MOLECULAR MARKERS ASSOCIATED WITH BARLEY YELLOW DWARF VIRUS TOLERANCE IN SPRING OAT AND THEIR UTILIZATION IN PREDICTIVE BREEDING

BY

BRADLEY JOSEPH FORESMAN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Crop Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Frederic L. Kolb, Chair
Professor Brian W. Diers
Assistant Professor Leslie Domier
Assistant Professor Patrick J. Brown

Abstract

Barley yellow dwarf viruses (BYDVs) are responsible for the disease barley yellow dwarf (BYD), which causes significant yield losses in many cereals including oat (*Avena sativa* L.). Phenotyping for disease sensitivity is time consuming, laborious and requires viruliferous aphids for inoculations. Until recently, the molecular marker technology in oat has not allowed for many marker-trait association studies to determine the genetic mechanisms for tolerance and as a result, marker-assisted selection (MAS) and genomic selection (GS) have not been extensively used in breeding for BYD tolerance. In the first study, a genome-wide association study (GWAS) was performed on 428 spring oat lines using a recently developed high-density oat single nucleotide polymorphism (SNP) array as well as a SNP-based consensus map. Marker-trait associations were performed using a Q-K mixed model approach to control for population structure and relatedness. Six significant SNP-trait associations representing two QTL were found on chromosomes 3C and 18D. This is the first report of BYDV tolerance QTL on chromosome 3C and 18D. Haplotypes using the two QTL were evaluated, and distinct classes for tolerance were identified based on the number of favorable alleles. In the second study, GS and MAS models were compared in their accuracy to predict barley yellow dwarf virus tolerance in 428 spring oat lines from North America and Europe. Several GS models were evaluated using 2305 SNPs and included models with previously identified or randomly selected markers as fixed effects. Model accuracies were evaluated using five-fold cross evaluation. GS models used ridge regression-best linear unbiased predictor (RR-BLUP) for marker effect estimation while MAS models used ordinary least square (OLS). Moderate to high prediction accuracies (0.5-0.9) were observed across the models. GS models containing fixed effects (GS-

GWAS, GS-3C18D) from previously identified QTL performed better than the GS model with all markers as random effects or the MAS models. Two MAS models (MAS-GWAS and MAS-3C18D) had prediction accuracies higher than the GS model with all markers as random effects. In the third study, GS was used to identify individuals with high BYDV tolerance for use in cross prediction. To do this, 2138 SNPs were used on a panel of 519 spring oat lines for barley yellow dwarf virus tolerance. Of the 519 oat lines, 428 lines had genotypes and phenotypes while 91 of the oat lines were only genotyped. Using the R package "PopVar", several GS models were compared for prediction accuracy. The BayesA model was identified as having the highest prediction accuracy and genomic estimated breeding values (GEBVs) were calculated for the 519 lines using the BayesA model. The top 10% of lines (52 lines) based on GEBVs for BYDV tolerance were selected to perform simulated crosses. A total of 1326 crosses were simulated, and the mean, genetic variance and mean of high/low superior progenies were calculated. From the 1326 crosses, 22 crosses were identified as having a balance between a low predicted mean (high tolerance) and high genetic variance. Because of the high tolerance and high genetic variance, the chance of obtaining transgressive segregants for BYDV tolerance is higher in these crosses. Using GS and simulated crosses gives breeders additional tools to improve breeding efficiency for BYDV tolerance and allows for better allocation of time and resources within the breeding program.

Acknowledgements

I first must acknowledge the funding for this research from the USDA-ARS, Quaker Oat, General Mills, North American Millers' Association (NAMA) and the University of Illinois at Urbana-Champaign/Illinois Crop Improvement Association J.C. Hackleman graduate fellowship.

To the research specialists involved with this project, Norman Smith and Allison Krill-Brown, your help with all of these projects was invaluable.

To my committee members, I want to thank you for your time, edits, support with this research and guidance through the process.

To my advisor, Dr. Fred Kolb, your guidance, support and criticism has helped me become a better scientist. I am grateful for your friendship, mentorship and giving me the opportunity to continue my education in your program.

To my friends and lab-mates Marcio Pais de Arruda and Olivia Jones, thank you for friendship, as well as your advice, criticisms and help throughout this research.

To my family, thank you for all of the love and support along this journey. This journey began many years ago and your advice not only in my education but also in my career search greatly helped me.

To my parents, Joe Foresman and the late Paula Foresman-Futer, thank you for instilling in me the drive to always be a better person in life and in academics, for instilling in me the value of hard work, and for the love and support throughout my life.

To my wife Erika, thank you for all of the support in our relationship and my journey through school. You have been by my side through it all and I cannot thank you enough for what you have done and the sacrifices you have made. Without you, this would not have been possible.

TABLE OF CONTENTS

CHAPTER 1

LITERATURE REVIEW

1.1 Introduction

Oat (*Avena sativa L.*) is an important cereal grain grown worldwide for consumption by both humans and livestock. Oat once ranked as high as sixth in the world in grain crop production, but production is trending downward as more and more producers place emphasis on crops such as corn and soybean (Marshall and Sorrells, 1992). In the U.S., oats are grown on over 1.2 million hectares (3 million acres) and in 2015, production totaled over 1.3 million tonnes (89 million bushels) (USDA, 2015). Because of the nutrition benefit of oats, livestock such as cattle and horses are regularly fed oats as part of their diet. In fact 75% of the consumption of oat are for livestock feed while 22% is used for human consumption and seed (Marshall and Sorrells, 1992). Oats are also commonly used as a cover crop. It is important to protect and maintain the worldwide supply of oats because of their contribution to the diets of both humans and livestock.

Barley Yellow Dwarf (BYD) is acknowledged to be one of the most destructive diseases of cereal crops worldwide and can have serious economic impacts on cereal grain production (Burnett and D'Arcy, 1995). Barley yellow dwarf viruses (BYDVs) are the causal agents of BYD which was first described in barley by Oswald and Houston in 1951, but also infects other grass species including oat. Symptoms tend to depend on cultivar and environment but normally include stunted growth, leaf discoloration, inhibition of root formation, and blasting of florets

(Jensen and D'Arcy, 1995). The discoloration in oat is usually yellowing or reddening of the leaf blades, especially at the leaf tips (Kosova et al., 2008). These changes in appearance are due to the reduction of chlorophyll and photosynthesis. The resulting reduction in the production of photosynthates leads to substantial yield loss. Root growth inhibition prevents the uptake of water and nutrients and may also lead to yield loss (Kolb et. al., 1991). The virus can also lead to the degradation of the phloem and the collapse of the sieve elements. Infected plants with unripe grains also become more sensitive to fungal infections further reducing yield (Jensen and D'Arcy, 1995). Reductions in photosynthesis and root growth inhibition are primary causes for grain yield loss.

