulting in matrix $\mathbf{G}_N$. The dimensions of $\mathbf{G}_N$ were reduced so only TP candidates were represented across the rows and only VP entries were represented across the columns. Thus, the mean of each row represents the mean relationship of individual TP candidates to the entire VP. The *n* TP candidates with the highest mean relationship to the VP are taken as the optimized TP subset.

StratCDmean (Rincent et al., 2012; Isidro et al., 2015): A within-subpopulation exchange algorithm was implemented using the CDmean script provided by R. Rincent (personal communication), which was modified according to the stratified sampling suggestion in Isidro et al. (2015). After selecting a random subset of *n* lines from the TP, each iteration of the algorithm exchanges an existing line in the subset with a TP candidate not presently included in the TP subset. If the exchange increases the trace of

$$\left[\frac{\mathbf{c}'(\mathbf{G}-\lambda_{\mathrm{CDmean}}(\mathbf{Z}'\mathbf{W}\mathbf{Z}+\lambda_{\mathrm{CDmean}}\mathbf{G}^{-1})^{-1})\mathbf{c}}{\mathbf{c}'\mathbf{G}\mathbf{c}}\right],$$

then the newly introduced candidate will be retained and the replaced entry will return to the set of non-sampled TP candidates. In the equation above, **c** is a matrix of contrasts comparing each column (i.e. a TP candidate) to the mean of all TP candidates, **Z** is the design matrix of random entry effects as defined above, $\mathbf{W}=\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$, where **X** is the design matrix of fixed effects as defined above, and **G** is a relationship matrix calculated from marker effects (Rincent et al., 2012). Rincent et al. (2012) found the algorithm to be robust to the value of $\lambda_{\mathrm{CDmean}}$, so its value was fixed at 1, which corresponds to a trait heritability of 0.50. Per the recommendation of Rincent et al. (2012) the allele frequency weighted relationship matrix ($\mathbf{G}_{\mathrm{freq}}$) was used. Given the *nTP* x *m* matrix $\mathbf{M}_{\mathrm{TP}}$ of TP genotypes coded as 1, 0, -1, **p** is then a vector of length *m* with each element p_i being the frequency of the allele coded as "1" at marker *i,* **P** is an *nTP* x *m* matrix such that all elements in column *i* are p_i, $\mathbf{C}=\mathbf{M}_{\mathrm{TP}}-\mathbf{P}$, and $\mathbf{s}=\mathbf{p}*(\mathbf{1}-\mathbf{p})$ so that

$$\mathbf{G}_{\text{freq}} = \frac{(\mathbf{C} * \mathbf{s}^{-\mathbf{1}})\mathbf{C}'}{m},$$

where * denotes the element-wise multiplication of matrices and vectors.

Isidro et al. (2015) suggested replacing the random sampling step of SS with the CDmean algorithm in structured populations as it considerably reduced computational time and resulted in comparable accuracies to the CDmean applied to the entire set of TP candidates. For each subpopulation cluster within the TP, the StratCDmean algorithm was iterated 10,000-12,000 times depending on *nTP*, which was sufficient for the algorithm to converge at a maximum for the subpopulation-specific CDmean criteria.

STPGA (Akdemir et al., 2015): STPGA was implemented using the R script provided in the supplemental materials of Akdemir et al. (2015). Given a phenotyped and genotyped TP and a set of genotyped selection candidates, STPGA identifies a TP subset that minimizes the ridge regression PEV among the selection candidates, calculated as:

$$\text{PEV}^{\text{Ridge}}(\mathbf{M}_{\text{VP}}) = \mathbf{M}_{\text{VP}}(\mathbf{M}'_{\text{TP}}\mathbf{M}_{\text{TP}} + \lambda_{\text{STPGA}}\mathbf{I})^{-\mathbf{1}}\mathbf{M}'_{\text{VP}},$$

where $\mathbf{M}_{\text{TP}}$ and $\mathbf{M}_{\text{VP}}$ are the genotype matrices of the TP and VP, respectively, and λ is set to $1/m$ which corresponds to a trait heritability of 0.50. To increase the computational speed, especially in cases of high marker density, the exact $\text{PEV}^{\text{Ridge}}(\mathbf{M}_{\text{VP}})$ is estimated by replacing the marker matrices $\mathbf{M}_{\text{TP}}$ and $\mathbf{M}_{\text{VP}}$ with principle components matrices calculated from the marker matrices themselves, or a genomic relationship matrix. STPGA parameters suggested in Akdemir et al. (2015) were modified in a manner that decreased computational

time while maintaining the accuracy of STPGA (data not shown). Specifically, in this study we replace $\mathbf{M}_{\mathbf{TP}}$ and $\mathbf{M}_{\mathbf{VP}}$ with *nTP* x 30 and *nVP* x 30 matrices, respectively, of the first 30 PCAs calculated from a genomic relationship matrix between the TP and VP as described for $\mathbf{G}_{\text{freq}}$ above. The first 30 PCAs explained 92.0% of the variation, whereas the first 40 explained 93.8%. Also, the number of populations simulated during each iteration of STPGA was reduced from 800 to 400. The number of iterations needed for the algorithm to converge ranged from 300 to 400 depending on *nTP*.

The following steps were implemented to compare the four TP optimization procedures:

1. Run StratCDmean to identify an optimized set of *n* lines from the TP candidates,
2. Sample a random set of 50 lines from the VP,
3. Sample a random set of *n* lines from the TP candidates,
4. Run SS, G_mean, and STPGA to identify optimal sets of *n* lines from the TP candidates,
5. Use the TP subsets identified in (3) and (4) to predict genotypic values using GBLUP and total genetic values using RKHS,
6. Calculate $r_{\hat{G},G}$ of the estimates from (5) and the BLUPs of the VP sample of (2),
7. Repeat (1) through (6) for 50 iterations,
8. Calculate the mean, median, and variance across the 50 $r_{\hat{G},G}$ coefficients from (6) for the randomly chosen TP in (3) and all optimization procedures in (4).

Significant pairwise difference between the mean prediction accuracies of the randomly sampled and optimized TP subsets were identified using the Tukey-Kramer multiple comparison procedure modified by Dunnet (1980). The proportion of times the $r_{\hat{G},G}$ of an optimization procedure is greater than the accuracy of the randomly sampled TP is also reported. The accuracy of GBLUP under each optimization procedure was compared to that of RKHS across the 50 iterations using a two-tailed t-test.

For DON, we used a TP size of $n = 300$ based on Lorenz et al. (2012) and Sallam et al. (2015), which both showed diminishing returns of $r_{\hat{G},G}$ as the TP size approaches and surpasses 300 individuals. For GY, Sallam et al. (2015) showed that $r_{\hat{G},G}$ in barley consistently increased as the TP size is increased, which appears to be consistent if we extrapolate the results of Lorenz et al. (2011) past 300 individuals; thus $n = 500$ for GY.

ANOVA

An analysis of variance (ANOVA) was conducted, where the aggregate of prediction accuracies over iterations, for each optimization procedure, across all cases tested during the 'TP Optimization' procedure was fit as the response variable. Explanatory variables include the main effects for the prediction model (GBLUP or RKHS), optimization procedure, and breeding cycle of the VP, as well as all of the appropriate two- and three- way interactions among those main effects and the TP. The main effect of TP was not included because the imbalance resulting from the nesting of TP within VP could result in misleading interpretations. The ANOVA was conducted in R (R Core Team, 2015) using type III sums of square, i.e. the significance of each effect was ascertained by comparing the change in the total sum of squares of the full model

with and without the explanatory variable fit. The model was then refit without the non-significant explanatory variables and differences between the levels of significant main effects were identified using Tukey's HSD (Lane and Salkin, 2010).

Results

Effect of GS on μ and σ_A^2 of Yield and DON

We observed significant gain for the two unfavorably correlated traits, yield and DON, across three cycles of genomic selection (Figure 1.2). The progeny-based $r_{\hat{G},G}$ of the original predictions of C1, C2, and C3 are 0.124, 0.264, and 0.415 for GY, and 0.288, 0.194, and 0.182 for DON, which were all lower than the accuracies of ORIG_TP predicted by 10-fold CV of 0.638 for GY, and 0.688 for DON. The average performance of GY, DON, and the GY.DON index in C3S were all significantly greater than their respective mean values in C1R (Figure 1.2). Across this contrast, GY increased 164.74 kg ha^{-1}, DON decreased 1.41 ppm, and the GY.DON index increased 23.36 points. While GY and DON did not individually respond to selection consistently across the cycles, the GY.DON index increased from 102.37 in C1S to 110.24 in C2S, which then significantly increased to 122.17 in CS3 (Figure 1.2).

The square root of the additive genetic variance for GY increased from 121.32 in C1 to 183.29 in C2, but then slightly decreased from C2 to 180.40 in C3 (Figure 1.2). The σ_A of DON decreased consistently over breeding cycles from 1.29 in C1 to 1.17 in C2, then to 1.10 in C3 (Figure 1.2). The square root of total genetic variance (σ_G) followed

the same trend as σ_A, and was on average 9.2% greater than σ_A for GY, and 23.6% greater than σ_A for DON (data not shown).

Effects of updating the TP

The addition of the BARI lines, which were not part of the original TP, improved $r_{\hat{G},G}$ for GY and DON in C1 and C2, but decreased $r_{\hat{G},G}$ for GY and DON in C3 (Table 1.1). The increase in accuracy from 0.288 to 0.723 when including BARI lines to predict DON in C1 was the only significant difference observed (Table 1.1).

Updating CAP_TP with recently phenotyped breeding lines always increased $r_{\hat{G},G}$ of GY, and generally increased $r_{\hat{G},G}$ for DON (Table 1.1). Furthermore, the greatest increase in $r_{\hat{G},G}$ is realized when just the selected lines from previous breeding cycle(s) are included in the TP to predict subsequent cycles (Table 1.1). Focusing on GBLUP results, $r_{\hat{G},G}$ of GY in C2 increases from 0.284 to 0.362 when C1 is added to CAP_TP, and to 0.424 when just C1S is added. Predictions for GY in C3 increased from 0.299 to 0.400 upon the addition of C1S, but to just 0.321 when C1 was added. Adding C1 and C2 to CAP_TP to predict GY in C3 increased $r_{\hat{G},G}$ to 0.477, but again, when C1S and C2S were added the $r_{\hat{G},G}$ increased to 0.535 (Table 1.1). For GY, the main effect of VP was not significant (Table 1.2).

Updating the training population for DON behaved differently than doing so for GY. Focusing again on GBLUP results, adding C1 or C1S to CAP_TP did not influence prediction accuracy of DON in C2, and the prediction accuracies of DON in C3 decreased from 0.121 to 0.097 and 0.115 when C1 and C1S were added, respectively (Table 1.1). When C1 and C2 were both added, however, the $r_{\hat{G},G}$ of DON in C3

increased from 0.121 to 0.197, and the addition of C1S and C2S to CAP_TP further increased $r_{\hat{G},G}$ to 0.237 (Table 1.1). For DON, the main effect of VP was significant in the ANOVA, with C1 being the best predicted followed by C2 and then C3 (Table 1.2).

Effects of optimizing the TP

We examined four TP optimizing procedures and a random selection of *n* lines for each of the nine TP updating cases described in the '*Updating TP*' section above, resulting in 45 total combinations. Based on these 45 combinations, the benefit of TP optimization differed for GY and DON (Supplementary Table 1.2).

For GY predictions by GBLUP, one or more optimized TP subset of size $n = 500$ resulted in greater $r_{\hat{G},G}$ than using CAP_TP in its entirety in eight of the nine cases (Supplementary Table 1.2). With one exception (drawing subsets from CAP_TP to predict C1), for each case there existed at least one optimization procedure that identified TP subsets than were, on average across the 50 iterations, significantly more predictive than a random subset of size *n* (Supplementary Table 1.2). For GY, G_mean resulted in the highest $r_{\hat{G},G}$ of any optimization procedure in four of the nine cases (Table 1.3). StratCDmean and STPGA resulted in the highest $r_{\hat{G},G}$ in two of the nine cases, each, while SS never resulted in the overall highest $r_{\hat{G},G}$ (Table 1.3). Furthermore, with the exception of the case noted above, G_mean always resulted in higher $r_{\hat{G},G}$ than a random TP subset, and in most cases the increase was significant (Supplementary Table 1.2). In contrast the average $r_{\hat{G},G}$ of the remaining STPGA, StratCDmean and SS were inferior to the random TP subset in the majority of cases (Table 1.3).

The main effect for the optimization procedure used in the ANOVA model for GY was highly significant, with G_mean being the most effective, followed by STPGA (Table 1.2). Stratified sampling and StratCDmean overall negatively impacted $r_{\hat{G},G}$, even compared to the random sampling procedure (Table 1.2). There was not a significant two-way interaction between the optimization procedure and statistical model used, however other two- and three-way interactions between the optimization procedure, TP, VP, and statistical model used were significant (Table 1.2).

Considering DON predictions by GBLUP, one or more optimized TP subset of size n = 300 resulted in greater $r_{\hat{G},G}$ than using CAP_TP in its entirety in all of the nine cases (Supplemental Table 1.2). Similarly, for each case, at least one optimization procedure identified TP subsets that were, on average across the 50 iterations, significantly more predictive than a random subset of size n (Supplementary Table 1.2). For DON, StratCDmean resulted in the highest $r_{\hat{G},G}$ of any optimization procedure in four of the nine cases, G_mean in three cases, and STPGA in the remaining two cases (Table 1.3). Again, SS never resulted in the overall highest $r_{\hat{G},G}$ (Table 1.3), but did result in higher $r_{\hat{G},G}$ than a random TP subset across a majority of the 50 iterations in seven of the nine cases, a higher frequency than any other optimization procedure (Table 1.3). However, the difference in average $r_{\hat{G},G}$ for SS is never significantly higher than that of the random subset, whereas the other optimization procedures do result in significantly greater $r_{\hat{G},G}$ in some cases (Supplementary Table 1.2).

The main effect of the optimization procedure in the ANOVA model for DON was highly significant with STPGA being the most effective, followed by StratCDmean and G_mean, which were not significantly different from one another (Table 1.2).

Prediction accuracies resulting from SS were the lowest and not significantly different than random sampling (Table 1.2). The two-way interactions between the optimization procedure and the main effects of TP and VP were significant, but the interaction with the statistical model used was not (Table 1.2).

GBLUP vs. RKHS

Whether GBLUP or RKHS was used significantly affected $r_{\hat{G},G}$ in 43 of the 45 combinations described above (Supplementary Table 1.2). RKHS resulted in higher $r_{\hat{G},G}$ in 39 of those 43 combinations, and the four combinations when GBLUP resulted in higher $r_{\hat{G},G}$ all occurred when TP subsamples were drawn from CAP_TP to predict C1 (Supplementary Table 1.2). Additionally, the main effect of the statistical model used was significant in the ANOVA analysis with RKHS resulting in significantly higher $r_{\hat{G},G}$ than GBLUP for GY (Table 1.2). The two-way interactions between the prediction model used and the main effects of TP and VP were also significant (Table 1.2). While the two-way interaction between the prediction model and optimization procedure was not significant, three-way interactions including the prediction model and optimization procedure were (Table 1.2).

The prediction model used significantly affected prediction accuracy for DON in 13 of the 45 combinations (Supplementary Table 1.2). In seven of those 13 combinations, all of which involved TP optimization by either G_mean or STPGA, RKHS resulted in higher $r_{\hat{G},G}$. The majority of the six instances where GBLUP resulted in significantly higher $r_{\hat{G},G}$ occurred when optimized subsets from CAP_TP predicted C3. The main effect of the prediction model used was not significant in the ANOVA analysis for DON

(Table 1.2), and only one interaction term involving the effect of prediction model – the three-way interaction between the prediction model, optimization procedure, and VP – was significant (Table 1.2).

Discussion

Response to genomic selection

Significant genetic gains were observed after three cycles of genomic selection for GY and DON independently, as well as an index expressing their joint response, despite their unfavorable genetic correlation of 0.752 (see supplemental method section for description). The case presented herein is representative of the dilemma often faced in breeding – the need to simultaneously improve multiple quantitative traits that may be unfavorably correlated. Considering the constraints placed on breeding programs that our study emphasizes, it is not surprising that the trait mean and genetic standard deviations of the breeding cycles did not always respond to selection in a consistent and predictable manner.

First, the accuracies of the original predictions were relatively low, especially for GY with an $r_{\hat{G},G}$ of 0.124. Such low $r_{\hat{G},G}$ were unexpected since the $r_{\hat{G},G}$ of ORIG_TP predicted by CV were 0.638 and 0.688 for GY and DON, respectively. This low selection pressure may have allowed genetic loci affecting GY to drift, resulting in an unimproved mean and change in genetic standard deviation that increases rather than the decrease expected in response to effective selection. Second, the unfavorable additive genetic correlation between GY and DON makes it hard to simultaneously improve GY and

DON. On account of this unfavorable genetic correlation, under the slightly stronger selection pressure for DON given its higher $r_{\hat{G},G}$ in C1, genetic loci influencing GY may have moved opposite of fixation towards more intermediate frequencies, again resulting in the lower mean and higher genetic standard deviation observed for GY in C2. Third, there were traits in addition to GY and DON considered when making selections, in particular a set of malting quality traits; as more traits are considered the expected response of any one trait decreases. Little can be done to mitigate the effect of the second and third constraints, but we have shown how actively maintaining the TP through updating and optimization can increase $r_{\hat{G},G}$ downstream.

Optimizing the TP improves prediction accuracy

We expected that more representative training populations would result in higher prediction accuracies. We retrospectively assessed the effect of three alternative selection strategies on the accuracy of genomic predictions. Amending ORIG_TP to include BARI breeding lines to more comprehensively represent the selection candidates in C1-C3 increased the $r_{\hat{G},G}$ of both traits in C1 and C2, but not C3. This is ostensibly explained by the addition of 300 random C1 lines to the original TP for original C3 predictions, however, that explanation seems inadequate given that the random C1 lines were not derived from any BARI progenitors. Comparing ORIG_TP for C3 consisting of 1,084 lines with 0% BARI representation, to the case where C1 is added to ORIG_TP to make a TP of 1,058 lines with approximately 13.4% BARI representation, the $r_{\hat{G},G}$ for GY and DON of the former is 0.415 and 0.182, whereas the $r_{\hat{G},G}$ of GY and DON of the latter is

0.321 and 0.097, respectively. Although the $r_{\hat{G},G}$ of the compared values are not significantly different, they certainly deviate from the overarching trend.

An alternative explanation is that the quality of the data collected on the random 300 C1 lines is better than the data collected on the C1 lines, which is known to affect $r_{\hat{G},G}$ (Combs and Bernardo, 2013a; Lado et al., 2013). This seems reasonable given that the random 300 C1 lines were grown in balanced experiments across two years in seven total environments (year-location combinations) for GY and eight total environments for DON (Lorenz and Smith, 2015). In contrast, the C1 lines were evaluated in a single year, with the 72 C1S lines being evaluated in three locations for GY and two locations for DON, and the 72 random C1 lines being evaluated in three locations for GY and four locations for DON (Supplemental Table 1.1).

The second strategy of updating the TP with data collected on lines from previous breeding cycles to the TP increased $r_{\hat{G},G}$. This practice was shown to be beneficial in the simulation experiment of Jannink (2010) and reportedly lead to increased $r_{\hat{G},G}$ in Rutkoski et al. (2015). Our results provide further empirical validation for the practice of adding lines from previous breeding cycles. However, we observed a difference between adding selected lines and adding all lines where the former was always better than the latter. The lower $r_{\hat{G},G}$ as the result of adding all lines, rather than only selected lines, occurred despite the increased mean kinship between the TP and VP ($\overline{K}_{ij}$; Figure 1.3) suggesting that improvements in $r_{\hat{G},G}$ are being driven by the marker-QTL LD in the selection candidates being better represented by the selected breeding lines (Habier et al., 2007).