In Illinois, the BYDV-PAV-IL isolate was estimated to cause 0.34 to 0.55% yield loss for each percent increase in virus infection of wheat (Perry et al., 2000). There has been little work done on the relationship between BYDV infection and yield loss in oat. One experiment by McKirdy et al. (2002) determined that yield losses in oat ranged from 13-25 kg/ha for each 1% increase in incidence of BYDV. Bauske et. al. (1997) determined a yield reduction of 4.5% for every 10% increase in incidence.

BYD is caused by a group of luteoviruses known as BYDVs. They are phloem limited viruses that are obligately transmitted to the grasses via their aphid vectors. Because of the phloem limitation the virus is found in low concentrations and infects only a few cells per plant. The virus particles are icosahedra and range from 22-25 nm. Aphids must feed on infected plants in order to transmit BYDV to other plants because the virus does not replicate in the aphid and does not persist transovarially (Hewings, 1995).

There are at least 25 aphid species that transmit BYDV, and the strains of BYDV are distinguished by the species of aphid vector (Halbert and Voegtlin, 1995). Historically,

taxonomy of these viruses has been based on the aphid specificity, their range of hosts, and cross protection of various strains. These specific virus strains have now been classified into two genera based on genome structure. BYDVs PAV (*Rhopalosiphum padi* and *Macrosiphum avenae), MAV (Macrosiphum avenae),* and SGV (*Schizaphis graminum*) are in the genus *Luteovirus*. Cereal yellow dwarf virus RPV (*Rhopalosiphum padi*; formerly known as barley yellow dwarf virus RPV) and Maize yellow dwarf virus RMV (*Rhopalosiphum maidis*) are in the genus *Polerovirus*. Viruses in the genus *Luteovirus* have replication-related protein related to members of *Tombusviridae*, while viruses in the genus *Polereovirus* have replication-related proteins similar to those of *sobemoviruses* (D'Arcy and Mayo, 1997).

The most common vector of BYDV in North America is the aphid genus *Rhopalosiphum*. This genus of aphid is primarily holocyclic and prefers damp environments. There are two main species of *Rhopalosiphum* that vector BYDV, *Rhopalosiphum maidis* and *Rhopalosiphum padi*. *Rhopalosiphum padi* is the primary vector for BYDV in Illinois. This vector overwinters anholocyclically when there are mild winters or when there is snow protection. During the fall, the aphids can be found at or below ground level. They may also colonize at leaf bases, stems, and heads (Halbert and Voegtlin, 1995).

The transmission of BYDVs involves three types of interactions. These include the direct interaction between the virus and vector, the direct interaction between the vector and host plant, and the indirect interactions between the virus and vectors that occur because of physiological changes in the host plant. Examining the direct interactions between virus and vector allow for determining the transmission phenotype (Power and Gray, 1995). There are many factors that influence the transmission phenotype. These include, variation in BYDV transmission within an aphid species, length of acquisition and inoculation access period, transmission efficiency related

to virus titer, environmental factors (such as temperature), and adaptability of BYDV transmission phenotype.

There are also many mechanisms that regulate the virus acquisition and transmission by aphids. BYDV is acquired by the aphids through feeding on infected plants. The virus is drawn into the food canal with the phloem sap. It then passes through the foregut, anterior and posterior midgut and then into the hindgut where it can be acquired into the aphid hemoceol. It is then believed to be taken up by cells through a receptor-mediated uptake method by the hindgut epithelial cells (Power and Gray, 1995). The virus is then transmitted to plants, most likely through the salivary system and more specifically the accessory salivary gland (Rochow, 1969).

There are also indirect virus-to-vector interactions, specifically within the host plant. It has been hypothesized that that the yellowing symptoms of leaves may lead to aphid attraction to infected plants (Kring, 1972). Feeding behavior may also be affected when feeding on infected plants. In fact, aphids increase speed of ingestion, and feed for longer amounts of time when feeding on infected plants. Kennedy et al. (1951) showed that there may be mutualism between virus and vector. The virus improves suitability of the plant for the vector and in turn the vector distributes the virus. However, this mutualism is diffuse at best and was most likely not strongly coevolved.

There are many characteristics or symptoms of BYD that can be evaluated; however, there is no one symptom that is an effective indicator of susceptibility (Endo and Brown, 1964). Traditionally, a 0 to 9 scale is used as a visual evaluation based on stunting, blasting and chlorosis (Qualset, 1984; Hewings et al., 1992). Using this scale, a rating of 0 would be a completely healthy plant, while a rating of 9 would be severely affected, mostly blasted and

highly stunted. The 0 to 9 scale can be used with most small grains including oat. Several other symptoms and characteristics of BYD can be evaluated for assessment of germplasm. Third leaf length has been reported to be longer in tolerant oat plants than in sensitive plants (Gellner et al. 1992).

There are two main methods used for the control to BYDVs. Since the virus is persistently spread by aphids, insecticides can be effective in controlling aphid populations. Synthetic pyrethroids and imidocloprid seed treatments significantly decrease BYDV incidence and thus increase grain yield (McKirdy and Jones, 1996). Organophosphates and carbamates have also been developed for aphid control (Mann, 1991). Combining seed treatment with imidocloprid and two foliar applications have been shown to decrease BYD by 88% and increase grain yield up to 76%; however, this is only economically feasible in highly intensive agricultural systems. The most practical form of control is the use of tolerant or resistance lines (McKirdy and Jones, 1996; Gourmet, 1996).

There has been much discussion of tolerance vs. resistance with BYDVs. Two types of resistance to BYDV have been distinguished; virus resistance and field resistance. Virus resistance is when there is a low virus titer in infected plants whereas field resistance is the reduction of symptoms of infection independent of the virus titer. Field resistance is usually referred to as tolerance (Kosova et. al 2008). In this dissertation, resistance will be defined as reduced viral replication in infected plants (Cooper and Jones, 1983). Tolerance will therefore be defined as the development of mild or negligible symptoms in infected plants. It can also be stated as the ability of plants to yield under BYDV infection. It is also important to recognize that the mechanisms of tolerance and resistance are not yet known.

Several quantitative trait loci (QTL) for BYD tolerance were identified in oat by Jin et al. (1998) and by Barbosa-Neto et al. (2000). It has been proposed that two to four genes are responsible for BYD tolerance (McKenzie et al. 1985). Zhu et al. (2003) examined BYD tolerance in cultivated oat by crossing the tolerant variety 'Ogle' with the sensitive line MAM17-5. A total of 272 RFLP, SSR and AFLP markers were used in the identification of QTL. Four QTL (BYDq1, BYDq2, BYDq3 and BYDq4) for BYD tolerance were identified and were found on linkage groups OM1, 5, 7 and 24, respectively. A significant epistatic effect was also found between some of the QTL. Their final model (including epistatic effects) explained 50.3-58.2% of the total variation for BYD tolerance. They also reported that some of the QTL for BYD tolerance were closely linked to QTLs for plant height and days to heading. Several QTL, including two large effect QTL were identified in two bi-parental recombinant inbred line (RIL) populations. The major QTL were identified on chromosomes 3C and 19A (Foresman, 2014). An update to the published consensus map would result in the relocation of the 19A QTL location to chromosome 18D (Chaffin et. al., 2015, submitted to The Plant Genome).