Using an optimization algorithm identified a TP subset that was more predictive than the complete TP, which ranged in size from 851 to 1,233, in all but one of the 18 cases under the GBLUP model. The $r_{\hat{G},G}$ of optimized TP subsets are significantly higher than $r_{\hat{G},G}$ of the randomly sampled TP subset in nearly all cases, suggesting the increases in $r_{\hat{G},G}$ are not spurious and truly result from the algorithms identifying more predictive TP subsets. The original publications of CDmean (Rincent et al., 2012), STPGA (Akdemir et al., 2015), and SS (Isidro et al., 2015) present the situation where genotypes are available for a diverse collection of germplasm, but given limited resources for phenotyping an optimization algorithm is implemented to identify an optimal TP subset that is more predictive than a randomly chosen subset.

We argue that this scenario is only valid for the initial stages of transitioning to GS, as was the situation of the case presented herein, and indeed we retroactively demonstrate that the use of a TP optimization algorithm would have, generally speaking, increased $r_{\hat{G},G}$ and therefore genetic gains (see Eq. 1.1). Of special interest to breeding programs routinely using GS would be a procedure that identifies selection candidates that, if advanced along with the selected lines would provide an additional benefit to the selected lines alone in terms of maintaining or improving $r_{\hat{G},G}$. The further addition of random lines to the TP increased K_{ij} (Figure 1.3) between the TP and selection candidates, but did not increase $r_{\hat{G},G}$, which suggests that G_mean may be insufficient for such a task. A more viable approach may be the use of STPGA to identify additional selection candidates that minimize the PEV among the candidates selected based on their predicted genotypic values.

All three strategies independently improve $r_{\hat{G},G}$, so the expectation is that their simultaneous use would result in the overall highest $r_{\hat{G},G}$ for predictions of C2 and C3. With the exception noted before, this is indeed what is observed among the 18 cases. However, if we disregard CAP_TP entirely and use only C1S and C2S (nTP = 173) to predict GY and DON in C3, the resulting $r_{\hat{G},G}$ are the highest observed in the study (Table 1.4). To eliminate the possibility that nTP = 173 is an ideal population size for predicting GY and DON we compared the $r_{\hat{G},G}$ rom the TP noted above to n = 173 random and optimized subsets drawn from a TP consisting of CAP_TP, C1S, and C2S; the $r_{\hat{G},G}$ of optimized subsets were not significantly different than $r_{\hat{G},G}$ from C1S and C2S, but they were numerically lower (Table 1.4).

Application to breeding

The breeder's equation relates the rate of gain to three factors that can all be positively manipulated through the use of GS. Here we focused on increasing $r_{\hat{G},G}$ via three strategies that used existing breeding program data and freely available optimization algorithms implemented in R (R Core Team, 2015). Such resources may not be available to all breeding programs and some investment into genotyping breeding lines, or perhaps the initiation of field evaluation of selected breeding lines may be required.

Implementing the strategy of updating the TP with data collected on lines from previous breeding cycles will likely benefit the $r_{\hat{G},G}$ of traits that can most readily be measured in the field, in which case the directly preceding breeding cycle can be included in the TP for the subsequent cycle. When making predictions for traits that take more

time to measure, the breeding lines that can be added to the TP will be, at best, one cycle removed from the selection candidates. For such traits it may be possible to improve $r_{\hat{G},G}$ by including a correlated trait (Yia and Jannink, 2012), from the directly preceding breeding cycle, that can readily be measured in the field.

We observed significant improvements of $r_{\hat{G},G}$ for GY under RKHS compared to GBLUP. Since relationships modeled by $\mathbf{K}_N$ in RKHS are not assumed to be linear, thus perhaps modeling non-additive effects (Heslot et al., 2012), RKHS could be used to identify selection candidates to be advance for eventual varietal release whereas GBLUP could be reserved for predicting elite parent candidates. This is not to suggest that RKHS should be used in this capacity for all traits, as it did not result in improved $r_{\hat{G},G}$ for DON. However it highlights an opportunity to expand the use of GS in the barley breeding program beyond early generation parent selection.

We cannot prescribe a single protocol for updating and optimizing the TP. In this study G_mean was the overall best optimization procedure for GY followed by STPGA, but for DON, STPGA was the best procedure, followed by StratCDmean and G_mean. However, the significant two- and three-way interactions observed between the main effects of the optimization procedure, TP, VP, and prediction model (Table 1.2) present an obstacle for identifying and suggesting an optimization procedure to use. Ultimately, if GS is to be used successfully a breeder must invest the time into understanding how their population responds to different optimization strategies and prediction models.

Table 1.1: Changes in prediction accuracies of each breeding cycle resulting from updating the TP. Each column specifies a TP used to predict the breeding cycle specified by the VP column. Confidence intervals of $r_{\hat{G},G}$ were calculated from 2,000 bootstrapped samples.

	VP[†]	ORIG_TP	CAP_TP +				
			ø[‡]	C1S	C1	C1S + C2S	C1 + C2
GY	C1	0.124 (-0.05, 0.28)	0.346 (0.15, 0.51)	--	--	--	--
	C2	0.264 (0.02, 0.46)	0.284 (0.11, 0.45)	0.424 (0.29, 0.55)	0.362 (0.22, 0.51)	--	--
	C3	0.415 (0.26, 0.55)	0.299 (0.10, 0.48)	0.400 (0.21, 0.56)	0.321 (0.15, 0.49)	0.535 (0.36, 0.69)	0.477 (0.31, 0.62)
DON	C1	0.288 (0.02, 0.50)	0.723 (0.61, 0.81)	--	--	--	--
	C2	0.194 (0.01, 0.37)	0.376 (0.20, 0.53)	0.380 (0.21, 0.53)	0.375 (0.21, 0.53)	--	--
	C3	0.182 (-0.01, 0.39)	0.121 (-0.11, 0.33)	0.115 (-0.09, 0.32)	0.097 (-0.11, 0.303)	0.237 (0.02, 0.45)	0.197 (0.0, 0.39)

[†] VP, validation population; C1, all cycle 1 lines; C2, all cycle 2 lines; C3, all cycle 3 lines

[‡] The following data from BREED_TP were added to CAP_TP: ø, no addition; C1S; cycle 1 selected lines; C1S + C2S; C1S and cycle 2 selected lines

Table 1.2: Significance of the variance explained by some factors manipulated in this study, including the prediction model and TP optimization procedure (Opt) used, the breeding cycle being predicted (VP), as well as the two- and three- way interactions among factors.

GY				DON			
Term[†]	**DF**	**F**[‡]	**Pr(>F)**[§]	**Term**[†]	**DF**	**F**[‡]	**Pr(>F)**[§]
Model				**Model**			
RKHS[A]	1	6.58	*	RKHS[A]	1	0.06	*NS*
GBLUP[B]				GBLUP[A]			
VP[¶]				**VP**[¶]			
C1[A]	2	1.42	*NS*	C1[A]	2	710.63	***
C2[A]				C2[B]			
C3[A]				C3[C]			
Opt				**Opt**			
G_mean[A]				STPGA[A]			
STPGA[B]				StratCDmean[B]			
Rand[C]	4	20.91	***	G_mean[B]	4	10.40	***
SS[D]				SS[C]			
StratCDmean[E]				Rand[C]			
Model • TP	4	4.58	**	**Model • TP**	4	1.10	***
Model • VP	2	11.30	***	**Model • VP**	2	1.98	*NS*
Model • Opt	4	0.73	*NS*	**Model • Opt**	4	0.44	*NS*
TP • Opt	16	36.15	***	**TP • Opt**	16	19.42	***
VP • Opt	8	78.77	***	**VP • Opt**	8	10.11	***
Model • TP • Opt	16	5.33	***	**Model • TP • Opt**	16	1.20	*NS*
Model • VP • Opt	8	7.87	***	**Model • VP • Opt**	8	3.00	**

[†] Model term fit in the ANOVA, including interactions denoted with an '•'. Levels of a term with different superscript letters indicates they are significantly different according to Tukey's HSD at $\alpha = 0.05$.

[‡] *F*-statistic for the model term

[§] *P*-value supporting the alternative hypothesis that the model term explains significant variation in the ANOVA model. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; *NS*, not significant

[¶] VP, validation population; C1, all cycle 1 lines; C2, all cycle 2 lines; C3, all cycle 3 lines

Table 1.3: The performance of an optimization relative to a random TP subset of size *n* across nine cases (i.e. a particular TP-VP combination). Actual values of $r_{\hat{G},G}$ summarized herein are presented in Supplemental Table 2.2.

Optimization	**Grain Yield**[†]		**DON**	
Procedure	> Rand[‡]	Best[§]	> Rand[‡]	Best[§]
SS	0	0	7	0
G_mean	8	4	4	3
StratCDmean	3	2	6	4
STPGA	4	2	5	2

[†] n = 500 for grain yield (kg ha^{-1}); n = 300 for DON (ppm)

[‡] > Rand; number of the nine cases the optimization procedure results in a higher $r_{\hat{G},G}$ value than a random TP subset

[§] Best, the number of cases the TP optimization procedure resulted in the overall highest $r_{\hat{G},G}$

Table 1.4: Prediction accuracies of C3 using three TP subsets of size $n = 173$.

		Subsets of CAP_TP + C1S + C2S	
Trait	C1S + C2S[†]	Random[‡]	Optimized[§]
GY	$0.562^{A¶}$	0.296^{B}	0.524^{A} (G_mean)
DON	0.473^{A}	0.221^{B}	0.409^{A} (STPGA)

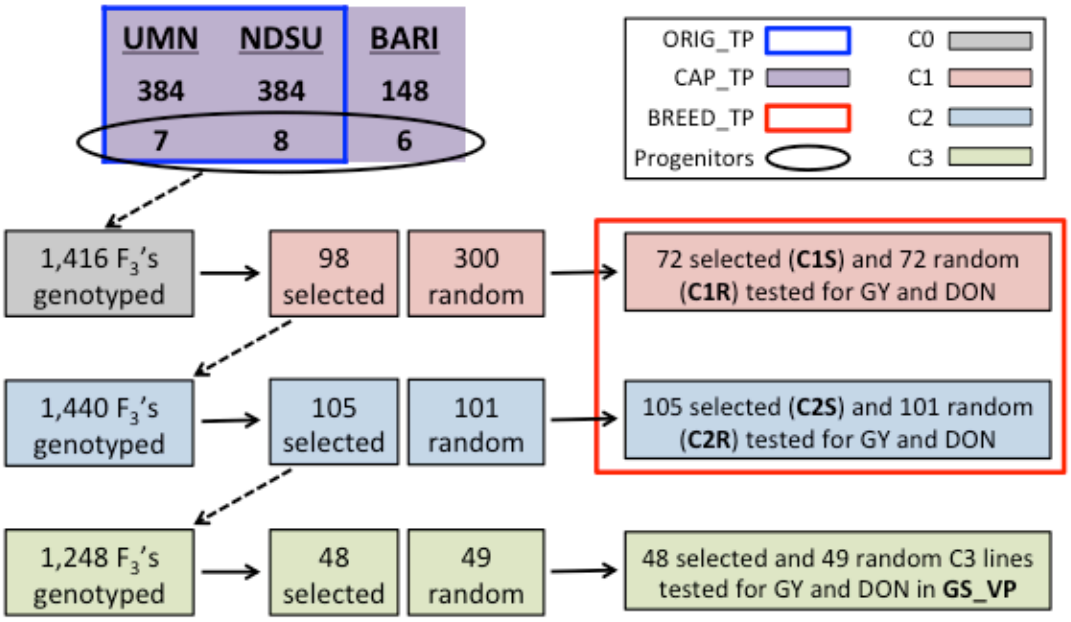
[†] C1S + C2S; cycle 1 selected lines plus cycle 2 selected lines, total 173 lines

[‡] Subset of 173 lines randomly chosen from the CAP_TP dataset plus C1S and C2S

[§] Subset of 173 lines chosen by G_mean (for GY) and STPGA (for DON) from the CAP_TP dataset plus C1S and C2S

[¶] Across rows, different superscript letters denote significant differences between the values according to Tukey's HSD at $\alpha = 0.05$.

Figure 1.1: Development of the University of Minnesota barley breeding program's composite GS population. The legend highlights datasets used in this study, including ORIG_TP, CAP_TP, and BREED_TP. Arrows with dotted lines indicate that intermating among lines occurred whereas solid lines indicate that lines were advanced a generation without intermating. C0 – C3, cycle 1 – cycle 3 lines; C1S, cycle 1 selected lines; C2S, cycle 2 lines.



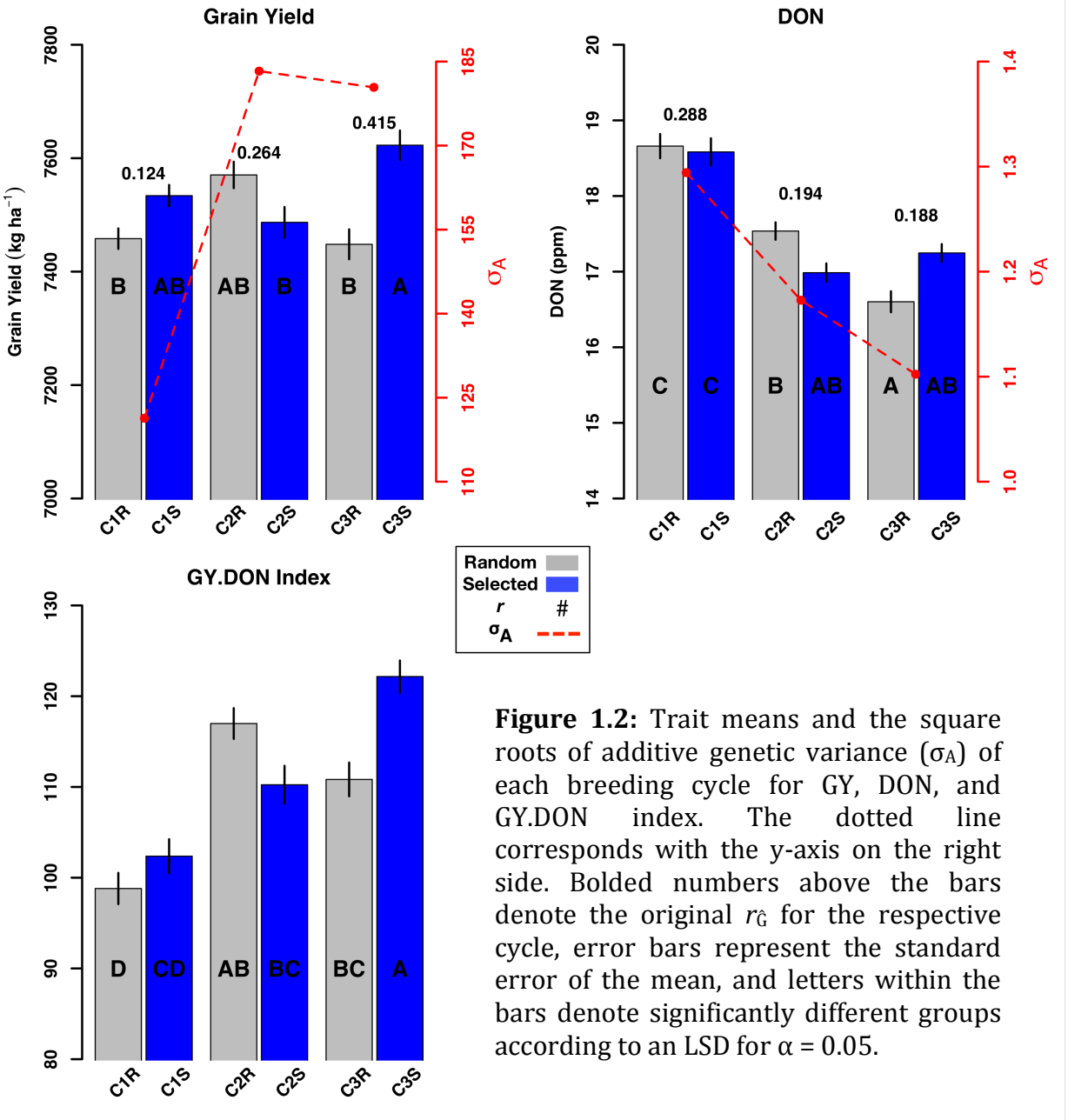
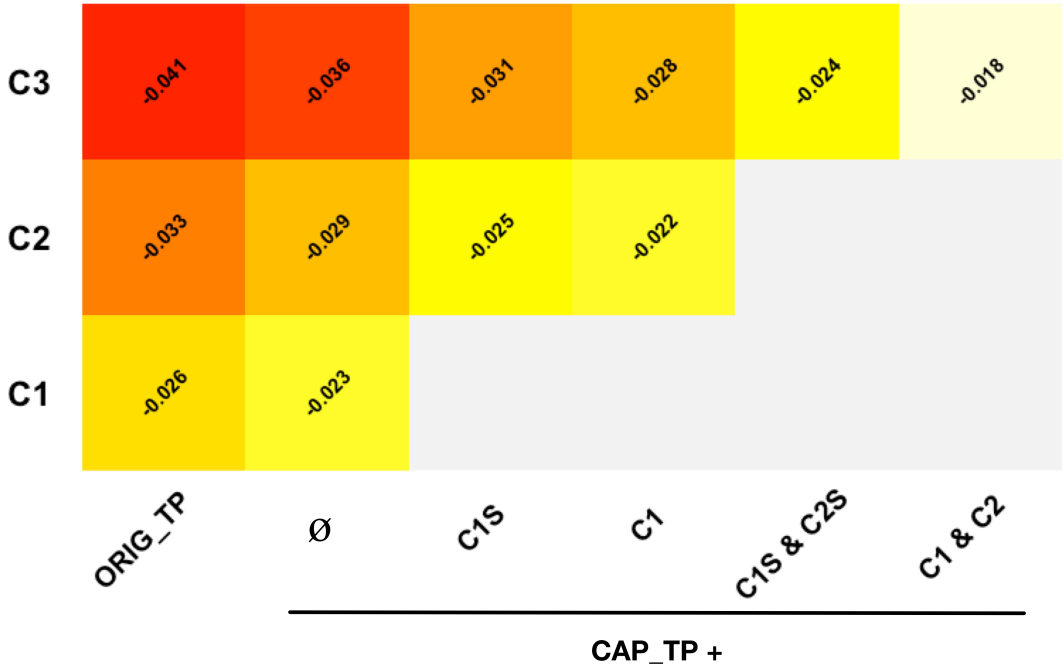


Figure 1.2: Trait means and the square roots of additive genetic variance (σ_A) of each breeding cycle for GY, DON, and GY.DON index. The dotted line corresponds with the y-axis on the right side. Bolded numbers above the bars denote the original r_G for the respective cycle, error bars represent the standard error of the mean, and letters within the bars denote significantly different groups according to an LSD for $\alpha = 0.05$.

Figure 1.3: Average realized genetic relationships (K_{ij}) for each TP-VP combination considered. Darker shades qualitatively express the exact K_{ij} values that appear in the center of each box. Axes are labeled as in Table 1.1.



Supplemental Methods 1

Calculating genetic correlation between GY and DON

For estimating additive genetic correlations (r_A) it would be ideal to have the same field plot (i.e. observation) evaluated for all traits of interest. However, quality GY data cannot be collected from disease nurseries, and likewise quality DON data is difficult to collect from yield trials. Therefore we used three yield trials and three disease nurseries from GS_VP that were evaluated in the same location in the same year to form the dataset GS_r_A.