Over the last few decades molecular marker technology has increased substantially, especially in crops such as maize and soybeans. Progress in oat, however, has been much slower because of the lack of investment in the crop. Various marker analysis methods have been developed and are continuing to develop each year. The type of markers used have changed from restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), randomly amplified polymorphic DNA (RAPD), simple sequence repeats (SSR) to single nucleotide polymorphisms (SNPs) (Bolstein et al., 1980, Chee et al. 1996; Vos et al., 1995; Weber and May, 1989; Williams et al., 1990). SSR and SNP markers with

coverage wide enough for fingerprinting studies have only recently become available in oat (Wight et al. 2010; Oliver et al., 2010; Oliver et al., 2011).

One viable option that has been used in oat and other crops is the use of a new marker platform called Diversity Array Technology (DArT). This high throughput technique is very useful because it does not require sequence information. A genomic representation is developed from amplified restriction fragments that can then be analyzed by the presence/absence of a particular clone. A microarray platform can then simultaneously type several thousand loci for an efficient genomic analysis (Jaccoud et al., 2001). These DArT markers have provided enhanced map coverage of the oat genome. Tinker et al. (2009) found more than 2000 polymorphic markers used in a global diversity study and identified 2700 potential polymorphic markers for future studies; however, the reliability of DArT marker data has been questionable and therefore may not be the best option for genotyping.

Recently, a high density oat single-nucleotide polymorphism (SNP) array containing 6000 SNPs was developed. The genotyping assay was performed using an Infinium assay developed by Illumina. These SNPs along with the previously mentioned markers were recently published in a physically anchored oat consensus map (Oliver et. al. 2013). This new SNP array has been reported to be extremely reliable and shows potential for being useful in breeding programs. Genotyping-by-sequencing has recently begun to be used in oat and appears to be how future genotyping will be performed (Huang, 2014).

The goal of QTL mapping is to identify genomic regions associated with traits of interest. The selection of identified regions in breeding programs can then be used in improving genetic gain for the trait (Lande and Thompson, 1990). One approach to QTL mapping is through linkage studies of previously developed populations (i.e. backcross, recombinant inbred etc.).

This approach (family mapping) is powerful in detecting QTL and does not require high marker densities; however, many QTL identified in these populations tend to be population specific (Jannink et al., 2001). Another approach would be to perform genome-wide association studies (GWAS). In this approach, a diverse population of lines is used (Zhu et al., 2008). This method requires a larger number and higher density of markers than in family mapping. Another issue that arises with GWAS is that the population structure in the panel of lines used as well as the relationship of the lines can affect the results obtained. These two issues can lead to spurious marker-trait associations (Kennedy et al., 1992). Because of this, controlling for both structure and the relationship of lines is important when performing GWAS, and research has shown these to be important to control in oat (Newell, et al., 2011). GWAS has been performed on several traits in oat including crown rust (*Puccinia cornata* f. sp. *avenae*), powdery mildew (*Blumeria graminis* f. sp. *avenae*), and beta-glucan concentration, (Montilla-Bascón et. al., 2015; Asoro et al., 2013; Newell et. al., 2012). The results of both family mapping and GWAS can be used in breeding programs through the use of marker-assisted selection (MAS) or genomic selection (GS).

MAS and GS, along with high density and high throughput marker platforms (SNP arrays or genotyping-by-sequencing) provide breeders with the tools to increase future genetic gains. With improvements in the understanding of the genetic architecture of oat, both technologies will likely become important tools for breeders to use in developing tolerant cultivars and should help breeding programs increase efficiency of their programs.

Strong marker-trait associations identified from QTL studies can be used in MAS. The central idea of MAS is to use the identified markers (usually from a combination of linkage mapping or genome-wide association studies) for indirect selection of the trait of interest (Ribaut

and Hoisington, 1998; Collard and Mackill, 2007). It can be particularly useful for traits that are difficult or destructive when phenotyped or have low heritabilities. Selection can also be performed in the seedling stage and undesirable genotypes can be discarded, saving time and resources throughout the breeding pipeline. MAS also provides an efficient method for back-crossing (marker-assisted backcrossing) specific genes into elite breeding lines. Gnanesh et. al. (2013) provides an example of the development of markers for use in a MAS system for crown rust resistance in oat. For complex traits (multiple small effect QTL), MAS may not be as successful as simple or single gene traits (Bernardo, 2008).

Genomic Selection (GS) can be viewed as an extension of marker-assisted selection where all marker information is used to calculate the genomic estimated breeding values (GEBV). As mentioned above, MAS uses only the markers associated with significant QTL, whereas GS uses all loci without consideration of the effect of each locus. This concept was first proposed in animal breeding by Meuwissen et al. (2001) and because all markers are used, the potential for all of the genetic variance can be explained by the markers. GS requires a large number of markers distributed across the genome, and with the advancements in marker technology, the increase in number and decrease in cost of markers has allowed for greater implementation in plant breeding programs.

Genomic selection models use the markers as the predictor variables (p) and the phenotypes as the response variable (n). With the increasing number of available markers, the number of markers often exceed the number of phenotypes (large p small n problem). Because of this, models that depend on shrinkage factors are needed to estimate the marker effects. There are several genomic selection models that can be used. Examples include ridge regression best linear unbiased predictor (RR-BLUP), genomic best linear unbiased predictor (G-BLUP), least

absolute shrinkage and selection operator (LASSO), elastic-net and several Bayesian models (BayesA, BayesB, BayesC) (Lorenz et al., 2011; Heslot et al, 2012).

The most popular model used is RR-BLUP where the markers are assumed to be from the same distribution and assumed to have a common variance. Because of this, the markers are equally shrunken towards zero (Bernardo and Yu, 2007; Heffner et al., 2009). Although, this is not truly realistic in plants, there are many examples of using this model effectively (Arruda et al., 2015; Lipka et al., 2014). The Bayesian models differ from the RR-BLUP models by allowing different variances for the predictor variables. This allows each marker to be shrunken towards zero to different degrees. BayesA, BayesB and BayesC are three slightly different types of Bayesian models with BayesA being the most basic model that allows for different variances for each predictor variable. BayesB is similar to BayesA, but each marker is given a probability that the effect is zero. BayesC uses the data to estimate the probability that each marker effect is zero.

Certain models are better suited for specific genetic architectures, traits and populations. When a trait is controlled by many small effect loci, RR_BLUP works very well. Under conditions where a few large-effect QTL explain the majority of the genetic variation, BayesB works well (Lorenz et al., 2011). In regards to the genetic control of BYDV tolerance in spring oat, previous research has shown that a few large-effect QTL seem to control the trait, and therefore, BayesB may be the best model for genomic selection.

Along with model selection, the development of a training population is an important step in the genomic selection process. The training population is used to estimate the model parameters and consists of individuals with phenotypic and genotypic data. The model can then be applied to a validation population (genotypes only) to produce GEBVs. Accuracy of the

prediction models can be measured by the correlation between the GEBV and the phenotypically estimated breeding value (PEBV), $cor(GEBV, PEBV)/\sqrt{h^2}$ (Lorenz et al., 2011).

GS can also be used for predictive breeding through the evaluation of simulated crosses based on marker profiles. By simulating potential crosses, or examining all possible crosses from the panel of lines, breeders can inspect the best crosses to make based on the marker data. Breeders can then check to see if certain crosses have already been made in the program or need to be made in the next breeding cycle. This can be extremely beneficial for breeding programs by allowing for more directed crosses for a specific trait or disease resistance. The ability to predict GEBVs for lines that have been genotyped and not phenotyped can also add valuable insight into the potential of the lines without the time and resource commitment required to phenotype them early in the breeding program. As a result undesirable lines can be discarded before they enter extensive testing.

Current phenotypic breeding methods for BYDV involve screening large numbers of breeding lines by inoculating with viruliferous aphids followed by rating of the disease symptoms. This phenotyping is time consuming and laborious and requires the availability of viruliferous aphids. Breeding lines are frequently not screened until the later stages the breeding program. Identifying QTL for BYDV is important for the use in a MAS or GS program. The ability to molecularly screen plants earlier in the breeding program through MAS or GS would allow breeders to increase efficiency and make more progress in developing tolerant BYDV cultivars.

1.2 References

Arruda, M.P., P.J. Brown, A.E. Lipka, A.M. Krill, C. Thurber and F.L. Kolb. 2015. Genomic selection for predicting fusarium head blight resistance in a wheat breeding program. The Plant Genome. 8.3:1-12.

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

Barbosa-Neto, J. F., W. Siripoonwiwat, L. S. O'Donoughue, S. M. Gray, D. M. Smith, F. L. Kolb, C. Gourmet, C. M. Brown and M. E. Sorrells. 2000. Chromosomal regions associated with barley yellow dwarf virus resistance in oat. Euphytica. 114:67-76.

Bauske, E.M., S. M. Bissonnette, and A.D. Hewings. 1997. Yield loss assessment of barley yellow dwarf disease on spring oat in Illinois. Plant Disease. 81.5:485-8.

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. and J. Yu. 2007. Prospects for genomewide selection for quantitative traits in maize. Crop Science. 47:1082-1090.

Bolstein, D., R. L. White, M. Skolnick and R. W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment polymorphsims. American Journal of Human Genetics. 32:314-31.

Burnett, P. A. and C. J. D'Arcy. 1995. Barley yellow dwarf: a brief introduction. *In*: Barley Yellow Dwarf, 40 Years of Progress. 1-5.

Chaffin, A.S., Y.-F. Huang, S. Smith, W.A. Bekele, E. Babiker, F. Nanjappa, B.J. Foresman, S. Blanchard, J. Jay, R. Reid, C.P. Wight, S. Chao, R.E. Oliver, E. Islamovic, F.L. Kolb, C. McCartney, J.W. Mitchell Fetch, A. Beattie, Å. Bjørnstad, A.M. Bonman, T. Langdon, C. Howarth, C. Brouwer, E.N. Jellen, J. Poland, T.-F. Hsieh, R. Brown, E.W. Jackson, J.A. Schlueter and N.A. Tinker. 2015. Submitted to Plant Genome.

Chee, M., R. Yang, E. Hubbell, A. Berno, X. C. Huang, D. Stern, J. Winkler, D. J. Lockhart, M.S. Morris and S.P. Fodor. 1996. Accessing genetic information with high-density DNA arrays. Science. 274:610-4.

Collard, B.C.Y. and D.J. Mackill. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society B. 363:557-572.

Cooper, J. I., and A. T. Jones. 1983. Responses of plants to viruses: proposal for the use of terms. Phytopathology. 73:127-128.

D'Arcy, C. J., and P. A. Burnett, eds. 1995. Barley yellow dwarf: 40 years of progress. St. Paul, Minnesota: The American Phytopathological Society.

D'Arcy, C.J., and M.Mayo. 1997. Proposals for changes in luteovirus taxonomy and nomenclature. Archives of Virology. 142:1 285-1 287.

Endo, R. M. and C. M. Brown. 1964. Barley yellow dwarf virus resistance in oat. Crop Science. 4:279-283.

Foresman, Bradley J. Identification of quantitative trait loci and comparison of selection methods for barley yellow dwarf tolerance in spring oat and winter wheat. M.S. The University of Illinois at Urbana-Champaign, 2014.

Gellner, J. L., J. S. Posch, D. L. Reeves, and R. W. Kieckhefer. 1992. Use of the third leaf to select for barley yellow dwarf resistance in oat. Cereal Research Communications. 20:201-205.

Gnanesh, B.N., J.M. Fetch, J.G. Menzies, A.D. Beattie, P.E. Eckstein and C.A. McCartney. 2013. Chromosome location and allele-specific PCR markers for marker-assisted selection of the oat crown rust resistance gene *Pc91*. Molecular Breeding. 32.3:679-686.

Halbert, S. and D. Voegtlin. 1995. Biology and taxonomy of vector of barley yellow dwarf viruses. *In:* Barley Yellow Dwarf, 40 Years of Progress. 217-258.

Heffner, E.L., M.E. Sorrells and J.-L. Jannink. 2009. Genomic selection for crop improvement. Crop Science. 49:1-12.

Heslot, N., H.-P. Yang, M.E. Sorrells and J.-L. Jannink. 2012. Genomic selection in plant breeding: a comparison of models. Crop Science. 52:146-160.

Hewings, A. D. 1995. Purification and virion characterization of barley yellow dwarf viruses. *In:* Barley Yellow Dwarf, 40 Years of Progress. 165-179.

Hewings, A. D., F. L. Kolb, G. R. Gregerson and E. M. Bauske. 1992. Field research and germplasm evaluation methodology for barley yellow dwarf viruses in cereals. *In*: The Proceedings at the Fourth International Oat Conference. 44-66.

Huang, Y.-F., J.A. Poland, C.P. Wight, E.W. Jackson, N.A. Tinker. 2014. Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. PLoS One. 9(7): e102448.

Jannink, J.-L., M.C.A.M. Bink and R.C. Jansen. 2001. Using complex plant pedigrees to map valuable genes. Trends in Plant Science. 6:337-342.