The following multivariate mixed model was fit:

$$\mathbf{Y} = \mathbf{bX} + \mathbf{SZ} + \mathbf{E},$$

where **Y** is a matrix with the first and second rows being the observations for GY and DON in the GS_r_A dataset, respectively, **X** is a matrix where the three rows represent the fixed main effect of each of the three trials included in GS_r_A, **b** is a vector of the fixed effect coefficients, and **Z** is the random effect design matrix with each row relating to a unique line in GS_r_A. **S** and **E** are matrices modeling the covariance structures of the random effects and random error terms as follows:

$$\text{Cov}(\mathbf{S}) = \mathbf{G}_N \otimes \begin{bmatrix} \sigma^2_{g_1} & \phi_{12} \\ \phi_{12} & \sigma^2_{g_2} \end{bmatrix} \text{ and } \text{Cov}(\mathbf{E}) = \mathbf{I}_{nGS\_r_A} \otimes \begin{bmatrix} \sigma^2_{e_1} & 0 \\ 0 & \sigma^2_{e_2} \end{bmatrix},$$

where $\mathbf{G}_N$ is the N x N additive relationship matrix as described above, $\sigma^2_{g_1}$ and $\sigma^2_{g_2}$ are the (genetic) variance components for the first and second trait, respectively, ϕ_{12} is the (co)variance component between the traits, and finally $\sigma^2_{e_1}$ and $\sigma^2_{e_2}$ are the (residual error)

variance components for the first and second trait, respectively. After solving the mixed model with package 'EMMREML' (Akdemir and Godfrey, 2015) the additive genetic correlation between GY and DON was calculated as:

$$r_{\mathrm{A}} = \phi_{12} \Big/ \sqrt{\sigma_{g_1}^2 * \sigma_{g_2}^2}\,.$$

Supplemental Table 1.1: Details of field trials data of BREED_TP and GS_VP.

Dataset[†]	Trait	Location	Year	Design[‡]	nReps[§]	Spatial Correction
breed_TP – C1S	GY	Crookston, MN	2011	RCBD	2	AR1
breed_TP – C1S	GY	Morris, MN	2011	RCBD	2	AR1
breed_TP – C1S	GY	St. Paul, MN	2011	RCBD	2	AR1
breed_TP – C1R	GY	Crookston, MN	2011	AD	1	MA
breed_TP – C1R	GY	Nesson Valley, ND	2011	AD	1	MA
breed_TP – C1R	GY	Saint Paul, MN	2011	AD	1	MA
breed_TP – C2	GY	Crookston, MN	2013	AD	1	MA
breed_TP – C2	GY	Saint Paul, MN	2013	AD	1	MA
breed_TP – C2	GY	Morris, MN	2013	AD	1	MA
breed_TP – C1S	DON	Crookston, MN	2011	RCBD	2	--
breed_TP – C1S	DON	Saint Paul, MN	2011	RCBD	2	--
breed_TP – C1R	DON	Crookston, MN	2011	RCBD	2	MA
breed_TP – C1R	DON	Langdon, ND	2011	RCBD	2	MA
breed_TP – C1R	DON	Osnabrock, ND	2011	RCBD	2	MA
breed_TP – C1R	DON	Saint Paul, MN	2011	RCBD	2	MA
breed_TP – C2	DON	Saint Paul, MN	2013	AD	1	MA
breed_TP – C2	DON	Crookston, MN	2013	AD	1	MA
GS_VP	GY	Morris, MN	2014	AD	1	MA
GS_VP	GY	Crookston, MN	2014	AD	1	MA
GS_VP	GY	Saint Paul, MN	2014	AD	1	MA
GS_VP	GY	Morris, MN	2015	AD	1	MA
GS_VP	GY	Crookston, MN	2015	AD	1	MA
GS_VP	DON	Crookston, MN	2013	AD	1	AR1
GS_VP	DON	Saint Paul, MN	2013	AD	1	AR1
GS_VP	DON	Crookston, MN	2014	AD	1	AR1
GS_VP	DON	Saint Paul, MN	2014	AD	1	AR1
GS_VP	DON	Crookston, MN	2015	AD	1	AR1

[†] C1S, cycle 1 selected lines; C1R, cycle 1 random lines; C2, all cycle 2 lines

[‡] Experimental design. RCBD, randomized complete block design; AD, augmented design

[§] Subset of 173 lines chosen by G_mean (for GY) and STPGA (for DON) from the CAP_TP dataset plus C1S and C2S

[¶] Method used to account for spatially related environmental error, see '*Spatial corrections*' section in the materials and methods. MA, moving average covariate; AR1, correlation structure was fit according to first order autoregressive procedure. An '--' indicates that row column coordinates for plots were not available

Supplemental Table 1.2 (part 1): Prediction accuracies of GY and DON for nine TP-VP cases using the whole TP (All), a random TP subset of *n* lines (Rand), and with the implementation of four optimization procedures; stratified sampling (SS), G_mean, stratified CDmean (StratCD), and STPGA.

	TP†	VP†	Model‡	All	Rand	SS	G_mean	StratCD	STPGA
GY¶	CAP_TP	C1	GBLUP	0.346	**0.358**^{A*§}	0.356^{A*}	0.341^{A*}	0.192^{B*}	0.313^{A}
			RKHS	0.304	0.302^{A*}	**0.311**^{A*}	0.285^{A*}	0.128^{B*}	0.296^{A}
	CAP_TP	C2	GBLUP	0.284	0.214^{B*}	0.198^{BC*}	**0.293**^{A*}	0.135^{C*}	0.212^{B*}
			RKHS	**0.359**	0.294^{B*}	0.291^{B*}	0.343^{A*}	0.225^{C*}	0.297^{B*}
	CAP_TP + C1S	C2	GBLUP	0.424	0.362^{C*}	0.321^{CD*}	0.437^{B*}	**0.519**^{A*}	0.281^{D*}
			RKHS	0.593	0.486^{BC*}	0.477^{C*}	0.605^{A*}	**0.627**^{A*}	0.536^{B*}
	CAP_TP + C1	C2	GBLUP	0.364	0.311^{C*}	0.300^{C*}	0.398^{B*}	**0.454**^{A*}	0.239^{C*}
			RKHS	0.535	0.444^{BC*}	0.424^{C*}	**0.550**^{A*}	0.529^{A*}	0.494^{B*}
	CAP_TP	C3	GBLUP	0.299	0.232^{B*}	0.175^{B*}	**0.353**^{A*}	-0.172^{C*}	0.190^{B*}
			RKHS	**0.464**	0.361^{B*}	0.313^{BC*}	0.423^{A*}	-0.002^{D*}	0.289^{C*}
	CAP_TP + C1S	C3	GBLUP	0.400	0.276^{B*}	0.275^{B*}	**0.459**^{A*}	0.156^{C*}	0.422^{A*}
			RKHS	0.605	0.480^{C*}	0.474^{C*}	0.564^{B*}	0.490^{C*}	**0.619**^{A*}
	CAP_TP + C1	C3	GBLUP	0.307	0.213^{BC*}	0.182^{CD*}	0.264^{B*}	0.141^{D*}	**0.326**^{A*}
			RKHS	0.461	0.344^{BC*}	0.330^{BC*}	0.362^{B*}	0.289^{C*}	**0.524**^{A*}
	CAP_TP + C1S + C2S	C3	GBLUP	**0.535**	0.404^{B*}	0.387^{B*}	0.498^{A*}	0.202^{C*}	0.481^{A*}
			RKHS	**0.615**	0.545^{AB*}	0.535^{AB*}	0.562^{A*}	0.494^{B*}	0.567^{A*}
	CAP_TP + C1 + C2	C3	GBLUP	0.477	0.238^{C*}	0.252^{C*}	0.435^{B*}	0.019^{D*}	**0.512**^{A}
			RKHS	**0.594**	0.378^{C*}	0.376^{C*}	0.480^{B*}	0.258^{D*}	0.538^{A}

† C1, all cycle 1 lines; C1S, cycle 1 selected lines; C2, all cycle 2 lines; C2S, cycle 2 selected lines

‡ Significant differences between $r_{\hat{G},G}$ from GBLUP and RKHS for a given TP-VP combination are denoted by an * (p-value < 0.05)

§ Bolded values in a specific row highlight the highest $r_{\hat{G},G}$ for the respective TP-VP-Model combination

¶ For grain yield (kg ha^{-1}) the size of the TP subset was n = 500

Supplemental Table 1.2 (part 2)

	TP[†]	VP[†]	Model[‡]	All	Rand	SS	G_mean	StratCD	STPGA
DON[¶]	CAP_TP	C1	GBLUP	0.723	0.692^{B}	0.711^{B}	0.638^{C}	**0.767**^{A§}	0.697^{B}
			RKHS	0.727	0.709^{BC}	0.730^{AB}	0.633^{D}	**0.756**^{A}	0.696^{C}
	CAP_TP	C2	GBLUP	0.376	0.339^{BC}	0.355^{BC}	0.300^{C}*	**0.426**^{A}*	0.392^{AB}*
			RKHS	0.389	0.362^{B}	0.374^{B}	0.367^{B}*	0.374^{B}*	**0.434**^{A}*
	CAP_TP + C1sel	C2	GBLUP	0.380	0.361^{BC}	0.392^{B}	0.408^{AB}*	0.330^{C}	**0.442**^{A}
			RKHS	0.392	0.384^{BC}	0.406^{ABC}	**0.447**^{A}*	0.365^{C}	0.416^{AB}
	CAP_TP + C1	C2	GBLUP	0.380	0.368^{B}	0.364^{B}	**0.449**^{A}	0.391^{B}	0.368^{B}
			RKHS	0.392	0.392^{B}	0.385^{B}	**0.460**^{A}	0.411^{B}	0.396^{B}
	CAP_TP	C3	GBLUP	0.121	0.137^{B}*	0.136^{B}*	0.025^{C}	0.133^{B}*	**0.218**^{A}*
			RKHS	0.080	0.088^{B}*	0.090^{B}*	0.035^{B}	0.047^{B}*	**0.167**^{A}*
	CAP_TP + C1_sel	C3	GBLUP	0.115	0.140^{B}	0.164^{AB}	0.048^{C}*	**0.226**^{A}*	0.146^{B}
			RKHS	0.098	0.109^{B}	0.125^{AB}	0.102^{B}*	0.131^{AB}*	**0.181**^{A}
	CAP_TP + C1	C3	GBLUP	0.097	0.163^{B}	0.187^{AB}	0.168^{B}	**0.236**^{A}	0.165^{B}
			RKHS	0.096	0.131^{C}	0.143^{BC}	0.178^{ABC}	0.189^{AB}	**0.206**^{A}
	CAP_TP + C1_sel + C2_sel	C3	GBLUP	0.237	0.223^{C}	0.234^{BC}	**0.407**^{A}*	0.292^{B}	0.366^{A}*
			RKHS	0.347	0.248^{B}	0.252^{B}	**0.454**^{A}*	0.284^{B}	0.443^{A}*
	CAP_TP + C1 + C2	C3	GBLUP	0.197	0.265^{BC}	0.217^{C}	**0.324**^{A}	0.254^{BC}	0.309^{AB}*
			RKHS	0.324	0.265^{B}	0.218^{B}	0.347^{A}	0.251^{B}	**0.366**^{A}*

[†] C1, all cycle 1 lines; C1S, cycle 1 selected lines; C2, all cycle 2 lines; C2S, cycle 2 selected lines

[‡] Significant differences between $r_{\hat{G},G}$ from GBLUP and RKHS for a given TP-VP combination are denoted by an * (p-value < 0.05)

[§] Bolded values in a specific row highlight the highest $r_{\hat{G},G,G}$ for the respective TP-VP-Model combination

[¶] For DON (ppm) the size of the TP subset was n = 300

Supplemental Table 1.3: Mean of GY, DON, and the GY.DON index in the breeding cycle subsets.

Cycle	Category	GY‡	DON	GY.DON Index
C1	Random	7458.17^{B}	18.66^{C}	98.81^{D}
C1	Selected	7534.00^{AB}	18.58^{C}	102.37^{CD}
C2	Random	7570.47^{AB}	17.54^{B}	116.99^{AB}
C2	Selected	7486.88^{B}	16.99A^{B}	110.24^{BC}
C3	Random	7448.15^{B}	16.60^{A}	110.83^{BC}
C3	Selected	7622.91^{A}	17.25^{AB}	122.17^{A}

† C1, cycle 1 lines; C2, cycle 2 lines; C3, cycle 3 lines

‡ Values for GY, DON, and GY.DON index with different superscripts letters are significantly different according to the least significant difference (LSD) for $\alpha = 0.05$.

Chapter 2: Exploiting genetic diversity from wild barley introgression lines to improve grain yield of an elite malting barley variety

Introduction

Genetic variation plays an active and proactive role in plant breeding. Active in driving the improvement of traits important to cultivation through selective breeding, and proactive in that it ensures plasticity of the germplasm pool to respond to future needs such as changing environmental conditions (biotic and abiotic) and end-use needs (Martin et al., 1991; Tanksley and Nelson, 1996; Tanksley and McCouch, 1997; Stuthman, 2002; Holland, 2004; Gutiérrez et al., 2009). Although genetic variability is expected to decrease over cycles of selection within a closed breeding population, novel epistasis resulting from recombination between tightly linked genetic loci or interactions with *de novo* variation arising from mutation, intragenic recombination, transposable elements, among others may contribute more to genetic variation than traditionally considered under the standard additive genetic model (Rasmusson and Phillips, 1997; Jannink, 2003). It is unlikely that this "passive" generation of genetic variation is sufficient to respond to rapidly changing breeding targets and to ensure the long term breeding progress needed to meet the challenges of a growing population. Therefore, breeders should intentionally manage the genetic variation within their program to drive the genetic gains necessary for meeting the grand challenges facing crop scientists (Zamir, 2001; Lauer et al., 2012).

Designing wide crosses to introgress novel genes and alleles from exotic

germplasm is one strategy breeders have used to create or increase genetic variability in elite breeding populations. This endeavor has often been coupled with the use of genetic markers to assist in the selection of favorable alleles at quantitative trait loci (QTL). While marker-assisted selection (MAS) of QTL has been successful for some qualitative, relatively simply inherited traits such as *Sub1* in rice for submergence tolerance (*Oryza sativa*; Xu and Mackill, 1996; Xu et al., 2006; Septiningsih et al., 2009), *Alp* in barley for tolerance to aluminum toxicity (Soto-Cerda et al., 2012), and *Fhb1* in wheat for resistance to Fusarium head blight (*Triticum aestivum*; Bai et al., 1999; Waldron et al., 1999; Anderson, 2007), it has not been widely successful for improving quantitative traits (Tanksley and Nelson, 1996; Bernardo, 2008; Bernardo, 2016).

Efforts to improve quantitative traits, rather, focus on narrow crosses between elite, highly related germplasm within the breeding population (Thomas et al., 1995; Rasmusson and Phillips, 1997; Condon et al., 2008). This 'best-by-best' breeding strategy is typical in malting barley breeding where elite lines are continuously recycled to be used as parents to maintain industry-defined malting profiles (Martin et al., 1991; Horsley et al., 1995; Peel and Rasmusson, 2001; Condon et al., 2008). Marker-assisted selection of QTL for quantitative traits is especially impractical in such crosses as there will be few genetic polymorphisms segregating, thus low genetic variability and mapping power (Lander and Botstein, 1989; Tanksley and Nelson, 1996).

Wide crosses of elite breeding lines to exotic donor sources such as landraces, wild relatives, or even elite varieties from other breeding programs could address the limited genetic variability within elite breeding populations (Zamir, 2001; Holland, 2004;

Steffenson et al., 2007; Warschefsky et al., 2014). The paradox, however, is that wide crosses are not routinely made in breeding programs because of linkage drag, which can lead to a reduction in the average population performance. Linkage drag, specifically, is the unintentional introgression of chromosomal segments neighboring the QTL in the donor parent that can be deleterious *per se*, result in unfavorable recurrent-by-donor epistatic interactions, and/or disrupt favorable epistatic interactions assembled in the elite recurrent parent over generations of breeding (Tanksley and Nelson, 1996; Rasmusson and Phillips, 1997; Ho et al., 2002). Regardless, in order to tap into the genetic diversity present in germplasm collections it will be necessary to make such wide crosses.

When using wide crosses to introgress exotic donor alleles into elite backgrounds breeders have often used additional backcrossing generations to recover the background of the recurrent parent and address the deleterious effects of linkage drag. Most notably Tanksley and Nelson (1996) proposed an "advanced backcross QTL analysis" (AB-QTL) strategy to join the process of QTL discovery and germplasm improvement for quantitative traits. The generations of backcrossing result in smaller introgressions from the donor parents, reducing linkage drag (Tanksley and Nelson, 1996; Ho et al., 2002). The AB-QTL strategy has been used to successfully map QTL for quantitative traits in cultivated tomato (*Solanum lycopersicum*; Fulton et al., 1997; Bernacchi et al., 1998), barley (*Hordeum vulgare*; Pillen et al., 2003; Von Korff et al., 2006; Yun et al., 2006; 2008; Saal et al., 2011; Sayed et al., 2012; Eshghi et al., 2013), rice (*Oryza sativa*; Xiao et al., 1998; Moncada et al., 2001; Brodandi et al., 2002; Wu et al., 2004; Nagata et al., 2015), and maize (*Zea mays*; Ho et al., 2002; Li et al., 2008). Additional examples are

presented in a recent review by Brozynska et al. (2015).

Pertinent to this study, however, are reports that demonstrate the use of advanced backcross-based strategies to improve quantitative traits of elite, cultivated germplasm through the introgression of exotic donor germplasm. For example, Xiao et al. (1998) found that exotic alleles from *Oryza rufipogon* backcrossed into a cultivated rice (*O. sativa*) background improved testcross performance of quantitative traits, including grains per plant and grain yield over the recurrent *O. sativa* parent crossed to the same tester. Another study conducted in maize showed that selection of QTL-alleles sourced from a different heterotic group during backcrossing resulted in lines whose testcross performance was greater than the recurrent parent's testcross performance as well as commercial hybrid check varieties (Ho et al., 2002). In barley, favorable introgressions from the wild progenitor (*H. vulgare* L. *ssp. spontaneum*) were identified in the background of an elite recurrent malting variety and introgression lines superior for 1000-grain weight were developed by marker-assisted selection (MAS) (Schmalenbach et al., 2009).

An alternative to using MAS to introgress specific chromosomal segments to increase genetic variation is to use genome-wide markers. Bernardo (2016) advocates the use of genome-wide selection, also known as genomic selection (GS) to identify the segregants to backcross in each generation. This strategy can mitigate the dilution effect of crossing to an exotic donor, with favorable QTL fixed at few loci, into an elite recurrent parent with favorable QTL fixed at many loci (Bernardo, 2016). Recently, genomic selection has been shown to be a viable selection methodology in plant breeding

by many empirical cross validation studies (reviewed in Nakaya and Isobe, 2012) and selection studies (Asoro et al., 2013; Combs and Bernardo, 2013b; Massman et al., 2013b; Rutkosko et al., 2015; Beyene et al., 2015; Tiede et al., *in press*). In particular, Combs and Bernardo (2013b) used GS to introgress exotic germplasm and demonstrated gains for a host of traits comprising a trait index, including grain yield, over five cycles of selection.

Barley researchers at the University of Minnesota have assembled collections of wild relatives (Steffenson et al., 2007) and developed structured populations to facilitate the movement of novel alleles into breeding germplasm (Yun et al., 2005). Of particular interest in this study is an advanced backcross nested association mapping population (AB-NAM) that was developed by making BC_2 lines by crossing 25 wild barley donors to a single recurrent parent, cv. Rasmusson (Nice et al., *in press*). Rasmusson (PI 658495; Smith et al., 2010) is an elite six-row malting variety that consistently tops the Minnesota variety trial list for grain yield (MAES 2015, https://www.maes.umn.edu/publications/field-crop-trials/2015). Since the original goal of the AB-NAM was to map QTL for grain yield, neither GS nor any selection was performed prior to each backcross (Nice et al., *in press*). Our interest is to determine whether superior individuals from the AB-NAM could be used in breeding to create progeny superior to Rasmusson for grain yield, a classically quantitative trait.