Jaccoud, D., K. Peng, D. Feinstein, and A. Kilian. 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Research. 29: e25.

Jensen, S. G. and C. J. D'Arcy. 1995. Effects of barley yellow dwarf on host plants. *In:* Barley Yellow Dwarf, 40 Years of Progress. 55-74.

Jin, H., L. L. Domier, F. L. Kolb, and C. M. Brown. 1998. Identification of quantitative loci for tolerance to barley yellow dwarf virus in oat. Journal of Phytopathology. 88.5:410-5.

Kennedy, B.W., M. Quinton and J.A.M. van Arendonk. 1992. Estimation of effects of single genes on quantitative trait. Journal of Animal Science. 70:2000-2012.

Kennedy, J. S. 1951. Benefits to aphids from feeding on galled and virus-infected plants. Nature. 168:890-894.

Kolb, F. L., N. K. Cooper, A. D. Hewings, E. M. Bauske and R.H. Teyker. 1991. Effects of barley yellow dwarf virus on root-growth in spring oat. Plant Disease. 75.2:143-5.

Kosova, K., J. Chrpova, and V. Sip. 2008. Recent advances in breeding of cereals for resistance to barley yellow dwarf virus - a review. Czech Journal of Genetics and Plant Breeding. 44.1:1-10.

Kring, J. B. 1972. Flight behavior of aphids. Annual Review of Entomology. 17:461-492.

Lande, R. and R. Thompson. 1990. Efficiency of marker-assisted selection in improvement of quantitative traits. Genetics. 124.3:743-756.

Lipka, A.E., F. Lu, J.H. Cherney, E.S. Buckler, M.D. Casler, and D.E. Costich. 2014. Accelerating the switchgrass (*Panicum virgatum* L.) breeding cycle using genomic selection approaches. PLoS ONE 9:e112227.

Lorenz, A.J., S. Chao, F.G. Asoro, E.L. Heffner, T. Hayashi, H. Iwata, K.P. Smith, M.E. Sorrells and J.-L. Jannink. 2011. Genomic selection in plant breeding: knowledge and prospects. Advances in Agronomy. 110:77-123.

Mann, B. P., S. D. Wratten, M. Poehling, and C. Borgemeister. 1991. The economics of reduced-rate insecticide applications to control aphids in winter-wheat. Annals of Applied Biology. 119:451-464.

Marshall, H. G., and M. E. Sorrells, eds. 1992. Oat Science and Technology. 1st ed. Madison, Wisconsin, USA: American Society of Agronomy, Inc.; Crop Science Society of America, Inc.

McKenzie, R. I. H., P. A. Burnett, C. C. Gill, A. Comeau and P. D. Brown. 1985. Inheritance of tolerance to barley yellow dwarf virus in oat. Euphytica. 34:681-7.

McKirdy, S. J., R. A. C. Jones, and F. W. Nutter Jr. 2002. Quantification of yield losses caused by barley yellow dwarf virus in wheat and oat. Plant Disease. 86:769-73.

McKirdy, S. J., and R. A. C. Jones. 1996. Use of imidacloprid and newer generation synthetic pyrethroids to control the spread of barley yellow dwarf luteovirus in Cereals. Plant Disease. 80:895-901.

Montilla-Bascón, G., N. Rispail, J. Sánchez-Martin, D. Rubiales, L.A.J. Mur, T. Langdon, C.J. Howarth and E. Prats. 2015. Genome-wide association study for crown rust (*Puccinia coronata* f. sp. *avenae*) and powdery mildew (*Blumeria graminis* f. sp. *avenae*) resistance in oat (*Avena sativa*) collection of commercial varieties and landraces. Frontiers in Plant Science. 6:103.

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

Newell, M.A., D. Cook, N.A. Tinker and J.-L. Jannink. 2011. Population structure and linkage disequilibrium in oat (*Avena sativa* L.): implications for genome-wide association studies. Theoretical and Applied Genetics. 122:623-632.

Newell, M.A., F.G. Asoro, M.P. Scott, P.J. White, W.D. Beavis and J.-L. Jannink. 2012. Genome-wide association study for oat (*Avena sativa* L.) Beta-glucan concentration using germplasm of worldwide origin. Theoretical and Applied Genetics. 125:1687-1696.

Oliver, R. E, D. E. Obert, G. Hu, J. M. Bonman, E. O'Leary-Jepsen, and E. W. Jackson. 2010. Development of oat-based markers from barley and wheat microsatellites. Genome. 53:458-71.

Oliver, R. E., G. R. Lazo, J. D. Lutz, M. J. Rubenfield, N. A. Tinker, J. M. Anderson, N. H. Wisniewski-Morehead, D. Adhikary, E. N. Jellen, P. J. Maughan, G. L. Brown-Guedira, S. Chao, A. D. Beattie, M. L. Carson, H. W. Rines, D. E. Obert, J. M. Bonman, and E. W. Jackson. 2011. Model SNP development for complex genomes based on hexaploid oat using high-throughput 454 sequencing technology. BMC Genomics. 12: 77.

Oswald. J.W. and B.R. Houston. 1951. A new disease of cereals, transmissible by aphids. Plant Disease Report. 43:471-475.

Perry, K. L., F. L. Kolb, B. Sammons, C. Lawson, G. Cisar, and H. Ohm. 2000. Yield effects of barley yellow dwarf virus in soft red winter wheat. Virology. 90:1043-8.

Power, A.G. and S.M. Gray. 1995. Aphid transmission of barley yellow dwarf viruses: interactions between viruses, vectors, and host plants. *In:* Barley Yellow Dwarf, 40 Years of Progress. 259-289.

Qualset, C.O. 1984. Evaluation and breeding methods for barley yellow dwarf resistance. In: Barley Yellow Dwarf: A Proceedings of the Workshop. 72-80.

Ribaut, J.-M. and D. Hoisington. 1998. Marker-assisted selection: new tools and strategies. Trends in Plant Science. 3:236-239.

Rochow, W. F. 1969. Biological properties of four isolates of barley yellow dwarf virus. Phytopathology. 59:1580-1589.

Tinker, N. A., A. Kilian, C. P. Wight, K. Heller-Uszynska, P. Wenzl, H. W. Rines, A. Bjornstad, C. J. Howarth, J. Jannick, J. M. Anderson, B. G. Rossnagel, D. D. Stuthman, M. E. Sorrells, E. W. Jackson, S. Tuvesson, F. L. Kolb, O. Olsson, L. C. Federizzi, M. L. Carson, H. W. Ohm, S. J. Molnar, G. J. Scoles, P. E. Eckstein, J. M. Bonman, A. Ceplitis and T. Langdon. 2009. New DArT markers for oat provide enhanced map coverage and global germplasm characterization. BMC Genomics 10:39.