In this study, we use the barley AB-NAM panel (Nice et al*., in press*) and biparental breeding populations from crosses among the 50 highest yielding AB-NAM entries to: 1) identify and characterize recombinants of crosses made between wild barley

introgression lines that outperform Rasmusson, for grain yield, 2) compare three different strategies of targeting parent combinations, and 3) evaluate whether GS is effective in selecting for favorable exotic alleles introgressed into a predominantly elite background of the recurrent parent Rasmusson.

Materials and Methods

Overview

The University of Minnesota maintains a Wild Barley Diversity Collection (WBDC) (Steffenson et al., 2007) consisting of 318 wild barley accessions. Twenty-five WBDC accessions, chosen to represent the majority of the allelic diversity found in the entire WBDC, were backcrossed twice to the maternal parent Rasmusson. The resulting 796 BC_2-derived $F_{4:6}$ lines across 25 families comprise the AB-NAM (Nice et al., *submitted*). Rasmusson is a six-row malting barley chosen as the recurrent parent for its elite malting quality and superior grain yield (Smith et al., 2010), as attested by state variety trial data (MAES 2015). The 50 highest yielding AB-NAM lines, identified by grain yield trials conducted in 2012 and 2013, were selected as parent candidates and crossed in combinations determined by three strategies (described below) to ultimately generate 1,000 segregating F_3 progeny representing multiple crosses per strategy. Genomic selection was then used to predict the 100 highest yielding F_3 progeny per crossing strategy to be advanced to multi-environment yield trials. Further details of each step can be found in their respective sections below.

AB-NAM of Nice et al.

The development and genetic characterization of the AB-NAM is described in detail in Nice et al. (*in press*). Here we describe the phenotypic evaluation of the AB-NAM as it pertains to our study. Grain yield evaluation of the AB-NAM was conducted in Crookston and Saint Paul, MN in 2012, and in Crookston, MN and Fargo, ND in 2013. We excluded the Bozeman, MT location of Nice et al. (*in press*) because it is outside the traditional Midwest region of six-row malting barley production. The trials were grown as a modified augmented design based on Lin and Poushinsky's (1985) type II design; for the AB-NAM a primary check was placed at the center of each 3 x 5 block and two additional check varieties were grown in six randomly-assigned blocks in the trial. Plots consisted of two 1.0 m long rows seeded at 300 seeds m^{-2}.

The following mixed-effects linear model was fit using the 'lme4' (Bates et al., 2014) package in R (R Core Team, 2015):

$$y_{ijrc} = \mu + trial_j + line_i + row_{r(j)} + column_{c(j)} + (line * trial)_{ij} + e_{ijrc}, \quad (1.1)$$

residual error term for each observation. The *line* term was fit as a fixed effect whereas all other model terms were considered random. The adjusted means of *line* were used as the phenotypic values for the AB-NAM lines in subsequent analyses.

Crossing strategies and developing breeding populations

Specific pairwise cross combinations among the 50 highest yielding AB-NAM lines (i.e. parent candidates) were identified by one of three strategies. The first strategy prioritizes crosses among the highest yielding parent candidates while ensuring that

crosses are not made among full-siblings, and is therefore referred to as 'best-by-best' (BxB). The second strategy prioritizes crosses among the most genetically distant (GD) parents as measured by the proportion of non-matching single nucleotide polymorphism (SNP) markers (Nei, 1974). The third strategy, termed 'idealized progeny value' (IPV), is a novel parent selection strategy that utilizes genomewide markers to identify combinations of parents that, upon "ideal" recombination, would produce the highest performing progeny. The IPV of a cross between parent candidates *i* and *j* is calculated from *n* markers as:

$$IPV_{ij} = \sum_{m=1}^{n} max\ (i_m * u_m, j_m * u_m) \quad (2.2)$$

where u_m is the estimated effect of marker *m*, i_m is the genotype of individual *i* at marker *m*, and j_m is the genotype of individual *j* at marker *m*. For this study the effects of 384 SNP markers scored across the 50 parent candidates were estimated by ridge regression best linear unbiased prediction (RR-BLUP) (Endelman, 2011) using a training population comprised of all 798 WBIP lines. The 1,225 possible cross combinations between the 50 parent candidates were prioritized based on IPV.

Experimental population development

The 50 parent candidates were grown in a greenhouse in Saint Paul, MN in the fall of 2013 and crosses were made following the priorities set by each strategy. Ultimately, 25 crosses representing 10 high-priority crosses for each strategy were advanced by single-seed decent to the F_3 generation, at which point selection was imposed. Some crosses were identified as high priority by more than one strategy.

Between 115 and 119 F_3 selection candidates per cross, for a total of 2,979, were genotyped using genotyping-by-sequencing (GBS) (USDA-ARS, Raleigh, NC) along with 689 AB-NAM lines for which remnant DNA was available from the original genotyping of the AB-NAM population (Nice et al., *in press*). The grain yield and marker data from these 689 AB-NAM lines was used as the training population (TP) for genomic prediction.

The TP was used to predict genotypic values of the 2,979 F_3 selection candidate using RR-BLUP. Other statistical prediction models were tested including Bayes A, Bayes B, Bayes Cπ, Bayesian Lasso, and Bayesian ridge regression using R package 'BGLR' (de los Campos and Pérez, 2015), but RR-BLUP resulted in the highest cross-validation accuracy or $r = 0.586$ across 500 iterations of the 60/40 sampling strategy (Mohammadi et al., 2015). The selection candidates with the highest predicted genotypic values within each strategy were advanced as $F_{3:4}$ head rows at an off-season nursery near Christchurch, NZ to increase seed.

The goal was to develop an experimental population with each selection strategy represented by its top 100 selection candidates. However, because of insufficient seed increase for some candidates the following was achieved: BxB was represented by 88 lines from 9 families, which ultimately represent 8 AB-NAM families and 7 wild barley parents; GD was represented by 80 lines from 9 families, which represent 10 AB-NAM families and 8 wild barley parents; IPV was represented by 90 selected lines encompassing 9 families, which represent 10 AB-NAM families and 8 wild barley parents. Again, some WIGS families represent more than one selection strategy

(Supplemental Table 2.1). In addition, 30 randomly selected candidates from each strategy, and the 30 selection candidates with the lowest predicted genotypic values were increased and subsequently bulked, by weight, to create 'mean' and 'low' bulks for each strategy.

An experimental population, hereafter called WIGS (wild introgression genomic selection) comprising 202 selected lines was grown in five trials during the 2015 growing season along with the mean and low bulks, each replicated five times per trial, and four check varieties, including Rasmusson. Plot sizes within each location were as follows: 0.61m x 2.4m in Saint Paul, MN; 1.2m x 2.4m in Crookston, MN; 1.5m x 2.4m in Buffalo County, WI; 1.5m x 3.0m in Roseau, MN; and 1.5m x 4.5m in Stephen, MN. Standard agronomic practices and a seeding rate of 1.2 million seeds per acre were used across trials. All trials were mechanically planted and harvested. Frequent precipitation and strong winds across MN and WI early in the 2015 growing season caused plots to lodge, so lodging was rated for each trial as a the proportion of the plot leaning in excess of 45 degrees from vertical.

WIGS was grown in an incomplete block design augmented with replicated checks. Briefly, a rectangular field is divided into incomplete blocks, each comprised of an equal number of subplots (just plots hereafter). A primary check variety is randomly assigned to a plot within each incomplete block. A set of secondary checks are then equitably replicated and assigned to plots in a semi-randomized manner that ensures the set of secondary checks is equitably distributed across incomplete blocks. In our case, Rasmusson was the primary check and the secondary checks were cultivars Lacey (PI

613603; Rasmusson et al., 2001), Quest (PI 663183; Smith et al., 2013), and Tradition (PI 612442; Busch Agricultural Resources, Inc.). Excess plots were filled with selected lines; the number and identity of replicated lines differed per trial. Similar to Eq. 2.1, the mixed-effect linear model fit to the multi-trial data can be generalized as:

$$y_{ijklrc} = \mu + trial_j + lodging_{rc(j)} + mv.cov_{rc(j)} + line_i + row_{r(j)} + column_{c(j)} + row.blk_{k(j)} + col.blk_{l(j)} + (line * trial)_{ij} + e_{ijklrc}, \quad (2.3)$$

where y_{ijklrc} is an observation of grain yield for line *i* located at row *r* and column *c*, in row-block *k* and column-block *l*, of trial *j*. The experiment-wide intercept is modeled by term μ, $trial_j$ denotes the fixed effect of the j^{th} trial, $lodging_{rc(j)}$ is a lodging covariate according the lodging score of the plot at row *r* and column *c* in trial *j*, $mv.cov_{rc(j)}$ is a moving average (Technow, 2015) covariate accounting for spatially-related error variance at the plot located at row *r* and column *c* in trial *j*, $entry_i$ is the fixed effect of the i^{th} experimental entry, $(line * trial)_{ij}$ is the random interaction of line *i* and trial *j*, and e_{ij} is the random error term for observation y_{ijk}. Lodging and moving average covariates and design effects, i.e. row, column, row-block, and column-block effects not significant in certain trials were removed to avoid over fitting the model.

Equation 2.3 indicates that a moving average covariate calculated using R package 'mvngGrAd' (Technow, 2015) modified according to Mohammadi et al. (2015) was used to account for spatial-related error in each trial. The model fitting this moving average covariate resulted in a better model fit, according to the Bayesian information criterion (Schwartz 1978), compared to a model fit with raw phenotypic or a model

including a two-dimensional separable autoregressive correlation structure (i.e. AR1⊗AR1; Cullis et al., 2006). The adjusted entry means from Eq. 2.3 were used as the phenotypic values of WIGs in subsequent analyses. The least significant difference at α = 0.05 ($LSD_{0.05}$) was calculated as $LSD_{0.05} = t(1 - 0.05/2, df_{error}) * ASED$, where *t* is the Student's *t* distribution for α = 0.05 and ASED is the average standard error of the differences for the adjusted means resulting from Eq. 3 (Vargas et al., 2013). Heritability on a line mean basis for grain yield was calculated as:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \bar{v}/2}, \quad (2.4)$$

where σ_g^2 is the genetic variance estimated from Eq. 2.3 when *line* is fit as a random effect and $\bar{v}$ is the mean pairwise variance for two fixed line effects (Cullis et al., 2006; Piepho and Möhring, 2007).

Marker datasets

The GBS markers were differentially filtered and parsed according to the needs of the analysis. First, perfectly correlated markers (i.e. Pearson's correlation coefficient is equal to 1 or -1) and those without two homozygous classes were removed, as were markers whose level of heterozygosity was greater than two standard deviations above the expectation, which equates to 6.1% in the AB-NAM (genotyped $BC_2F_{4:5}$ bulks) and 24% in the WIGS (genotyped single F_3 plants).To identify SNPs that are minimally segregating in the WIGS, markers with an MAF < 2.5% or missing more than 15% data were removed. To conduct GS we needed to identify high-quality GBS markers in

common between the AB-NAM TP and the 2,981 WIGS selection candidates. To that end, genetic markers missing more than 15% data or with an MAF < 3% in the TP, and TP entries missing more than 20% data were removed. Markers passing the filtering criteria in the TP, but either monomorphic or missing more than 15% data in the selection candidates were removed. Selection candidates missing more than 20% data were removed. Finally, to conduct a genome-wide association study (GWAS) markers with an MAF lower than 5% or missing more than 15% data across the 203 WIGS lines plus Rasmusson were removed. None of the WIGS lines nor Rasmusson were removed during the filtering process for the GWAS dataset. After completing all filtering steps, missing marker data in each dataset were imputed using the EM-algorithm (Endelman, 2011).

Genomic selection

As mentioned, GS was performed using ridge-regression BLUP (RR-BLUP) since it resulted in the highest cross-validation accuracy of $r = 0.586$ among all prediction models tested. After filtering, 680 AB-NAM lines scored with 6,904 markers remained to train the following ridge-regression model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \tag{2.5}$$

where **y** is a 680 x 1 vector of adjusted means for grain yield of the AB-NAM lines comprising the TP, **X** is the design matrix for the fixed effects represented in the vector **b**, which in this case are a 680 x 1 vector of ones and a scalar value representing the mean of the TP, respectively. **Z** is the design matrix for the random effects represented in vector **u**, which in the case of RR-BLUP is the matrix of marker genotypes coded as {1,

0, -1} to represent the homozygous, heterozygous, and the other homozygous genotypes and the solutions of **u** are the RR-BLUP estimates of each of the 6,904 marker effects with $\mathrm{var}(\mathbf{u}) \sim \mathcal{N}(0, \sigma_G^2)$, where σ_G^2 is the genetic variance. Finally, **e** is a vector of residual errors with $\mathrm{var}(\mathbf{e}) \sim \mathcal{N}(0, \sigma_E^2)$, where σ_E^2 is the error variance. For the original genomic predictions, the mixed model was solved using the 'mixed.solve' function of the 'rrBLUP' package in R (Endelman, 2011). All other mixed models were solved using the 'EMMREML' package (Akdeir and Godfrey, 2015).

Marker interpolation

Markers filtered for GS were also used to visualize and quantify exotic introgressions, and to identify haplotypes. These mapped markers were used to interpolate genotypes of pseudomarkers at 0.1 cM intervals, similar to the procedure of Nice et al. (*in press*), which were in turn used to visualize Figure 2.1 and estimate the proportion, frequency, and size of exotic introgressions within the WIGS lines and AB-NAM parents. Briefly, markers from the exotic donors were assigned a value of 1, markers from Rasmusson a value of 0, and heterozygotes a value of 0.5. The interpolated position of a pseudo marker *M* at the i^{th} interval between two mapped markers, A and B having values *A* and *B*, could then be calculated as:

$$\mathrm{M}_i = A\left(\frac{i-a}{b-a}\right) + B\left(\frac{b-i}{b-a}\right), \tag{2.6}$$

where *a*, the mapped position of A is always assumed to be the upstream locus and for simplicity assigned a value of zero, and *b* and *i* are positions relative to *a*. The proportion of exotic introgression was estimated as the summation of pseudo-markers for a

particular genotype, be it a WIGS line or AB-NAM parent, divided by the number of pseudomarkers. For estimating the frequency and size of exotic introgressions a threshold of 4.0 cM was allowed between non-zero interpolated markers before exotic introgressions were considered independent.

Association analysis

Two approaches to association analysis were taken. Each approach was performed using 3,971 markers available after filtering. The first approach used a unified mixed linear model (MLM) (Yu et al., 2006) implemented in the GAPIT software package (Lipka et al., 2012). Control of spurious associations arising from population structure comes from incorporating a random polygenic term whose covariance structure is defined by a marker-based kinship matrix (i.e. K model) among entries into the unified MLM (Zhang et al., 2010; Lipka et al., 2012; Segura et al., 2012). A false discovery rate (FDR) (Benjamini and Hochberg, 1995) threshold of 0.05 was used to identify significant associations resulting from the unified MLM model.

The second approach was to use the multi-locus mixed-model (MLMM) procedure of Segura et al. (2012). The MLMM approach expands the K model by including stepwise selection of markers to enter the model as cofactors to further control for inflated test statistics and false positives expected to arise in structured populations. The MLMM implements a forward selection step to incorporate marker cofactors into the linear mixed-model, followed by a backwards elimination step (Segura et al., 2012). In our study the inclusion or removal of a marker cofactor was based on whether its *p*-value

was less than the Bonferroni-adjusted significance level of $\alpha = {}^{0.05}/_{m} = 1.26\text{E}^{-5}$, as m = 3,971. The MLMM R script is available at the website provided by Segura et al. (2012).

To estimate additive marker effects Eq. 2.5 was first expanded to include **Ws**, where **W** is a design matrix relating observations in **y** to the random variable **s**, where $\text{var}(\mathbf{s}) \sim \mathcal{N}(0, \mathbf{K} \otimes \sigma_{\text{G}}^2)$ and **K** is a marker-based kinship matrix, as fit in the association analysis models, calculated according to the second method of VanRaden (2008). Single markers were then passed to **Z** (Eq. 2.5) to estimate **u**, with $\text{var}(\mathbf{u}) \sim \mathcal{N}(0, \mathbf{1} \cdot \sigma_{\text{G}}^2)$ where **1** is a scalar variable of 1. The mixed model to estimate marker effects was solved using the EMMREML package in R (Akdeir and Godfrey, 2015).

Haplotype characterization

Nice et al. (*in press*) found that linkage disequilibrium (LD) within the WBDC decayed to less than 0.20 within 0.37 cM on average. To be conservative we identify haplotypes just among the markers that map to the same cM bin as the significant SNP per marker-trait association according to POPSEQ map positions (Mascher et al., 2013). Spurious haplotypes could arise from genotyping errors so each haplotype must be observed more than once in the WIGS to be considered real. Additionally, haplotypes were confirmed in the AB-NAM and WBDC progenitors that were included in the re-genotyping effort using GBS. Pairwise marker LD was estimated using the 'genetics' package in R (Warnes et al., 2013; R Core Team, 2016) as the squared correlation

coefficient (r^2) and the standardized deviation from expected haplotype frequency (D') between marker genotypes.

Results

Top performing line

With an adjusted mean of 6357 kg ha^{-1}, the top performing experimental line was identified by GD and yielded 407 kg ha^{-1}, or 6.8% more than Rasmusson. This increase was not significant according to an $LSD_{0.05}$ = 781.41 kg ha^{-1}, which equates to a 13.1% difference from Rasmusson or 14.2% difference from the population mean. Nonetheless, a total of 11.4% of the experimental lines yielded higher than Rasmusson. In contrast, at 5950 kg ha^{-1} Rasmusson yielded significantly ($p < 0.05$) higher than 15.3% of the experimental lines. Rasmusson also yielded significantly higher than the check varieties Quest and Tradition, which is consistent with the MN variety trial results conducted over the last three years (MAES, 2015). To give additional perspective, the 25 AB-NAM parents represent the top 3.1% of the AB-NAM, and the average of these parents' adjusted means is 2.9% higher than the adjusted mean of Rasmusson across AB-NAM trials considered in this study. In comparison, the average of the top six WIGS lines, or 3.0% of the population, is 4.8% higher than Rasmusson. The average grain yield of Rasmusson in the AB-NAM trials, however, is 836.9 kg ha^{-1} lower than the check variety Tradition, which is not consistent with the MN variety trial (MAES, 2015), suggesting

that any comparison to Rasmusson should be considered cautiously. Tradition yielded 202.5 kg ha^{-1} higher than the highest yielding AB-NAM line.

Of the 23 lines that performed numerically better than Rasmusson, four lines yielded higher than Rasmusson in four of the five trials, eight lines yielded higher in three of the five trials, and the remaining 11 lines performed better than Rasmusson in just one or two of the individual trials. The top nine lines, which all yielded higher than Rasmusson in three or four individual trials, are the focus of Table 2.1.

Performance by crossing strategy

The performance of the three crossing strategies was assessed in multiple ways. First, crosses targeted by BxB resulted in the highest yielding progeny, on average, followed by GD and then IPV. Best-by-best and GD were not significantly different from each other, but both were significantly ($p < 0.05$) better than IPV (Table 2.2). Similarly, the second criterion considers the average performances of the crosses (i.e. families) targeted by each crossing strategy. The top performing family was identified by GD, and the second and third highest performing families were identified by IPV (Supplemental Table 2.2); those three families were each represented by four or fewer individuals. Of the families represented by at least ten individuals, the highest performing family was identified by GD and BxB, the second by GD alone, the third by BxB, and the fourth by IPV.