USDA. 2015. Small Grains 2015 Summary. United States Department of Agriculture National Agriculture Statistics Service.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M, Hornes, A, Frijters, J Pot, J Peleman, and M. Kuiper. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research. 23:4407-14.

Weber, J., and P. May. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. American Journal of Human Genetics. 44:388-96.

Wight, C. P., W. Yan, J. M. Fetch, J. Deyl and N. A. Tinker. 2010. A set of new simple sequence repeat and avenin DNA markers suitable for mapping and fingerprinting studies in oat (*Avena* Spp.). Crop Science. 50:1207-18.

Williams, J., A. R. Kubelik, K. J. Livak, J. A. Rafalski and S. V. Tingey. 1990. DNA Polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research. 18:6531-5.

Zhu, C., M. Gore, E. Buckler, S. Edward and J. Yu. 2008. Status and prospects of association mapping in plants. The Plant Genome. 1:5-20.

Zhu, S., F. L. Kolb, and H. F. Kaeppler. 2003. Molecular mapping of genomic regions underlying barley yellow dwarf tolerance in cultivated oat (*Avena sativa* L.). Theoretical and Applied Genetics. 106:1300-1306.

CHAPTER 2

GENOME-WIDE ASSOCIATION MAPPING OF BARLEY YELLOW DWARF VIRUS TOLERANCE IN SPRING OAT (*AVENA SATIVA* L.)

2.1 Abstract

Barley yellow dwarf (BYD) is the most important viral disease of cereal grain crops worldwide. Barley yellow dwarf viruses (BYDVs) are responsible for BYD and affect many cereals including oat (*Avena sativa* L.). Until recently, the molecular marker technology in oat has not allowed for many marker-trait association studies to determine the genetic mechanisms for tolerance. A genome-wide association study (GWAS) was performed on 428 spring oat lines using a recently developed high-density oat single nucleotide polymorphism (SNP) array as well as a SNP-based consensus map. Marker-trait associations were performed using a Q-K mixed model approach to control for population structure and relatedness. Six significant SNP-trait associations representing two quantitative trait loci (QTL) were found on chromosomes 3C and 18D. This is the first report of BYDV tolerance QTL on chromosomes 3C and 18D. Haplotypes using the two QTL were evaluated and distinct classes for tolerance were identified based on the number of favorable alleles. A large number of lines carrying both favorable alleles were observed in the panel.

2.2 Introduction

Oat (*Avena sativa* L.) is an important cereal crop grown worldwide with nutritional benefits for both livestock and humans (Andon and Anderson, 2008). Barley yellow dwarf (BYD) is one of the most destructive viral diseases of small grains. The disease was first described in barley by Oswald and Houston in 1951 and affects all major cereal crops (rice, maize, wheat, oat and rye) as well as other grass species. The disease is caused by a group of phloem limited luteoviruses known as barley yellow dwarf viruses (BYDVs) that are obligately transmitted via aphid vectors (Hewings, 1995). Symptoms on oat depend on cultivar and environment but normally include leaf discoloration, stunted growth and blasting of florets. Economic losses due to BYDVs in oat range from 13-25 kg/ha for each 1% increase in incidence (McKirdy et al., 2002). Methods of control for BYDVs include insecticides to control aphid populations; however, insecticides may only be feasible in highly intensive agricultural systems. The most effective way to control BYD is by planting tolerant cultivars (Burnett et al., 1995).

Two types of resistance to BYDV have been distinguished: virus resistance and field resistance. Field resistance is usually referred to as tolerance (Kosova et al., 2008). Virus resistance refers to low virus titer in infected plants whereas field resistance (tolerance) refers to the reduction of symptoms of infection independent of the virus titer. In this paper, resistance will be defined as reduced viral replication in infected plants (Cooper and Jones, 1983). Tolerance will therefore be defined as the development of mild or negligible symptoms in infected plants. It can also be stated as the ability of plants to yield under BYDV infection.

Historically, molecular marker technology in hexaploid oat has lagged behind that of maize, soybeans, and other diploid crops. Therefore, phenotyping for BYDV has been the only

reliable method for screening breeding material for tolerance. Marker assisted breeding (MAB) would allow for more efficient selection of tolerant lines by aiding in the introgression of multiple genes controlling the trait. Until recently, the only available marker platforms in oat were restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), and random amplified polymorphic DNA (RAPD) markers (Tanksley et al., 1989; Vos et al.. 1995; Williams et al., 1990). As molecular technology has improved, newer and less expensive options have become available to oat breeders. Advancements in abundance and coverage with simple sequence repeats (SSR) markers have increased in oat, and the development of Diversity Arrays Technology (DArT) markers, has provided another option for a high-throughput assay (Becher, 2007; Li et al., 2000; Pal et al., 2002; Oliver et al., 2010; Tinker et al., 2009). Genotyping-by-sequencing also has been applied in oat (Huang et al., 2014). Recently, single nucleotide polymorphism (SNP) technology has been developed for oat from expressed sequence tag (EST) information for SNP genotyping and a high-throughput 6K oat SNP array has been developed (Oliver et al., 2011; Tinker et al., 2014).

Breeding for host-plant tolerance to BYD is the most effective method used to combat the destructive nature of the disease. It is believed that two to four genes are responsible for BYDV tolerance in oat (McKenzie et al. 1985). Four QTL (BYDq1, BYDq2, BYDq3 and BYDq4) for BYD tolerance were identified and were found on linkage groups OM1, 5, 7 and 24 (Zhu et al., 2003). Several other chromosomal regions with BYDV tolerance QTL have been identified in other studies. Barbosa-Neto et al. in 2003 found 21 chromosomal regions distributed over 16 linkage groups that were associated with tolerance to BYD in oat. Using the new 6K high density oat SNP array, several QTL were identified in two bi-parental populations. Two large effect QTL were identified on chromosome 3C and 19A (Foresman, 2014).

Genome-wide association studies (GWAS) are powerful in determining the genetic structure of complex traits in crops, and a few studies have been used to determine the relationship between markers and BYDV tolerance in oat. GWAS detects marker-trait associations by exploiting linkage disequilibrium between a marker allele and the causative QTL allele. However, population structure within the panel of lines and genetic relationships within the population can lead to false positive associations and therefore need to be taken into consideration (Yu, et al., 2006; Stich et al., 2008). GWAS with population structure control has been successfully used to detect marker-trait associations for beta-glucan concentration in oat (Asoro et al., 2013; Newell et al., 2012). Marker-trait associations that are identified will enable breeders to use marker assisted selection (MAS) as well as genomic selection models. Using a set of oat cultivars and breeding lines as a GWAS panel, our objectives were to (1) assess population structure in the GWAS panel, (2) identify markers associated with BYDV tolerance and (3) to examine haplotypes within the population using the identified QTL.