The bulk representing the mean of BxB yielded the highest followed by the bulk representing the mean of IPV, and then the mean bulk for GD, although none of these

differences were significant (Table 2.2). Finally, of the top nine lines that performed better than Rasmusson in at least three of the five trials, six were identified solely by GD, two solely by IPV, and one, albeit the top performing line (3502-012; Table 2.1), was targeted by BxB and GD. In contrast, eight of the nine worst performing WIGS lines (Figure 2.1) were identified by IPV, with the ninth worst line being identified by GD. If the scope of comparison is expanded, 20 of the 25 worst lines were identified by IPV, three by GD alone, and two by GD and BxB.

Accuracy of GS

The original progeny-based prediction accuracy, measured as the Pearson correlation (r) between the predicted genotypic value and adjusted means of all WIGS lines, was $r = -0.01$, with a 95% confidence interval calculated from 2,000 bootstrapped samples of -0.15 to 0.13. The quartiles of prediction accuracy among the 15 individual families of WIGS containing five or more lines are $Q_1 = -0.06$, $Q_2 = 0.05$, and $Q_3 = 0.15$. In hindsight, however, the effect of AB-NAM family was significant ($p < 0.05$) in an ANVOA and when we replace $\mathbf{X}$ in Eq. 2.5 with a design matrix accounting for family membership among AB-NAM lines, prediction accuracy of the WIGS population significantly increased to $r = 0.37$ (95% CI: 0.23 – 0.48). And again, the quartiles of prediction accuracy among the 15 families of WIGS containing five or more lines are $Q_1 = -0.20$, $Q_2 = 0.15$, and $Q_3 = 0.42$.

The performance of the low bulk was lower than the performance of the mean population bulk for all three strategies (Table 2.2). Although not significant, the

difference was greatest for IPV and least for BxB (Table 2.2). The mean of the WIGS lines selected by strategies BxB and GD were greater than their mean population bulks, but the population bulk for IPV was slightly higher than the mean of the WIGS lines selected by IPV (Table 2.2).

Marker-yield associations

A total of 27 significant marker-trait associations for grain yield, spanning five chromosomes and clustering into seven groups (i.e. putative QTL) were identified by the two association analysis procedures (Figure 2.1, Table 2.3). Twenty-four of those associations were identified by the unified MLM approach. One marker-yield association on chromosome 6H identified a genomic region for which the exotic marker had a positive effect on grain yield (Table 2.3). For the remaining 23 putative associations, the positive marker effect for grain yield was associated with the recurrent genotype.

Five marker-yield associations were found significant by the MLMM association analysis. Two of these associations overlapped with markers identified by the unified MLM approach, including the most significant SNP at 108.6 cM on chromosome 3H (Table 2.3). The non-overlapping associations are located in regions of the genome not identified by the unified MLM approach, including a marker on the short arm of 6H and two markers on chromosome 7H (Figure 2.1; Table 2.3). Furthermore, the non-overlapping marker-yield associations on 6H identified by MLMM has a positive marker effect associated with the exotic marker (Table 2.3).

The recurrent parent genotype was highly represented (83 – 100%) in the nine highest performing WIGS lines across the 27 significantly associated markers. The two markers on 6H, for which the exotic marker is favorable, were also largely represented by the recurrent parent genotype rather than the favorable exotic genotype (Table 2.3). In contrast, the recurrent parent was less represented (17-94%) in the nine poorest performing nine lines, and as a result the exotic genotype was better represented at the two QTL on 6H for which the exotic marker is favorable (Table 2.3).

Exotic donor representation

The GBS marker data was filtered to remove perfectly correlated markers or those missing more than 15% marker data across WIGS, resulting in 8,626 high quality marker, including monomorphic markers. After additional filtering to remove markers with an MAF below 1.5% across the 202 WIGS lines plus Rasmusson there remained 5,304 markers, suggesting that 61.5% of the genome is represented by a wild introgression (i.e. a polymorphic marker with Rasmusson) in at least three lines (203 * 0.015 = 3.03) in the WIGS population. Individual WIGS lines are, on average, composed of 10.76% wild-derived introgressions. The representation of wild alleles in the top nine lines is numerically lower ($p = 0.07$) at 8.66% (Table 2.1), while the average representation of wild alleles in the poorest performing nine lines, at 18.64%, is significantly higher ($p < 0.001$) than the top nine lines and the average across WIGS. The AB-NAM parents of WIGS that were genotyped with GBS had an average introgression frequency 12.96%, which comports well with the observed frequency across the whole AB-NAM of 12.3%

reported in Nice et al. (*in press*) and the expected frequency of 12.5% in a BC_2-derived population.

The number of exotic introgressions in WIGS lines ranged from 3 to 19, but averaged 9.7 introgressions per line and averaged 16.28 cM in length. An average of 9.7 introgressions per WIGS line is a significant decrease over the average of 16.38 introgressions per AB-NAM line that was found in parents of WIGS and genotyped with GBS. However, the average size of the introgressions in the AB-NAM lines was significantly smaller at 12.07 cM.

The top nine lines carried an average of 11 introgressions per line, which was not statistically different ($p = 0.17$) than the WIGS as a whole; however, the average introgression size among the WIGS lines was significantly smaller ($p = 0.04$) at 13.39 cM. Furthermore, the top nine lines carried significantly fewer ($p = 0.05$) exotic introgressions than the aggregate of their AB-NAM parents, which harbor 14.17 introgressions per line on average; however, the size of the introgressions remained virtually unchanged. Furthermore, the poorest performing nine WIGS lines carried on average 14.33 introgressions per line that were on average 21.66 cM, both of which were significant increases over the WIGS population as a whole and the top nine WIGS lines ($p < 0.01$). Finally, the AB-NAM parents of the poorest performing nine WIGS lines carried 17.83 introgressions, which is significantly more ($p = 0.04$) introgressions than the respective nine WIGS lines, but the average introgression size was significantly smaller ($p < 0.01$) in the AB-NAM parents.

Haplotype characterization

The regions visualized (Figure 2.2) correspond to SNPs identifying a QTL that was identified as the most significant marker-trait association by the unified MLM and the MLMM approaches, and another QTL identified by MLMM for which the wild allele was identified as the increasing allele. The first QTL was associated with a SNP (SNP_{QTL-1}) at position 6,843 on Morex contig 42,877 of the barley reference sequence (International Barley Genome Sequencing Consortium, 2012), which maps to 108.57 cM on chromosome 3H according to the POPSEQ map (Mascher et al., 2013). The second QTL was associated with a SNP (SNP_{QTL-2}) at position 10,722 on Morex contig 41,407, which is not included in the POPSEQ map, but would be positioned between Morex contigs 7,606 and 42,9678, which both map to 1.84 cM on chromosome 6H.

SNP_{QTL-1} consistently segregates between nucleotides T and C for the Rasmusson and non-Rasmusson allele, respectively. When considered in concert with two markers in high LD with SNP_{QTL-1} (Figure 2.2) that cluster within an 800 bp region on the adjacent Morex contig 43,416, which also maps to the 108.57 cM position, we identify three distinct exotic haplotypes (Figure 2.2). The mean of the Rasmusson haplotype at 5620.5 kg ha^{-1} was significantly ($\alpha = 0.05$) higher than the three exotic haplotype class means according to Tukey's honestly significant different test for multiple comparisons (Lane and Salkin, 2010). The three exotic haplotype class means ranged from 4751.0 to 4990.5 kg ha^{-1} and were not significantly different from one another (Figure 2.2).

Similarly, considering SNP_{QTL-2} together with two markers in high LD with SNP_{QTL-2} on adjacent Morex contig 7,606 results in five total haplotypes, including the Rasmusson haplotype, four of which would remain if SNP_{QTL-2} were ignored (Figure 2.2). In contrast to SNP_{QTL-1}, SNP_{QTL-2} does not discretely segregate between Rasmusson and non-Rasmusson haplotypes (Figure 2.2). Furthermore, the class mean of the exotic haplotype "ACC" at 5869.6 kg ha^{-1} was greater ($p = 0.09$) than the Rasmusson haplotype class mean at 5540.3 kg ha^{-1}, and is significantly greater than the two exotic haplotype class means represented by more than two observations (Figure 2.2).

Discussion

Until this point, exotic germplasm has largely been used as a source of genetic variation to improve traits conditioned by one or a small number of genes. This has been accomplished with genes that could be identified and selected upon solely by their phenotypic expression (i.e. qualitative traits), or with one or a few large-effect QTL uncovered through genetic mapping (Tanksley and McCouch, 1997) and introgressed using marker-assisted selection (Tanksley et al., 1996; Xu and Mackill, 1996; Bai et al., 1999; Waldron et al., 1999; VonKorff et al., 2004, 2005; Xu et al., 2006; Anderson, 2007; Soto-Cerda et al., 2012). Cultivated, or elite barley represents only a portion of the genetic variation present in the primary gene pool (Harlan and de Wet, 1971; Ellis et al., 2000; Kilian et al., 2006; Ren et al., 2013). While our study did not identify any experimental lines yielding significantly higher than Rasmusson, our results indicate that

exotic donors, which have historically been mined for single genes and large-effect QTL, may also be used to improve quantitative traits governed by many small-effect QTL.

Quantitative traits can be improved by exotic introgressions

Several previous research groups have identified experimental lines with exotic introgressions in a predominantly elite recurrent background that performed better than the recurrent parent for quantitative traits (Xiao et al., 1998; Ho et al., 2002; Peel and Rasmusson, 2000; Schmalenbach et al., 2009; Nice, 2015). With the exceptions of Peel and Rasmusson (2000), the aims of these studies were foremost to map QTL for grain yield and productivity-related traits in an advanced backcross population; therefore, the gains were observed absent any selection during backcrossing generations. Even Peel and Rasmusson (2000), whose objective was to improve genetic variability and grain yield by introgressing elite exotic germplasm, avoided imposing selection until the completion of multiple backcross generations. In contrast, the WIGS population was developed to explore breeding-focused objectives and combines multiple exotic introgressions into individual WIGS lines, and resulted in a breeding population with heritable genetic variation ($H^2 = 0.57$) for grain yield. Although it could not be determined whether GS was effective in this study, it has been demonstrated to be effective in numerous other empirical studies and, upon considering the suggestions we make herein, should be strongly considered in future efforts to introgress exotic germplasm into elite breeding pools. Taken together, WIGS embodies a contemporary paradigm for using exotic donors in plant breeding.

Importantly, nine WIGS lines performed numerically equal to or better than Rasmusson in a majority of the five MN and WI testing locations, supporting the use of exotic donors in plant breeding to improve quantitative traits. Despite the lack of statistical significance ($\alpha = 0.05$) we, as breeders, feel that the lines performing consistently better than Rasmusson across trials deserve additional consideration as breeding materials. We highlight the case of the highest yielding WIGS line, 3502-112, which had an average grain yield 6.8% greater than Rasmusson across five MN and WI locations and carried an exotic introgression in 3.87% of its genome. Again, while this yield increase is not significant, line 3502-112 performed better than Rasmusson in four of the five trials, and equal with Rasmusson in the fifth (Table 2.1), suggesting that the genetic variance arising from exotic introgressions had a positive effect on grain yield.

The question now becomes whether a 6.8% increase is meaningful in the context of breeding. Rasmusson was grown in five MN locations in the 2015 MN state variety trials, four of which overlap with the WIGS trials (MAES 2015). Across these five trials Rasmusson was the highest yielding variety, significantly ($\alpha = 0.05$) out-performing the most recent six-row variety release from the UMN program, Quest, by 12.1%. Rasmusson also out-performed the most widely grown malting variety in MN, Lacey, by 5.7%, which was not a significant difference. The three-year average of Rasmusson across these five MN locations is again the highest among all varieties, yielding 7.9% higher than Lacey and 6.9% higher than Quest, both of which are significant increases ($\alpha = 0.05$) (MAES 2015).

From the variety trial data we also see that our $LSD_{0.05}$ value is within expected values. The $LSD_{0.05}$ values of the 2015 variety trials conducted in Crookston, Stephen, Saint Paul, and Roseau, MN, locations that overlap our WIGS trials, are equal to 7%, 19%, 11%, and 12% of the trial means, respectively. In the WIGS trial the $LSD_{0.05}$ was 14.2% of the experiment mean across all locations. The variety trials included fewer entries but were grown in larger plots in randomized complete block designs with three replications per trial. Overall, Rasmusson is clearly the highest yielding barley variety currently available to growers in MN, making the 6.8% yield increase observed for WIGS line 3502-112 substantial.

These high yielding lines could, upon further testing for other important agronomic and malting quality traits, be recommended for variety release. More likely, however, is their use as parents for further breeding efforts to increase the genetic variability in the elite barley breeding population for grain yield and other quantitative traits. For example, the nine top performing lines in WIGS that were largely enriched for the recurrent alleles, which contributed the favorable marker effect at 25 of 27 loci detected in the association analysis, could presumably be further improved by crossing to other WIGS or AB-NAM lines carrying, for example, an exotic introgression at the 6H locus where a positive marker effect was associated with exotic genotypes.

The *ad hoc* use of association analysis in this study to identify lines' composition for favorable markers in regions of putative yield QTL, which, as suggested, could be used to determine parent combinations, is a relaxed form of marker-assisted selection. In *elite x exotic* crosses the largest effect QTL may often correspond to domestication genes

for grain shattering, florescence architecture, photoperiod sensitivity, and tillering, among others (Doebley et al., 2006; Pourkheirandish and Komatsuda, 2007; Sang, 2009); such genes deserve special attention in any marker-assisted selection scheme or genomic prediction model.

In fact, the QTL chromosome 3H (39.5 – 46 cM) maps to a pair of domestication genes, *brittle rachis 1* (*btr1*) and *brittle rachis 2* (*btr2*) on chromosome 3H (Komatsuda et al., 2004; Pourkheirandish and Komatsuda, 2007), which were also identified in the AB-NAM (Nice, 2015) and another AB-QTL study (VonKorff et al., 2006). Nice (2015) identified a QTL associated with grain yield on 2H at 19.5 cM that coincides with the domestication gene, *Ppd-H1*. This region was not detected in the WIGS population, however, possibly because the nearest segregating markers were approximately 4 cM away and existed at a low frequency (MAF ≈ 0.05). The QTL on 5H identified herein co-localizes to a QTL region identified in the AB-NAM (Nice, 2015) and AB-QTL population of VonKorff et al. (2006) for grain yield, but does not appear to be associated with a domestication gene. Identifying these large-effect QTL, which in some cases may correspond to domestication genes, and selecting upon them by, perhaps, incorporating the associated markers as fixed effects in the GS model (Bernardo, 2014; Spindel et al., 2016), will facilitate a more rapid recovery of the elite parent's phenotype and growth habit.

Not as simple as Rasmusson vs. Wild

The association analyses and genomic prediction models effectively contrasted the 'recurrent allele' with the 'exotic allele', which ignores the possibility of multiple wild alleles and likely hindered power and accuracy. This simplifying assumption would be appropriate if all exotic alleles had the same effect and SNPs perfectly discriminated between the recurrent Rasmusson allele and all exotic alleles. However, the cases of $SNP_{QTL\text{-}1}$ and $SNP_{QTL\text{-}2}$ highlight that neither assumption is appropriate.

$SNP_{QTL\text{-}1}$ is the most significant marker identified by either association analysis with an FDR-adjusted *p*-value of 0.006, but when the linear effect estimates of the three haplotypes defined by the two markers not identified as significant by the association analysis at QTL-1 (Figure 2.2) are fit as a covariate in the unified MLM, the effect of $SNP_{QTL\text{-}1}$ becomes non-significant ($p = 0.12$). Although the haplotypes at QTL-1 are not significantly different than one another (Figure 2.2), this result nonetheless suggests that the genetic variation attributed to $SNP_{QTL\text{-}1}$ is better explained by the multiple haplotypes than $SNP_{QTL\text{-}1}$ alone. Unlike $SNP_{QTL\text{-}1}$, $SNP_{QTL\text{-}2}$ did not perfectly discriminate between Rasmusson and exotic alleles (Figure 2.2). Again, when the effects of the three haplotypes defined by the two markers not identified as significant by the association analysis at QTL-2 (Figure 2.2) were fit as covariates in the MLMM analysis the significance of $SNP_{QTL\text{-}2}$ decreased from *p*-value = $5.94E^{-7}$ to $p = 0.56$.

These two examples briefly illustrate that biallelic markers are insufficient to characterize the alleles in this multi-parent population. Accounting for multiple alleles from the wild accession donors may lead to more power in an association analysis and a

reduced number of false positives. For example, the grouping of the "ACG" haplotype with the recurrent haplotype at QTL-2 would have strengthened the contrast between the "G" and "C" variants at $SNP_{QTL\text{-}2}$, perhaps leading to a false positive. One can easily imagine the opposite scenario where the inability to resolve multiple exotic haplotypes results in the dilution of an individual exotic haplotype and a false negative result. Likewise, accounting for haplotypes in genomic prediction models may lead to more accurate genome-based predictions.

Such prediction models that consider haplotype regressors rather than single SNPs have been suggested and have led to improved genomic prediction accuracies in simulation and empirical studies (Calus et al., 2008; Villumsen et al., 2009; Calus et al., 2009; Cuyabano et al., 2014; Cuyabano et al., 2015). Improvements under haplotype-based models in these animal studies were often modest, but given the structure of WIGS and our knowledge that biallelic SNPs are insufficient to capture variation at some loci, haplotype-based methods may prove to be beneficial. A foreseeable complication of investigating haplotype based GS with this dataset is the extensive amount of missing data, a hallmark of GBS. Incomplete genotypic data leads to uncertainty in haplotypes; for example, AB-NAM line W350-04, a parent of multiple WIGS populations, is missing genotypic data for all SNPs characterizing the haplotypes of the QTL on chromosome 6H, so its haplotype was inferred from relatives. More complete sequencing data from the WBDC founder lines, or AB-NAM parental lines could be used to impute missing marker variants among WIGS lines or to derive pedigree-based haplotypes. Testing such methods is beyond the scope of the current study, but the WIGS population is a good

candidate to investigate the use of haplotype-based models for breeding with wild relatives.

Effectiveness of marker-based strategies inconclusive

We expected that IPV would identify elite parent combinations that would otherwise be ignored under traditional methods such as GD and BxB. The results, however, do not support the use of IPV. While two of the top three crosses, by family mean, were identified by IPV, these families were represented by only one to three progeny (Supplemental Table 2.2). Similarly, while four of the progeny within the 23 lines that performed better than Rasmusson were identified by IPV, so were 34 of the lowest 50 performing. The resulting performance of IPV may have been due to an oversampling of specific AB-NAM parents, not realizing the necessary recombination between the complementary portions of the parental genomes, or inaccurate marker effect estimates, as suggested by the low progeny-based prediction accuracy. Here again, we see a potential benefit to the use of haplotype-based effect estimates. Alternatively, the criteria of BxB and GD do not depend on accurate effect estimates, but rather accurate genotyping and phenotyping of AB-NAM parents, which is supported by a heritability estimate of $H^2 = 0.61$ for the AB-NAM trials (Nice, 2015), and accurate genotyping, respectively.

Genomic selection is effective in many breeding applications (reviewed in Bernardo, 2016 and Tiede et al., *in preparation*), including scenarios of introgressing exotic germplasm into elite breeding material (Bernardo, 2009; Combs and Bernardo,

2013b; Bernardo, 2016). In our study the cross-validation accuracy of $r = 0.586$ observed in the AB-NAM TP suggests that population-wide selection should have been effective. With a progeny-based prediction accuracy of $r = -0.01$, selection was essentially absent and we cannot make a determination as to whether GS would be generally effective under our scheme of introgressing exotic germplasm. The improvement in prediction accuracy observed upon including a fixed term for family membership in the prediction model was significant. This highlights the importance of diligently evaluating and updating the TP when data on breeding material becomes available.

Breeding with exotic germplasm

Integrating exotic germplasm into an elite breeding population is a substantial undertaking and requires a radically different approach than the traditional *elite x elite* breeding. Take as an example the University of Minnesota's barley breeding program. When breeding solely with elite germplasm, crosses are largely made between highly related individuals (Peel and Rasmusson, 2000; Condon et al., 2008), resulting in a collection of highly related biparental breeding populations. In such a case, a single TP representative of the breeding populations is expected to be indicative of the marker-QTL linkage disequilibrium and genetic relationships present in the breeding populations, and would therefore be suitable for accurate genomic predictions to identify parent candidates and candidates for preliminary varietal testing.

In contrast, breeding with wide *elite x exotic* crosses requires the additional steps of identifying suitable exotic sources, likely from a germplasm collection like the WBDC

or National Small Grains Collection (NSGC), pre-breeding to introgress the exotic alleles into a majority recurrent background, and extra care in developing a training population and genomic selection pipeline. The AB-NAM and other NAM populations are useful tools for bridging the gap between exotic germplasm, which in NAM populations are largely wild relatives and landraces, and modern cultivated varieties for breeding purposes. While existing datasets for grain yield and other quantitative traits will likely be already available to inform decisions on which exotic source to focus on, NAM resources are limited in that they often use a single recurrent parent genotype, whereas breeders will likely opt to introgress exotic donors into a number of relevant varieties as well as advanced breeding lines as in Peel and Rasmusson (2000). Additionally, NAM populations are generally not advanced backcross populations like the AB-NAM, so additional backcrossing generation will be needed to reduce the negative effects of linkage drag of deleterious alleles from exotic sources.

Furthermore, implementing marker-based breeding in exotic-derived, multi-family breeding populations will require specified training populations and prediction models that accommodate more than two alleles. Herein the AB-NAM population was used to train the prediction model; however, a more balanced F_2- or BC_1-derived population of WIGS crosses, as was done in Combs and Bernardo (2013b) and suggested by Bernardo (2016) may have been more effective; doing so would require at least an additional year for phenotyping, but the more related population will likely result in higher prediction accuracies (Clark et al., 2012; Lorenz and Smith, 2015). Additionally, Bernardo (2016), focusing on a single biparental population, proposes the use of GS to

identify progeny to backcross to the recurrent parent in each cycle, a process which is repeated five or more generations depending on the genetic architecture of the trait and other variables. In a biparental population, biallelic markers are sufficient to discriminate between the recurrent and donor parent, so extending this approach to a breeding population with multiple recurrent parents and multiple exotic donors will likely require prediction models using haplotype regressors or, perhaps, hierarchical Bayesian models (Gelman, 2006; Gelman and Pardoe, 2006) for partial pooling of training population data to estimate family-specific marker effects that are informed, through the Bayesian framework, by priors estimated from the whole population (Technow and Totir, 2015).

Conclusion

Absent any selection during the development of the advanced backcross population, we were able to derive progeny from crosses among AB-NAM lines that consistently performed numerically better than the elite recurrent parent, Rasmusson. The top 20 WIGS lines will be evaluated in multi-environment field trials during the 2016 growing season (April – July) to further support, or not, the gains in grain yield observed across the 2015 WIGS trials. It is likely an approach involving GS during backcrossing would have resulted in even greater gains for grain yield. In addition, we suggest a series of marker-based protocols and methodologies, based on our observations herein and other work (Peel and Rasmusson, 2000; Combs and Bernardo, 2013b; Bernardo, 2016). While the IPV parent selection strategy did target crosses that more traditional phenotypic and marker-based methods such as BxB and GD, respectively, would have ignored, it was not

shown to consistently identify high yielding lines in this study. Therefore, GD and BxB appear to be the most effective in identifying superior crosses between the top performing parent candidates.

Table 2.1: Important summary items for c.v. Rasmusson and the nine highest yielding WIGS lines. Includes the percent the genome represented by exotic introgressions, the overall adjusted mean of the line measured across five trials, as well as the line's grain yield (kg ha^{-1}) within each individual trial. Shaded values indicate the line's grain yield was numerically higher than Rasmusson in that trial.

				Grain Yield				
Line[†]	**Strategy**[‡]	**% Exotic**	$\overline{Y}$[§]	**WI**[¶]	**Crk**	**Roseau**	**Stephen**	**StP**
Rasmusson	--	--	5.95	5.84	8.11	8.02	8.20	7.46
3502-012	GD, BxB	3.87	6.36	5.84	8.59	9.02	8.48	7.60
3507-034	GD	6.78	6.27	5.76	8.24	8.96	8.47	7.65
3526-024	GD	10.82	6.21	7.29	8.21	--	8.45	7.07
3507-073	GD	9.81	6.21	7.04	8.28	8.43	7.15	7.46
3513-020	IPV	11.38	6.20	6.23	7.75	8.15	8.83	7.87
3526-086	GD	7.49	6.19	6.80	8.16	8.11	8.17	7.01
3513-086	IPV	10.97	6.18	5.41	8.35	8.10	8.38	8.24
3501-115	GD	5.06	6.15	6.76	7.88	8.37	8.57	6.85
3526-012	GD	11.73	6.15	6.80	7.83	8.13	7.91	7.45
		$LSD_{0.05}$[#]	0.78	1.65	0.91	1.61	1.27	1.54

[†] Line name of WIGS entry. The first four numbers designate the WIGS family, and the following three designate the specific progeny with a family

[‡] GD, genetic distance; BxB, best-by-best; IPV, idealized progeny value

[§] Adjusted phenotypic mean (see Eq. 2.3) in kg ha^{-1}

[¶] WI, Buffalo County, WI; Crk, Crookston, MN; Roseau, MN; Stephen, MN; StP, Saint Paul, MN

[#] Least significant differences at $\alpha = 0.05$, pertains to each column of data

Table 2.2: Means of the selected lines and bulks of lines representing the population mean and lowest lines, as predicted by genomewide markers, per crossing strategy.

Crossing Strategy[†]	Selected Mean[‡]	Mean Bulk[§]	Low Bulk
BxB	5653.65	5403.98	5289.04
GD	5635.0	5272.36	5144.043
IPV	5351.59	5368.37	4647.33

[†] GD, genetic distance; BxB, best-by-best; IPV, idealized progeny value
[‡] Mean of the selected WIGS lines per crossing strategy
[§] Mean and low bulks consisted of a random 30 selection candidates, and the 30 lowest predicted selection candidates, per cross
[¶] Least significant differences at $\alpha = 0.05$, pertains to all values

Table 2.3: Significant marker associations with grain yield identified by either or both the unified MLM and MLMM association analyses. Shaded rows indicate that the marker was found significant by both association analyses.

Chr[†]	cM[†]	*P*-value[‡]	Effect[§]	Frequency[¶]	
				Highest	*Lowest*
2H	58.8	3.74E-05	434.2	0.94	0.56
3H	39.5	1.55E-05	309.9	1.0	0.39
	43.9	2.23E-04	321.2	1.0	0.17
	45.8	2.86E-04	241.4	1.0	0.17
	46.0	6.05E-05	310.4	1.0	0.39
	106.1	1.30E-04	255.0	1.0	0.17
	108.4	4.16E-06	273.2	1.0	0.39
	108.4	1.76E-05	270.3	1.0	0.17
	108.4	7.60E-05	274.1	1.0	0.33
	108.6	3.44E-12	300.8	1.0	0.17
	108.6	4.54E-05	251.2	1.0	0.67
	108.6	2.14E-04	215.4	1.0	0.33
	108.6	2.41E-04	301.1	1.0	0.17
	108.8	6.07E-05	259.4	1.0	0.17
5H	39.8	2.81E-05	362.3	1.0	0.50
	43.8	2.25E-05	355.2	1.0	0.50
	43.8	7.79E-05	332.8	1.0	0.50
	48.4	9.77E-05	316.6	1.0	0.50
	48.4	2.86E-04	283.2	1.0	0.44
	49.7	3.77E-08	322.6	1.0	0.50
	49.7	3.55E-05	356.0	1.0	0.44
	50.0	2.03E-04	328.7	1.0	0.56
	50.4	3.31E-05	364.3	1.0	0.50
6H	1.8	5.94E-07	-40.4	0.17	0.28
	65.4	2.20E-04	-322.6	0.06	0.6
7H	47.0	4.50E-08	311.1	0.89	0.94
	54.4	4.20E-06	162.3	1.0	0.89

[†] Position of Morex contig (IBGSC, 2012) containing the significantly associated SNP, according to the POPSEQ map (Mascher et al., 2013). Chr, chromosome; cM, centi-Morgan

[‡] Unadjusted *P*-value of the association analysis. If a SNP was found significant in both association analyses the lower *P*-value of the two is presented.

[§] Estimate of the additive effect of the marker inherited from the recurrent parent, c.v. Rasmusson, on grain yield (kg ha^{-1})

[¶] Frequency of the marker having the positive additive marker effect estimate among the nine *highest* performing WIGS lines, and *lowest* performing nine WIGS lines

Figure 2.1: Regions of exotic introgressions in the highest and lowest performing WIGS lines, highlighting QTL regions identified by association analysis with a p-value ≤ 0.20. A: Manhattan plot where shades and and of points refer to the association analysis with which the association was identified, and the source of the favorable marker effect (exotic or recurrent), respectively. B and C: Genomic regions, based on the POPSEQ map (Mascher et al., 2013), of the nine highest (B) and nine lowest (C) performing entries in WIGS, harboring exotic introgressions. Identify the maternal or paternal origin of the exotic parent from the legend. Shading of the introgression regions is relative to the interpolated value, with the darkest color referring to the homozygous genotype of the respective exotic parent.

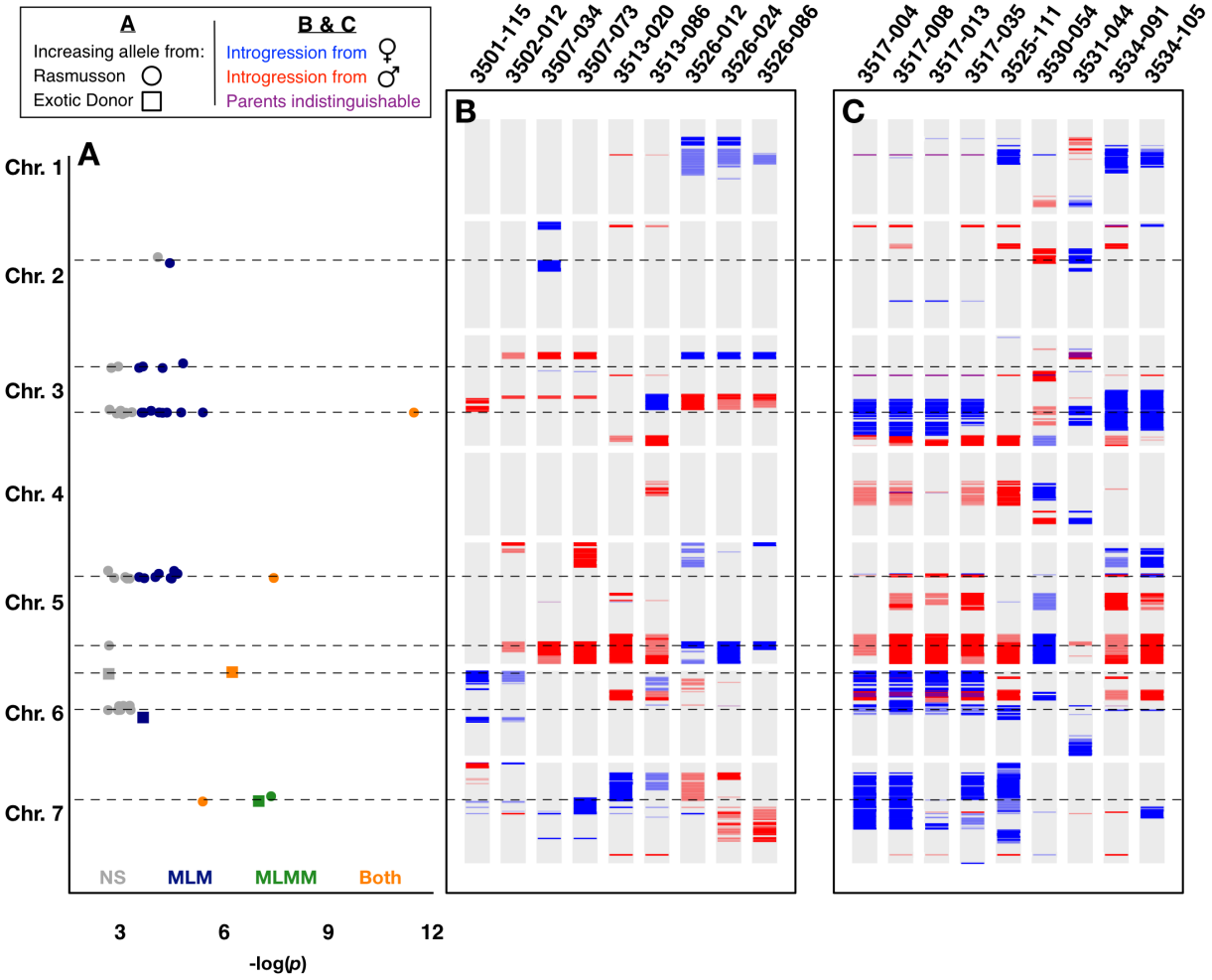
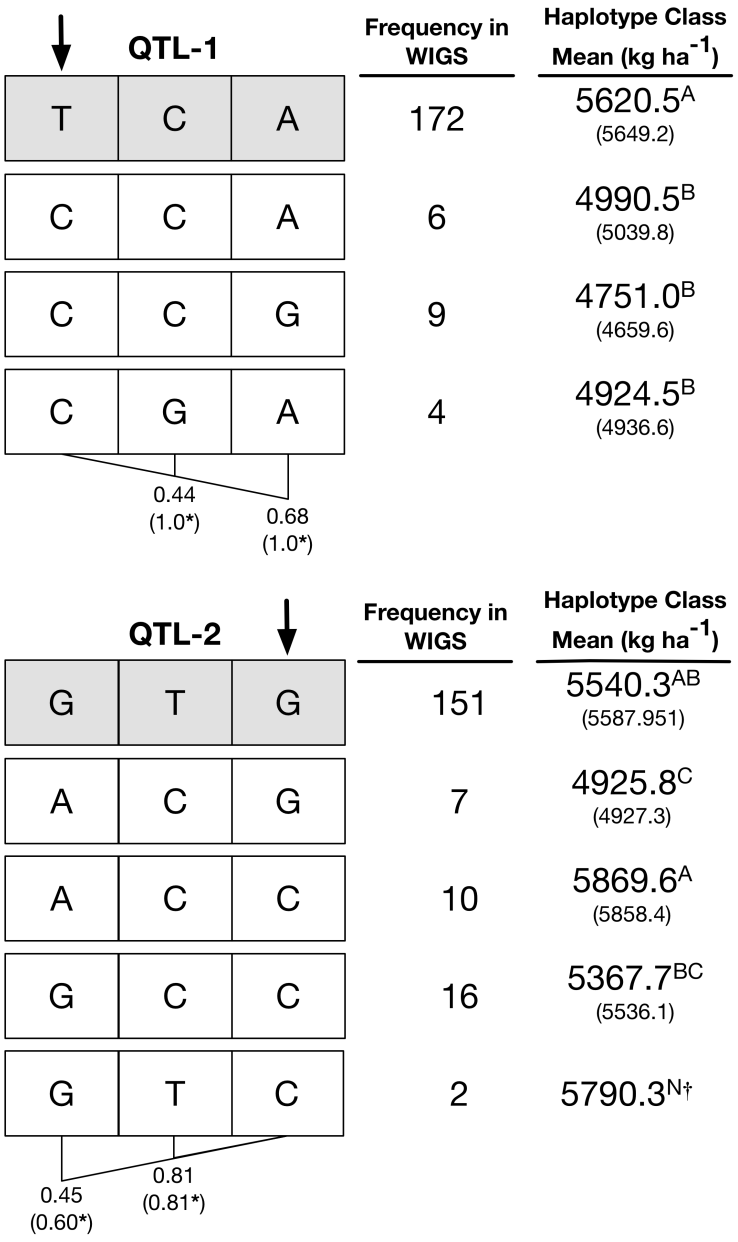


Figure 2.2: Haplotypes at QTL-1 and QTL-2 within WIGS. The arrow refers to the SNP identified in the association analysis, which is in linkage disequilibrium (LD) with the adjacent markers according to measures of r^2 and (D'), where an asterisk indicates that D' is significantly different (p-value < 0.05) than zero (linkage equilibrium). The frequency of unambiguous haplotype classes in WIGS and the class mean (and median) of WIGS lines carrying that haplotype are indicated. Letter superscripts indicate significant differences (p-value < 0.05) between haplotype class means according to Tukey's HSD. †Haplotype class consisted of only two WIGS lines and was not compared to other class means.



Supplemental Table 2.1: Pedigrees of highest nine yielding WIGS lines.

Line[†]	WIGS Family	Maternal Parent	Paternal Parent
3501-115	3501	W016-27	W103-01
3502-012	3502	W016-27	W302-13
3507-034	3507	W042-13	W302-13
3507-073	3507	W042-13	W302-13
3513-020	3513	W092-30	W350-04
3513-086	3513	W092-30	W350-04
3526-012	3526	W302-13	W092-30
3526-024	3526	W302-13	W092-30
3526-086	3526	W302-13	W092-30

[†] Line name of WIGS entry. The first four numbers designate the WIGS family, and the following three designate the specific progeny with a family

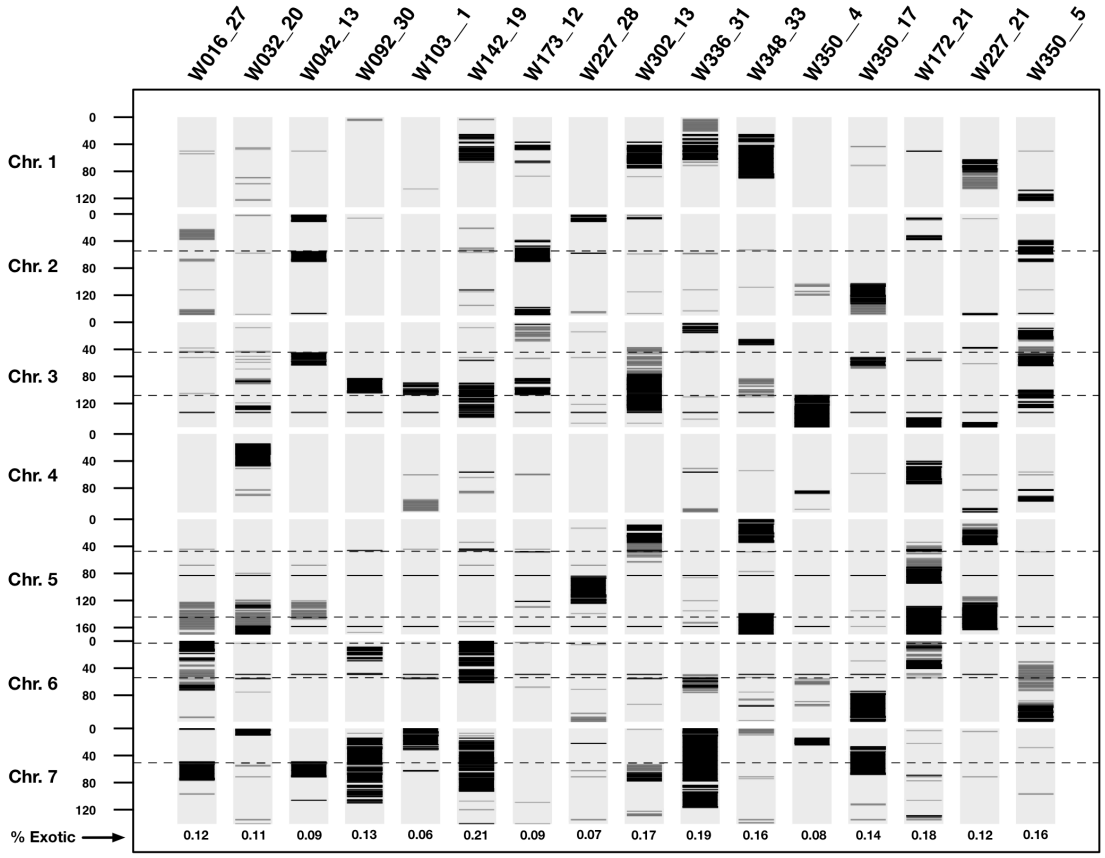
[‡] The maternal and paternal parents are from the AB-NAM population (Nice et al., *in press*).