2.3 Materials and Methods

Plant material and disease assessment

The association panel was developed by the Collaborative Oat Research Enterprise (CORE), group of research scientists from North America and several other countries worldwide. Oat breeders submitted lines to the panel which includes a high level of diversity. A total of 428 spring oat lines were included in the final panel and phenotyped for BYDV tolerance.

Tolerance for BYDV was evaluated at the University of Illinois Crop Sciences Research and Education Center, Urbana, IL in 2010 and 2011. Two replications of hills were planted in a randomized complete block design with 15 seeds per hill in the BYDV nursery. When the seedlings were in the three leaf stage (approximately 20 days after planting), the hills were inoculated with viruliferous *Rhopalosiphum padi* (L.) carrying the Illinois isolate BYDV-PAV. The plants were sprayed with insecticide (Cygon 2E™) after one week to kill the aphids. BYDV tolerance was evaluated after stem elongation was completed, using a scale from 0 to 9 with 0 being assigned to the most tolerant plants and 9 indicating the most sensitive (Qualset, 1983; Hewings et al., 1992).

Phenotypic data analysis

Best linear unbiased predictors (BLUPs) for BYDV tolerance were calculated using a mixed model:

$$Y_{ijk} = \mu + year_i + block(year)_{ij} + line_k + (year\ x\ line)_{ik} + \varepsilon_{ijk} \qquad \text{(i)}$$

Where Y_{ijk} is the observed BYDV phenotype, μ is the overall mean, $year_i$ is the random effect of the *i*th year, $block_j$ is the random effect of the *j*th block within the *i*th year, $line_k$is the random effect of the *k*th line, $year\ x\ line_{ik}$ is the random effect of the interaction between the *i*th year and the *k*th line, and ε_{ijk} is the random error term. Broad-sense heritabilities on an entry-mean basis (H^2) were calculated for BYDV tolerance using the variance components from the mixed model above (i) and equation (ii):

$$H^2 = \frac{\sigma_g^2}{\sigma_e^2/rt + \sigma_{ge}^2/t + \sigma_g^2} \qquad \text{(ii)}$$

where H^2 is the entry-mean heritability, σ_g^2 is the genetic variance, σ_e^2 is experimental error, σ_{ge}^2 is the variance due to the interaction of genotype and environment, r is the number of replications, and t is the number of environments.

Genotypic data

A total of 428 spring oat lines were included in the final panel and were genotyped using a high density oat SNP array containing 6,000 SNPs. Genotyping was performed at the USDA-ARS Small Grains Genotyping Lab in Fargo, ND using the Infinium assay developed by Illumina (Oliver et al., 2011). A total of 2,305 SNPs were identified to be polymorphic in the population. The marker number was further reduced by filtering for minor allele frequencies below 0.05 and markers with the proportion of missing genotypes greater than 0.10. Furthermore, by using the LDTagSNP Selection function in JMP Genomics 7 (SAS, Cary, NC) markers showing linkage disequilibrium (r^2) higher than 0.8 were binned, and a representative SNP from each bin was used (Carlson et al., 2004). This process was to help reduce the redundancy of markers explaining the same information and led to a final number of 1,402 SNPs used in the association analysis.

A previously developed SNP-based physically anchored consensus map was used for marker locations (Oliver et al., 2013). The map was developed from 390 recombinant inbred lines from six bi-parental populations, which included 985 SNPs and 68 previously-published markers. The final map consisted of 21 linkage groups having a total map distance of 1,838.8 cM.

Genome-wide association analysis

Principal component analysis (PCA) was performed in order to examine the level of genetic structure in the panel (Q-matrix) via the PCA for Population Stratification function. The amount of relatedness (marker based kinship matrix, K-Matrix) was performed via the Relationship Matrix function. The above functions, as well as the marker-trait associations were performed in JMP Genomics 7 (SAS, Cary, NC). A Q-K mixed model was used for the marker-trait associations using the PCA values to form the Q-matrix, treated as fixed effect and the identity-by-descent (IBD) values for the K-matrix, treated as random effect. IBD values were calculated using equation (iii);

$$IBD_{i,j} = \left(X_{i,l} - 2p\right) * (X_{j,l} - 2p)/2pq \qquad \text{(iii)}$$

Where $X = 0, 1, 2$ correspond to genotype BB, AB, AA at marker l, and p and q are the allele frequencies for allele A and B. The measure is averaged over all loci. The mixed model procedure was performed using the Q-K mixed model function with a false discovery rate α=0.05 for multiple testing correction.

2.4 Results

Phenotypic data

A wide range of phenotypic variation was observed for BYDV tolerance (Table 1). This was expected due to the quantitative nature of the trait. The BYDV tolerance had a mean rating of 4.54 with maximum of 8.58 and a minimum of 1.33. High broad-sense heritability was

observed for BYDV tolerance with a value of 0.91, which is consistent with recently observed heritability but is higher than historical observations (Barbosa-Neto et al., 2000; McKenzie et al., 1985).

Genotypic data and population structure

The level of population structure was examined to gain an understanding of the possible effect on the association analysis. Principal component analysis (PCA) using eigenvalues on marker data showed that PCA 1 accounted for 8.7% of the variation in the data. The first five PCAs accounted for a total of 23.3% variation in the data, showing there may be some slight population structure in the panel of lines. Five principal components were determined to be sufficient and used in the final model based on the inflection point of the scree plot.

Marker-trait associations for BYDV Tolerance

A total of six SNPs were significantly associated with BYDV tolerance on chromosomes 3C and 18D (Figure 1). SNP GMI_ES22_c20081_313 on chromosome 3C accounted for 17% of the variance with an effect of 0.82 per favorable allele (Table 2). There were three other markers that were significant on chromosome 3C that explained between 6% and 7% of the variation and had effects between -0.49 and -0.31. On chromosome 18D, SNP GMI_ES05_c3073_282 had the highest significance and explained 6% of the variation with an effect of 0.50. One other marker, SNP EMI_ES17_c6498_89, was also significant and explained 3% of the variation with an effect of -0.31.

Using the most significant marker on chromosome 3C (GMI_ES22_c20081_313) and 18D (GMI_ES05_c3073_282), haplotypes were compared using contrasts (Table 3). For ease of use, the two SNP loci will be referred to SNP 3C and SNP 18D moving forward. Two entries were not used in the haplotype analysis due to missing data at one of the two markers. All other entries fell into the four possible haplotypes (haplotypes 1 through 4) (Figure 2). Haplotype 1 contained the favorable allele ("+ +") at both loci and exhibited the lowest mean for BYDV tolerance (3.41). Haplotype 4 was constituted by having both unfavorable alleles ("- -") at SNP 3C and SNP 18D. This haplotype displayed the highest BYDV sensitivity at 7.46. Haplotypes 2 (favorable allele at SNP 3C and unfavorable allele at SNP 18D, "+ -") and haplotype 3 (unfavorable allele at SNP 3C and the favorable allele at SNP 18D, "- +") both had means that fell in between haplotype 1 and 4 (5.23 and 5.96). Contrasts between the four haplotypes were all significant at $\alpha = 0.05$.