Supplemental Table 2.2: Summary of WIGS families.

WIGS Family	Crossing Strategy[†]	n[‡]	Rank[§]	Family Average[¶]
3501	IPV, GD	8	10	5652.1
3502	GD, BxB	13	4	5762. 0
3503	BxB	1	5	5756.7
3504	BxB	15	13	5592.9
3507	GD	4	1	5928.4
3511	IPV	3	3	5799.0
3513	IPV	24	11	5632.1
3514	BxB	19	8	5715.5
3515	GD, IPV	18	14	5578.0
3517	IPV	17	20	5077.4
3519	IPV	8	17	5314.0
3520	BxB	2	15	5467.5
3522	BxB	3	6	5733.7
3525	IPV	7	21	4983.4
3526	GD	12	7	5730.5
3527	ALL	5	9	5672.9
3528	GD	7	16	5420.0
3529	IPV	2	2	5865.1
3530	IPV	7	19	5104.9
3531	GD	1	22	4022.6
3533	GD, BxB	12	12	5622.1
3534	IPV	14	18	5248.6

[†] GD, genetic distance; BxB, best-by-best; IPV, idealized progeny value
[‡] Number of individuals within a family
[§] Rank of family among all WIGS families, according to grain yield
[¶] Mean grain yield (kg ha^{-1}) of family, based on adjusted entry means

Supplemental Figure 2.1: Regions of exotic introgressions in the AB-NAM parents of the whole WIGS population. The first four characters of the AB-NAM parents designates family membership within AB-NAM, and the following numbers designate the specific progeny with the family. '% Exotic' expresses the proportion of the genome represented by an exotic introgression, as determined from the interpolated genotypic data. Shading of the introgression regions is relative to the interpolated genotypic value, with the darkest color referring to the homozygous genotype of the respective exotic parent.



Chapter 3: Future directions

I have found the direction of research to be more web-like than linear. Answering one question only opens doors to more unknowns. In this chapter I will propose some logical next steps to expand upon what my thesis research has so far developed and uncovered. I will also discuss further research projects to address new directions and questions that have arisen during the course of my graduate work. Definitions and abbreviations of terms may be carried from the chapters into their respective sections. I must also note that some of these ideas have been developed through discussion with others, including Dr. Kevin Smith and Dr. Aaron Lorenz.

Chapter 1

At this time, the six-row malting barley breeding program has undergone five cycles of selection. Genomic predictions for the fourth and fifth cycles, however, were made using different marker datasets and training populations (TP). Nonetheless, a set of cycle four (C4) and C5 progeny could be evaluated for grain yield (GY) and deoxynivalenol (DON) with a subset of the lines in GS_VP, which represents the first three breeding cycles, to measure additional gain from selection and changes in genetic variance. Alternatively, bulks of each population could be grown together in a number of replicated trials, which would greatly reduce the size of the experiment, but genetic variance per cycle would not be estimable.

In Chapter 1 we defined TP sizes of 500 and 300 for GY and DON, respectively, to accommodate the optimization algorithms used. The original publications of these

optimization algorithms often cite the case that breeders have limited resources and can evaluate a set number of lines to form the TP. This is true to an extent, but it may very well be that breeders would be willing to shift resources to TP development and maintenance if it would ensure higher prediction accuracy in the next breeding cycle, thereby increasing genetic gain and possibly reducing the resources needed downstream for preliminary field trials. For this, it would be useful to identify a general metric to inform the optimal size of the TP when using an optimization algorithm. For example, using G_mean, if the metric is demonstrated to be a $K_{ij} = 0.25$, then only TP candidates with an average kinship to the selection candidates of 0.25 or greater would be included in the TP.

The resources presented in Chapter 1 would provide three validation populations (VP) and a number of TPs that can be used to empirically investigate such metrics. Since measuring gain from selection would no longer be the focus, traits that can be measured on the grain (e.g. malt quality, test weight, 1000 kernel weights) could be included with DON and GY. Then, continuing with the G_mean approach, the response in prediction accuracy as TP candidates are sequentially added, from most related on average to least related, can identify a point (i.e. the metric) at which there is no benefit from further additions to the TP. If consistent across TP-VP combinations then that metric could be used when making genomic predictions in subsequent breeding cycles. To make that metric more widely applicable it would be helpful to translate the marker-based estimate of relatedness, in the case of G_mean, to pedigree-based relationship coefficients in a series of crops.

As empirical datasets across multiple selection cycles have become available in the barley breeding lab we are observing that cross-validation (CV) based prediction accuracies overestimate progeny-based prediction accuracies. This was observed in Chapter 1, where the CV-based accuracies were 0.638 for GY and 0.688 for DON, but progeny-based prediction accuracies in C1, C2, and C2 were 0.124, 0.264, and 0.415 for GY, and 0.288, 0.194, and 0.182 for DON, respectively. An extreme example was observed in Chapter 2 where the CV-based accuracy in the AB-NAM TP was 0.586, but the actual correlation between the predicted genotypic values and adjusted line means of WIGS lines was approximately zero. The literature on GS depends heavily on CV procedures, and while there are studies comparing CV methodologies, there does not appear to be any studies comparing CV prediction accuracies to progeny-based prediction accuracies in actual breeding populations under selection. Resources from the statistics field (Rao et al., 2008) and the theoretical underpinning of prediction accuracy (Daetwyler et al., 2008; Goddard, 2009; Goddard et al., 2011; Combs and Bernardo, 2013a) would provide a good starting point for an in-depth investigation into the translation of CV-based prediction accuracies to empirical, progeny-based prediction accuracies.

Specifically, a TP could be constructed for each breeding cycle, which can be considered as independent VPs, based on what was learned in Chapter 1. The TP could be subsampled in a manner similar to the widely used ten-fold CV strategy, and during each iteration the statistics outlined in the studies cited above – TP size, heritability, and effective number of chromosomes, which is a function of the effective population size,

number of chromosomes segregating in the population, and the average length of a chromosome, and the trait heritability – can be reported along with the CV-based prediction accuracy of the subset and empirical prediction of the subset. Doing this over a number of iterations should provide a landscape of data to draw conclusions from and enable us to provided the breeding community using GS insight into whether CV accuracy, in their case, will be indicative of progeny-based prediction accuracy and, if not, steps that can be taken to mitigate the disconnect.

Chapter 2

Integrating exotic germplasm into elite breeding programs is challenging for traits conditioned by one or a few genes, and certainly an even greater challenge for quantitative traits. I believe that using genomewide marker-based prediction strategies for progeny and parent selection will hasten the process so that a breeder may realize success in a relatively short time period, perhaps less than 10 years. Granted, the progeny-based prediction accuracy of $r \approx 0$ does not allow us to draw a conclusion about the effectiveness of GS, nor does it give us confidence that the merit of a strategy such as IPV was sufficiently evaluated in Chapter 2. This leads to the obvious need to develop, or implement a proposed model (Calus et al., 2008) to estimate haplotype effects rather than individual marker effects. Whether haplotype-based models can improve the accuracy of predicted genotypic values can readily be tested – simply train such a model using the AB-QTL population and see if the prediction accuracy of WIGS lines improves

significantly. An improvement of GS accuracy would suggest that estimating effects of haplotypes would also benefit IPV, although that is not as easily verified since crosses between AB-NAM parents would need to be remade, which would essentially dismiss the yield gains observed in the WIGS population reported in Chapter 2. Rather, crosses between the top performing WIGS lines could be targeted by a marker-based crossing strategies, which could again be compared to GD or BxB.

Since the use of IPV during the development of the WIGS population there has been another marker-based parent selection strategy proposed. The more recent procedure estimates genetic variance by simulating recombination between two parent candidates, which can then be used in concert with the estimated mean of the resulting biparental population (i.e. the mid-parent value) to identify crosses that would produce the highest performing progeny (Bernardo, 2014; Mohammadi et al., 2015). The simulation step of this new procedure, which we call 'PopVar', would replace the 'idealized' recombination requirement of IPV with a probabilistic approach based on recombination rates according to a genetic map. Again, crosses targeted by PopVar and either GD or BxB could be made among top WIGS lines, thereby capitalizing on the gains in grain yield observed in Chapter 2.

If haplotype-based predicted genotypic values estimates are shown to significantly improve prediction accuracy then it would make sense to re-predict genotypic values for all ~3,000 genotyped F_3 selection candidates (~1,000 representing each parent selection strategy). Any selection candidates predicted to be high performing that do not already exist in WIGS could be advance to field testing and/or included in the

parent candidates of the study proposed in the previous paragraph. Additionally, as the AB-QTL population is evaluated for additional traits important to the breeding program, such as Fusarium head blight resistance, lodging resistance, malting quality, etc., genotypic values can be predicted for all ~3,000 selection candidates and the top predicted lines can be moved forward for field testing.

Gorjanc et al. (2016) recently discussed initiating a pre-breeding program directly from maize landraces. This is an interesting proposition, but likely a futile endeavor in barley since barley is an inbred crop and cannot be testcrossed to elite parents. Rather, I would propose that a synthetic population be derived from combining top-performing AB-NAM and WIGS lines with advanced breeding germplasm from the UMN barley breeding population. Even more, elite two-row cultivars could be included, resulting in a synthetic segregating for two- and six-row head morphology that combines favorable alleles selected upon for decades within the different breeding backgrounds. After a few cycles of recombination without intentional selection, selections can begin to be made out of the synthetic and inbred via SSD or DH for field evaluation. As phenotypic information is gathered the inbred lines can also be genotyped to develop a TP that could be used for to identify lines from the synthetic to evaluate in field trials to be used as parents for the next cycle, advanced for potential variety release, and/or TP maintenance and improvement.

Bibliography

Akdemir D., Sanchez J.I. & Jannink J.L. 2015. Optimization of genomic selection training populations with a genetic algorithm. *Genet Sel Evol* 47: 38.

Akdemir, D. & Godfrey, O.U. (2015). EMMREML: Fitting Mixed Models with Known Covariance Structures. R package version 3.1. http://CRAN.R-project.org/package=EMMREML

Anderson J.A. 2007. Marker-assisted selection for Fusarium head blight resistance in wheat. *International journal of food microbiology* 119: 51-53

Ariyadasa R, Mascher M, Nussbaumer T, Schulte D, Frenkel Z, Poursarebani N, Zhou R, Steuernagel B, Gundlach H, Taudien S et al. 2014. A sequence-ready physical map of barley anchored genetically by two million single-nucleotide polymorphisms. *Plant Physiol.* 164:412-423.

Asoro F.G., Newell M.A., Beavis W.D., Scott M.P. & Jannink J.-L. 2011. Accuracy and Training Population Design for Genomic Selection on Quantitative Traits in Elite North American Oats. *The Plant Genome Journal* 4: 132.

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 Sci.* 53: 1894.

Bai G., Kolb F.L., Shaner G. & Domier L.L. 1999. Amplified fragment length polymorphism markers linked to a major quantitative trait locus controlling scab resistance in wheat. *Phytopathology* 89: 343-348.

Bates D., Mächler M., Bolker B. & Walker S. 2014. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1-48

Beaulieu J., Doerksen T.K., MacKay J., Rainville A. & Bousquet J. 2014. Genomic selection accuracies within and between environments and small breeding groups in white spruce. *BMC Genomics* 15: 1048.

Benjamini Y. & Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 289-300.

Bernacchi D., Beck-Bunn T., Emmatty D., Eshed Y., Inai S., Lopez J., Petiard V., Sayama H., Uhlig J. & Zamir D. 1998. Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from Lycopersicon hirsutum and L. pimpinellifolium. *Theoretical and Applied Genetics* 97: 170-180.

Bernardo R. & Yu J. 2007. Prospects for Genomewide Selection for Quantitative Traits in Maize. *Crop Sci.* 47: 1082.

Bernardo R. 1994. Prediction of Maize Single-Cross Performance Using RFLPs and Information from Related Hybirds. *Crop. Sci.* 34: 20.

Bernardo R. 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Science* 48: 1649-1664.

Bernardo R. 2014. Genomewide selection of parental inbreds: classes of loci and virtual biparental populations. *Crop Sci.* 54(6): 2586–2595

Bernardo R. 2014. Genomewide Selection when Major Genes Are Known. *Crop Science* 54: 68.

Bernardo, R. 2016. Genomewide Predictions for Backcrossing a Quantitative Trait from an Exotic to an Adapted Line. *Crop Sci.* 56(2). doi: 10.2135/cropsci2015.09.0586

Beyene Y., Semagn K., Mugo S., Tarekegne A., Babu R., Meisel B., Sehabiague P., Makumbi D., Magorokosho C., Oikeh S., Gakunga J., Vargas M., Olsen M., Prasanna B.M., Banziger M. & Crossa J. 2015. Genetic Gains in Grain Yield Through Genomic Selection in Eight Bi-parental Maize Populations under Drought Stress. *Crop Science* 55: 154.

Brondani C., Rangel P., Brondani R. & Ferreira M. 2002. QTL mapping and introgression of yield-related traits from Oryza glumaepatula to cultivated rice (Oryza sativa) using microsatellite markers. *Theoretical and Applied Genetics* 104: 1192-1203.

Brozynska M., Furtado A. & Henry R.J. 2015. Genomics of crop wild relatives: expanding the gene pool for crop improvement. *Plant Biotechnol J*

Calus M.P., Meuwissen T.H., de Roos A.P. & Veerkamp R.F. 2008. Accuracy of genomic selection using different methods to define haplotypes. *Genetics* 178: 553-561.

Calus M.P., Meuwissen T.H., Windig J.J., Knol E.F., Schrooten C., Vereijken A.L. & Veerkamp R.F. 2009. Effects of the number of markers per haplotype and clustering of haplotypes on the accuracy of QTL mapping and prediction of genomic breeding values. *Genet Sel Evol* 41: 11.

Clark, S.A., Hickey J.M., Daetwyler H.D., & van der Werf J.H.J. 2012. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genet Sel Evol* 44: 4

Close T.J., Bhat P.R., Lonardi S., Wu Y., Rostoks N., Ramsay L., Druka A., Stein N., Svensson J.T., Wanamaker S., Bozdag S., Roose M.L., Moscou M.J., Chao S., Varshney R.K., Szucs P., Sato K., Hayes P.M., Matthews D.E., Kleinhofs A., Muehlbauer G.J., DeYoung J., Marshall D.F., Madishetty K., Fenton R.D., Condamine P., Graner A. & Waugh R. 2009. Development and implementation of high-throughput SNP genotyping in barley. *BMC Genomics* 10: 582.

Combs E. & Bernardo R. 2013a. Accuracy of Genomewide Selection for Different Traits with Constant Population Size, Heritability, and Number of Markers. *The Plant Genome* 6.

Combs E. & Bernardo R. 2013b. Genomewide Selection to Introgress Semidwarf Maize Germplasm into U.S. Corn Belt Inbreds. *Crop Sci.* 53: 1427.

Condón F., Gustus C., Rasmusson D.C. & Smith K.P. 2008. Effect of advanced cycle breeding on genetic diversity in barley breeding germplasm. *Crop science* 48: 1027-1036.

Crossa J., Campos G.L., Pérez P., Gianola D., Burgueño J., Araus J.L., Makumbi D., Singh R.P., Dreisigacker S., Yan J., Arief V., Banziger M. & Braun H.J. 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* 186: 713-724.

Cullis B., Gogel B., Verbyla A.N. & Thompson R. 1998. Spatial analysis of multi-environment early generation variety trials. *Biometrics* 1-18.

Cullis B.R., Smith A.B. & Coombes N.E. 2006. On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics* 11: 381-393.

Cuyabano B.C., Su G., Rosa G.J., Lund M.S. & Gianola D. 2015. Bootstrap study of genome-enabled prediction reliabilities using haplotype blocks across Nordic Red cattle breeds. *J Dairy Sci* 98: 7351-7363.

Cuyabano B.C.D., Su G. & Lund M.S. 2014. Genomic prediction of genetic merit using LD-based haplotypes in the Nordic Holstein population. *BMC genomics* 15: 1.

Daetwyler H.D., Villanueva B. & Woolliams J.A. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS One* 3: e3395.

Daetwyler H.D., Villanueva B., Bijma P. & Woolliams J.A. 2007. Inbreeding in genome wide selection. *Journal of Animal Breeding and Genetics* 124: 369-376.

de los Campos G.P.P. 2015. BGLR: A Statistical Package for Whole Genome Regression and Prediction. https://cran.r-project.org/package=BGLR

Doebley J.F., Gaut B.S. & Smith B.D. 2006. The molecular genetics of crop domestication. *Cell* 127: 1309-1321.

Dunnett, C.W. 1980. Pairwise Multiple Comparisons in the Unequal Variance Case. *Journal of the American Statistical Association*. 75 (372): 796-800.

Efron B. 1979. Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7: 1–26.

Ellis R.P., BP Forster, DR Robinson 2000. Wild barley: a source of genes for crop improvement in the 21st century? *Journal of Experimental Botany* 9.

Endelman J.B. 2011. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome Journal* 4: 250.

Eshghi R., Salayeva S., Ebrahimpour F., Rahimi M., Baraty M. & Ojaghi J. 2013. Advanced-backcross QTL analysis in hulless barley: I. Detection of exotic alleles for yield and yield components introgressed from Hordeum vulgare ssp. spontaneum. *Intl J Agri Crop Sci* 5: 95-100.

Fulton T.M., Beck-Bunn T., Emmatty D., Eshed Y., Lopez J., Petiard V., Uhlig J., Zamir D. & Tanksley S.D. 1997. QTL analysis of an advanced backcross of Lycopersicon peruvianum to the cultivated tomato and comparisons with QTLs found in other wild species. *Theoretical and Applied Genetics* 95: 881-894.

Gianola D. & van Kaam J.B. 2008. Reproducing kernel hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics* 178: 2289-2303.

Goddard M. 2009. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica* 136: 245-257.

Goddard M.E., Hayes B.J. & Meuwissen T.H. 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. *J Anim Breed Genet* 128: 409-421.

Gorjanc G., Jenko J., Hearne S.J. & Hickey J.M. 2016. Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. *BMC Genomics* 17: 30.

Grenier C., Cao T.-V., Ospina Y., Quintero C., Châtel M.H., Tohme J., Courtois B. & Ahmadi N. 2015. Accuracy of genomic selection in a rice synthetic population developed for recurrent selection breeding. *PloS one* 10: e0136594.

Guo Z., Tucker D.M., Basten C.J., Gandhi H., Ersoz E., Guo B., Xu Z., Wang D. & Gay G. 2014. The impact of population structure on genomic prediction in stratified populations. *Theor Appl Genet* 127: 749-762.

Gutiérrez L., Nason J.D. & Jannink J.-L. 2009. Diversity and mega-targets of selection from the characterization of a barley collection. *Crop science* 49: 483-497.

Habier D., Fernando R.L. & Dekkers J.C.M. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177: 2389-2397.

Hamblin M.T., Close T.J., Bhat P.R., Chao S., Kling J.G., Abraham K.J., Blake T., Brooks W.S., Cooper B. & Griffey C.A. 2010. Population structure and linkage disequilibrium in US barley germplasm: implications for association mapping. *Crop Sci.* 50: 556-566.

Harlan J.R. & de Wet J.M.J. 1971. Toward a rational classification of cultivated plants. *Taxon* 509-517.

Heffner E.L., Jannink J.-L. & Sorrells M.E. 2011. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *The Plant Genome* 4: 65-75.

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

Heslot N., Yang H.-P., Sorrells M.E. & Jannink J.-L. 2012. Genomic Selection in Plant Breeding: A Comparison of Models. *Crop Sci.* 52: 146.

Ho J., McCouch S. & Smith M. 2002. Improvement of hybrid yield by advanced backcross QTL analysis in elite maize. *Theoretical and Applied Genetics* 105: 440-448.

Holland J.B. 2004. Breeding: Incorporation of exotic germplasm. *Encyclopedia of Plant and Crop Science. Marcel Dekker, Inc., New York* 222-224.

Horsley R.D., Schmierer D., Maier C., Kudrna D., Urrea C.A., Steffenson B.J., Schwarz P.B., Franckowiak J.D., Green M.J. & Zhang B. 2006. Identification of QTLs associated with Fusarium head blight resistance in barley accession CIho 4196. *Crop Sci.* 46: 145-156.

Horsley R.D., Schwarz P.B. & Hammond J.J. 1995. Genetic diversity in malt quality of North American six-rowed spring barley. *Crop science* 35: 113-118.

International Barley Genome Sequencing Consortium, Mayer KF, Waugh R, Brown JW, Schulman A, Langridge P, Platzer M, Fincher GB, Muehlbauer GJ, Sato K et al. 2012. A physical, genetic and functional sequence assembly of the barley genome. *Nature*. 491:711-716.

Isidro J., Jannink J.L., Akdemir D., Poland J., Heslot N. & Sorrells M.E. 2015. Training set optimization under population structure in genomic selection. *Theor Appl Genet* 128: 145-158.

Jannink J.-L. 2003. Selection dynamics and limits under additive× additive epistatic gene action. *Crop science* 43: 489-497.

Jannink J.-L. 2010. Dynamics of long-term genomic selection. *Genetics Selection Evolution* 42

Jannink J.L., Lorenz A.J. & Iwata H. 2010. Genomic selection in plant breeding: from theory to practice. *Brief Funct Genomics* 9: 166-177.

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

Kilian B., Ozkan H., Kohl J., von Haeseler A., Barale F., Deusch O., Brandolini A., Yucel C., Martin W. & Salamini F. 2006. Haplotype structure at seven barley genes: relevance to gene pool bottlenecks, phylogeny of ear type and site of barley domestication. *Mol Genet Genomics* 276: 230-241.

Komatsuda T., Pourkheirandish M., He C., Azhaguvel P., Kanamori H., Perovic D., Stein N., Graner A., Wicker T. & Tagiri A. 2007. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proceedings of the National Academy of Sciences* 104: 1424-1429.

Lado B., Matus I., Rodríguez A., Inostroza L., Poland J., Belzile F., del Pozo A., Quincke M., Castro M. & von Zitzewitz J. 2013. Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data. *G3* 3: 2105-2114.

Lander E.S. & Botstein D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185-199.

Lane D. & Salkin N. 2010. Tukey's honestly significant difference (HSD). *N. Salkind der.), Encyclopedia of Research Design içinde, Thousand Oaks, CA: SAGE Publications* 1: 1566-1571.

Lauer J.G., Bijl C.G., Grusak M.A., Baenziger P.S., Boote K., Lingle S., Carter T., Kaeppler S., Boerma R. & Eizenga G. 2012. The scientific grand challenges of the 21st century for the Crop Science Society of America. *Crop science* 52: 1003-1010.

Li Y., Dong Y., Niu S., Cui D., Wang Y., Liu Y., Wei M. & Li X. 2008. Identification of agronomically favorable quantitative trait loci alleles from a dent corn inbred Dan232 using advanced backcross QTL analysis and comparison with the F2: 3 population in popcorn. *Molecular Breeding* 21: 1-14.

Lian L, Jacobson A, Zhong S. & Bernardo R. 2015. Prediction of genetic variance in biparental maize populations: genomewide marker effects versus mean genetic variance in prior populations. *Crop Sci*. 55(3):1181–1188.

Lin C.-S. & POUSHINSKY G.R.E.G. 1985. A modified augmented design (type 2) for rectangular plots. *Canadian journal of plant science* 65: 743-749.

Lipka A.E., Tian F., Wang Q., Peiffer J., Li M., Bradbury P.J., Gore M.A., Buckler E.S. & Zhang Z. 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28: 2397-2399.

Lorenz A.J. & Smith K.P. 2015. Adding Genetically Distant Individuals to Training Populations Reduces Genomic Prediction Accuracy in Barley. *Crop Sci.* 55: 2657.

Lorenz A.J., S. Chao, F.G. Asoro, E.L. Heffner, T. Hayashi, H. Iwata, K.P. Smith, M.E. Sorrells & J.-L. Jannink. 2011. Genomic Selection in Plant Breeding(ed.) 77-123.

Lorenz A.J., Smith K.P. & Jannink J.-L. 2012. Potential and Optimization of Genomic Selection for Fusarium Head Blight Resistance in Six-Row Barley. *Crop Sci.* 52: 1609.

Lorenzana R.E. & Bernardo R. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theor. Appl. Genet.* 120: 151-161.

Ma Z., Steffenson B.J., Prom L.K. & Lapitan N.L.V. 2000. Mapping of quantitative trait loci for Fusarium head blight resistance in barley. *Phytopathology* 90: 1079-1088.

Martin J.M., Blake T.K. & Hockett E.A. 1991. Diversity among North American spring barley cultivars based on coefficients of parentage. *Crop Science* 31: 1131-1137.

Mascher M., Muehlbauer G.J., Rokhsar D.S., Chapman J., Schmutz J., Barry K., Muñoz Amatriaín M., Close T.J., Wise R.P. & Schulman A.H. 2013. Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). *The Plant Journal* 76: 718-727.

Massman J., Cooper B., Horsley R., Neate S., Dill-Macky R., Chao S., Dong Y., Schwarz P., Muehlbauer G.J. & Smith K.P. 2011. Genome-wide association mapping of Fusarium head blight resistance in contemporary barley breeding germplasm. *Mol. Breeding* 27: 439-454.

Massman J.M., Gordillo A., Lorenzana R.E. & Bernardo R. 2013a. Genomewide predictions from maize single-cross data. *Theor. Appl. Genet.* 126: 13-22.

Massman J.M., Jung H.-J.G. & Bernardo R. 2013b. Genomewide Selection versus Marker-assisted Recurrent Selection to Improve Grain Yield and Stover-quality Traits for Cellulosic Ethanol in Maize. *Crop Sci.* 53: 58.

Mayor P.J. & Bernardo R. 2009. Genomewide Selection and Marker-Assisted Recurrent Selection in Doubled Haploid versus F Populations. *Crop Sci.* 49: 1719-1725.

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

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

Minnesota Agricultural Experiment Station (MAES). 2015. "2015 Barley Field Crop Trials Results." https://www.maes.umn.edu/sites/maes.umn.edu/files/barley_2015_final.pdf

Mirocha C.J., Kolaczkowski E., Xie W., Yu H. & Jelen H. 1998. Analysis of deoxynivalenol and its derivatives (batch and single kernel) using gas chromatography/mass spectrometry. *Journal of Agricultural and Food Chemistry* 46: 1414-1418.

Mohammadi M., Tiede T. & Smith K.P. 2015. PopVar: A Genome-Wide Procedure for Predicting Genetic Variance and Correlated Response in Biparental Breeding Populations. *Crop Sci.* 55: 2068.

Moncada P., Martinez C.P., Borrero J., Châtel M., Gauch Jr H., Guimaraes E., Tohme J. & McCouch S.R. 2001. Quantitative trait loci for yield and yield components in an Oryza sativa× Oryza rufipogon BC2F2 population evaluated in an upland environment. *Theoretical and Applied Genetics* 102: 41-52.

Nagata K., Ando T., Nonoue Y., Mizubayashi T., Kitazawa N., Shomura A., Matsubara K., Ono N., Mizobuchi R. & Shibaya T. 2015. Advanced backcross QTL analysis reveals complicated genetic control of rice grain shape in a japonica× indica cross. *Breeding science* 65: 308.

Nakaya A. & Isobe S.N. 2012. Will genomic selection be a practical method for plant breeding. *Ann Bot* 110: 1303-1316.

Nei M. 1974. A new measure of genetic distance(ed.) *Genetic distance*: 63-76. Springer.

Nice L.M., Steffenson B.J., Brown-Guedira G.L., Akhunov E.D., Liu C., Kono T.J.Y., Morrel P.L., Horsley R.D., Smith K.P., Muehlbauer G.J. Development and genetic characterization of an Advanced Backcross – Nested Association Mapping (AB-NAM) population of wild x cultivated barley. *Genetics*. In press.

Nice L.M. 2015. Assessing wild barley germplasm in multiparent and advanced backcross populations for mapping, gene discovery, and improvement of malting barley (Doctoral dissertation). Received through personal communication with L.M. Nice.

NSGC. http://www.ars.usda.gov/Main/docs.htm?docid=21891. April 3, 2016.

Piepho H.-P. & Möhring J. 2007. Computing heritability and selection response from unbalanced plant breeding trials. *Genetics* 177: 1881-1888.

Pillen K., Zacharias A. & Léon J. 2003. Advanced backcross QTL analysis in barley (Hordeum vulgare L.). *Theoretical and Applied genetics* 107: 340-352.

Pourkheirandish M. & Komatsuda T. 2007. The importance of barley genetics and domestication in a global perspective. *Annals of Botany* 100: 999-1008.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Rao R.B., Fung G. & Rosales R. 2008. On the Dangers of Cross-Validation. An Experimental Evaluation. SDM: 588-596.

Rasmusson D.C. & Phillips R.L. 1997. Plant breeding progress and genetic diversity from de novo variation and elevated epistasis. *Crop Science* 37: 303-310.

Rasmusson D.C., Smith K.P., Dill-Macky R., Schiefelbein E.L. & Wiersma J.V. 2001. Registration of Lacey barley. *Crop science* 41: 1991-1991.

Ren X., Nevo E., Sun D. & Sun G. 2013. Tibet as a potential domestication center of cultivated barley of China. *PloS one* 8: e62700.

Riedelsheimer C., Endelman J.B., Stange M., Sorrells M.E., Jannink J.L. & Melchinger A.E. 2013. Genomic predictability of interconnected biparental maize populations. *Genetics* 194: 493-503.

Rincent R., Laloë D., Nicolas S., Altmann T., Brunel D., Revilla P., Rodríguez V.M., Moreno-Gonzalez J., Melchinger A., Bauer E., Schoen C.C., Meyer N., Giauffret C., Bauland C., Jamin P., Laborde J., Monod H., Flament P., Charcosset A. & Moreau L. 2012. Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (Zea mays L.). *Genetics* 192: 715-728.

Robbins K.R., Backlund J.E. & Schnelle K.D. 2012. Spatial Corrections of Unreplicated Trials using a Two-dimensional Spline. *Crop Sci.* 52: 1138.

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. *The Plant Genome* 8: 0.

Saal B., von Korff M., Léon J. & Pillen K. 2011. Advanced-backcross QTL analysis in spring barley: IV. Localization of QTL × nitrogen interaction effects for yield-related traits. *Euphytica* 177: 223-239.

Sallam A.H. 2014. Assessing Genomic Selection Prediction Accuracy in a Dynamic Barley Breeding Population and Comparing Gain between genomic and Phenotypic Selection in Barley (Doctoral dissertation). Retrieved from http://hdl.handle.net/11299/170192

Sallam A.H., Endelman J.B., Jannink J.-L. & Smith K.P. 2015. Assessing Genomic Selection Prediction Accuracy in a Dynamic Barley Breeding Population. *The Plant Genome* 8.

Sang T. 2009. Genes and mutations underlying domestication transitions in grasses. *Plant Physiol* 149: 63-70.

Sayed M.A., Schumann H., Pillen K., Naz A.A. & Léon J. 2012. AB-QTL analysis reveals new alleles associated to proline accumulation and leaf wilting under drought stress conditions in barley (Hordeum vulgare L.). *BMC genetics* 13: 61.

Schmalenbach I., Léon J. & Pillen K. 2009. Identification and verification of QTLs for agronomic traits using wild barley introgression lines. *Theoretical and Applied Genetics* 118: 483-497.

Schwarz G. 1978. Estimating the dimension of a model. *The annals of statistics* 6: 461-464.

Segura V., Vilhjálmsson B.J., Platt A., Korte A., Seren Ü., Long Q. & Nordborg M. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature genetics* 44: 825-830.

Septiningsih E.M., Pamplona A.M., Sanchez D.L., Neeraja C.N., Vergara G.V., Heuer S., Ismail A.M. & Mackill D.J. 2009. Development of submergence-tolerant rice cultivars: the Sub1 locus and beyond. *Annals of Botany* 103: 151-160.

Smith K.P., Budde A., Dill-Macky R., Rasmusson D.C., Schiefelbein E., Steffenson B., Wiersma J.J., Wiersma J.V. & Zhang B. 2013. Registration of 'Quest' Spring Malting Barley with Improved Resistance to Fusarium Head Blight. *Journal of Plant Registrations* 7: 125.

Smith K.P., Rasmusson D.C., Schiefelbein E., Wiersma J.J., Wiersma J.V., Budde A., Dill-Macky R. & Steffenson B. 2010. Registration of 'Rasmusson 'barley. *Journal of plant registrations* 4: 167-170.

Soto-Cerda B.J., Peñaloza E.H., Montenegro A.B., Rupayan A.R., Gallardo M.H. & Salvo-Garrido H. 2013. An efficient marker-assisted backcrossing strategy for enhancing barley (Hordeum vulgare L.) production under acidity and aluminium toxicity. *Mol Breeding* 31: 855-866.

Spindel J., Begum H., Akdemir D., Virk P., Collard B., Redoña E., Atlin G., Jannink J.-L. & McCouch S.R. 2015. Genomic Selection and Association Mapping in rice (Oryza sativa): Effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. *PLoS Genet* 11: e1004982.

Steffenson B.J. 2003. Fusarium head blight of barley: impact, epidemics, management, and strategies for identifying and utilizing genetic resistance. In: Leonard KL, Bushnell WR (eds) Fusarium head blight of wheat and barley. *The American Phytopathological Society*, St. Paul, pp 241–295.

Steffenson B.J., Olivera P., Roy J.K., Jin Y., Smith K.P. & Muehlbauer G.J. 2007. A walk on the wild side: mining wild wheat and barley collections for rust resistance genes. *Crop and Pasture Science* 58: 532-544.

Stuthman D.D. 2002. Contribution of durable disease resistance to sustainable agriculture. *Euphytica* 124: 253-258.

Tanksley S.D. & McCouch S.R. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277: 1063-1066.

Tanksley S.D. & Nelson J.C. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theoretical and Applied Genetics* 92: 191-203.

Tanksley S.D. 1997. Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. *Science* 277: 1063-1066.

Tayeh N., Klein A., Le Paslier M.C., Jacquin F., Houtin H., Rond C., Chabert-Martinello M., Magnin-Robert J.B., Marget P., Aubert G. & Burstin J. 2015. Genomic Prediction in Pea: Effect of Marker Density and Training Population Size and Composition on Prediction Accuracy. *Front Plant Sci.* 6: 941.

Technow F. & Totir L.R. 2015. Using Bayesian Multilevel Whole Genome Regression Models for Partial Pooling of Training Sets in Genomic Prediction. *G3 (Bethesda)* 5: 1603-1612.

Technow, F. 2015. R package mvngGrAd: moving grid adjustment in plant breeding field trials. R package version 0.1.5. http://CRAN.R-project.org/package=mvngGrAd.

Thomas W.T.B., W. Powell, R. Waugh 1995. Detection of quantitative trait loci for agronomic, yield, grain and disease characters in spring barley (Hordeum vulgare L.). *Theor Appl Genet* 1037.

Tiede, T. & Smith K.P. Evaluating genomewide marker-based breeding methods in traditional and wild relative-derived barley populations (Doctoral dissertation). *In preparation*.

VanRaden P.M. 2008. Efficient methods to compute genomic predictions. *J Dairy Sci* 91: 4414-4423.

Vargas M., Combs E., Alvarado G., Atlin G., Mathews K. & Crossa J. 2013. META: A suite of SAS programs to analyze multienvironment breeding trials. *Agronomy Journal* 105: 11-19.

Villumsen T.M., Janss L. & Lund M.S. 2009. The importance of haplotype length and heritability using genomic selection in dairy cattle. *J Anim Breed Genet* 126: 3-13.

Von Korff M., Wang H., Léon J. & Pillen K. 2006. AB-QTL analysis in spring barley: II. Detection of favourable exotic alleles for agronomic traits introgressed from wild barley (H. vulgare ssp. spontaneum). *Theoretical and Applied Genetics* 112: 1221-1231.

Waldron B.L., Moreno-Sevilla B., Anderson J.A., Stack R.W. & Frohberg R.C. 1999. RFLP mapping of QTL for Fusarium head blight resistance in wheat. *Crop Science* 39: 805-811.

Warnes G., G. Gorjanc, F. Leisch & M. Man. (2013). genetics: Population Genetics. R package version 1.3.8.1. http://CRAN.R-project.org/package=genetics.

Warschefsky E., Penmetsa R.V., Cook D.R. & von Wettberg E.J. 2014. Back to the wilds: tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. *Am J Bot* 101: 1791-1800.

Wu J.-L., Sinha P.K., Variar M., Zheng K.-L., Leach J.E., Courtois B. & Leung H. 2004. Association between molecular markers and blast resistance in an advanced backcross population of rice. *Theoretical and Applied Genetics* 108: 1024-1032.

Xiao J., Li J., Grandillo S., Ahn S.N., Yuan L., Tanksley S.D. & McCouch S.R. 1998. Identification of trait-improving quantitative trait loci alleles from a wild rice relative, Oryza rufipogon. *Genetics* 150: 899-909.

Xu K. & Mackill D.J. 1996. A major locus for submergence tolerance mapped on rice chromosome 9. *Molecular Breeding* 2: 219-224.

Xu K., Xu X., Fukao T., Canlas P., Maghirang-Rodriguez R., Heuer S., Ismail A.M., Bailey-Serres J., Ronald P.C. & Mackill D.J. 2006. Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* 442: 705-708.

Yun S.J., Gyenis L., Bossolini E., Hayes P.M., Matus I., Smith K.P., Steffenson B.J., Tuberosa R. & Muehlbauer G.J. 2006. Validation of quantitative trait loci for multiple disease resistance in barley using advanced backcross lines developed with a wild barley. *Crop Science* 46: 1179-1186.

Zamir D. 2001. Improving plant breeding with exotic genetic libraries. *Nature reviews genetics* 2: 983-989.

Zhang Z., Ersoz E., Lai C.-Q., Todhunter R.J., Tiwari H.K., Gore M.A., Bradbury P.J., Yu J., Arnett D.K. & Ordovas J.M. 2010. Mixed linear model approach adapted for genome-wide association studies. *Nature genetics* 42: 355-360.

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: 567-576.