Two different sets of near-isogenic lines (NILs) were included in the panel from the University of Illinois at Urbana-Champaign (Kolb et al., 2006). Each set contains four lines that were developed using phenotypic selection. NIL family IL2250 contained four lines that ranged in BYDV rating from 3.14 to 6.98 (Table 4). NIL family IL2294 exhibited a similar range of BYDV ratings between 3.37 and 6.98.

The haplotypes using SNP 3C and SNP 18D exhibited similar patterns in both families. In family IL2250, the two tolerant lines (IL2250-18 and IL2250-14) both had BYDV BLUPs of 3.14 and 3.37, respectively, and both lines had a "+ +" haplotype. The moderately tolerant line IL2250-3 had BLUP of 4.95 with a haplotype of "- +" and the sensitive line IL2250-15 with a BLUP of 6.98 had the unfavorable allele at both SNPs ("- -").

In the second family, the two tolerant lines IL2294-3 and IL2294-8 had BLUPs of 3.37 and 4.73. Both lines had the favorable allele at both SNPs ("+ +"). IL2294-1 had a BLUP of 6.31 and had the favorable allele at SNP 3C and the unfavorable allele at SNP 18D ("+ -"). IL2294-2 was the most sensitive line in the family and had a BLUP of 6.98. As expected this line had the unfavorable allele at both SNP markers ("- -").

2.5 Discussion

Barley yellow dwarf is one of the most destructive diseases of cereal crops worldwide. New breeding strategies such as marker-assisted selection and genomic selection along with high-throughput genotyping platforms can help to provide breeders with the tools necessary to introgress tolerance into elite cultivars. The identification of QTL associated with BYDV tolerance is an important first step to understanding the genetics of the tolerance mechanisms. In this study, genome-wide association mapping was performed on a panel of 428 spring oat cultivars using a 6K oat SNP array. This panel of lines includes diverse oat germplasm from several countries and breeding programs. A total of 1402 SNPs were used for genome-wide association for BYDV tolerance. A broad and continuous distribution was observed for BYDV tolerance across the panel and is in agreement with previous work that shows host plant tolerance for BYDV is multigenic (McKenzie et al., 1985; Jin et al., 1998; Zhu et al. 2003;).

Population structure can result in false associations between markers and traits and therefore should be evaluated for proper analysis (Matthies et al., 2012). Principal component analysis with the SNP markers was used to determine the level of population structure in the panel. A moderate level of structure was observed via PCA analysis and is likely due to multiple

oat breeding programs in the CORE submitting lines. Relatedness (identity-by-descent) of lines was also evaluated for K-matrix calculation. To control false positive associations due to a high number of lines submitted from each program, the marker-trait association analysis was performed with a Q-K model containing both the matrices.

Previous studies have identified QTL for BYDV tolerance; however, many were performed with a very low number of markers, and the new SNP-based consensus map did not use any of the markers reported before, making it difficult to compare the linkage maps and the consensus map. Two bi-parental recombinant inbred lines populations using a previous version of the consensus map and the same 6K high density oat SNP array identified two large effect QTL that mapped to chromosomes 3C and 19A (Barbosa-Neto et al., 2000). In this study, two QTL were identified on chromosome 3C and 18D. These two QTL were located at 114.5 cM and 147.7 cM on their respective chromosomes.

The most recent update to the consensus map included some changes to some marker locations; however, markers in each of the QTL areas grouped together in a manner similar to previous versions. SNP GMI_ES05_c3073_282, the most significant marker in the previous bi-parental study, was formerly mapped to chromosome 19A. In this study, due to the updated consensus map, the marker was located on 18D but remained the most significant marker at the QTL. On chromosome 3C, the two studies identified a different significant marker, but this is likely due to the LDTagSNP selection function that bins markers and tags a representative SNP for each bin. It is also important to point out that the two tolerant parents (IL 86-1156 and IL 86-6404) and the susceptible parent (Clintland 64) from the bi-parental studies were included in this association panel. It appears that the two QTL identified in this GWAS are in agreement with the results from the bi-parental study.

Examining the potential haplotypes for the identified QTL showed that having the favorable alleles at both loci ("++") leads to improved levels of tolerance (lower BYDV scores). Although, all the contrasts comparing the four haplotypes were significant, classes of tolerance are visible based on the number of beneficial alleles. From a plant breeding perspective, this is important because there are different implications when making selections. Identifying the most tolerant lines is important because this group can be selected to move forward in a breeding program or could be used as parents to improve other lines without tolerance. Haplotype 1 ("++") is representative of this "high tolerance" class. The second group has medium levels of tolerance for BYDV. This class contains both haplotype 2 ("+-") and haplotype 3 ("-+"). These classes contain cultivars with one positive allele. Even though there is a significant difference between haplotype 2 and 3, from a breeding perspective both haplotypes have medium levels of tolerance and therefore can be grouped together in the "moderate tolerance" class. The "sensitive" class is also important to be identified because these lines do not carry any BYDV tolerance and could be discarded from a breeding program or if they contain other beneficial traits could be crossed with more tolerant lines to improve them.

From a breeding perspective, it is important to identify the tolerant lines that were submitted as well as the programs they came from. The complete list of lines is included in the supplementary material as well as a breakdown of the haplotypes by programs (Supplementary Tables 1 and 2). Overall, 259 out of 428 lines included in the panel have the haplotype "+ +", 114 lines contained one favorable allele and one unfavorable allele ("+ -" or "- +") and 55 lines had two unfavorable alleles. The results appear to show that the oat community has effectively selected for BYDV tolerance or these alleles were at a higher frequency in the founder lines used

thought-out breeding programs. Using BYDV tolerant lines as breeding parents can help to combine BYDV tolerance with other beneficial traits.

The two families of NIL lines that were included in this study were both developed using phenotypic selection during the 1990s. The marker haplotype data agree with the phenotypically selected lines and a similar breakdown into distinct classes could be seen. Several other families of NILs were not included in the panel and therefore should also be genotyped in the future to further examine the breakdown of the high tolerance, moderately tolerant and sensitive groups.

2.6 Conclusions

Host plant tolerance for BYD is the most effective mechanism for reducing losses to the disease. In this study, two QTL were identified on chromosomes 3C and 18D using a 6K high density oat SNP array and recently published consensus map. These results are consistent with a previous bi-parental study using the same SNP array. This study also identified three main levels of BYDV tolerance based on the number of favorable alleles at the two loci. The largest group with over half the lines in the panel formed the "tolerant" class and had favorable alleles at both 3C and 18D. The SNPs identified in this study can be used in marker-assisted selection or genomic selection programs to better improve host plant tolerance for BYD.

2.7 Tables and Figures

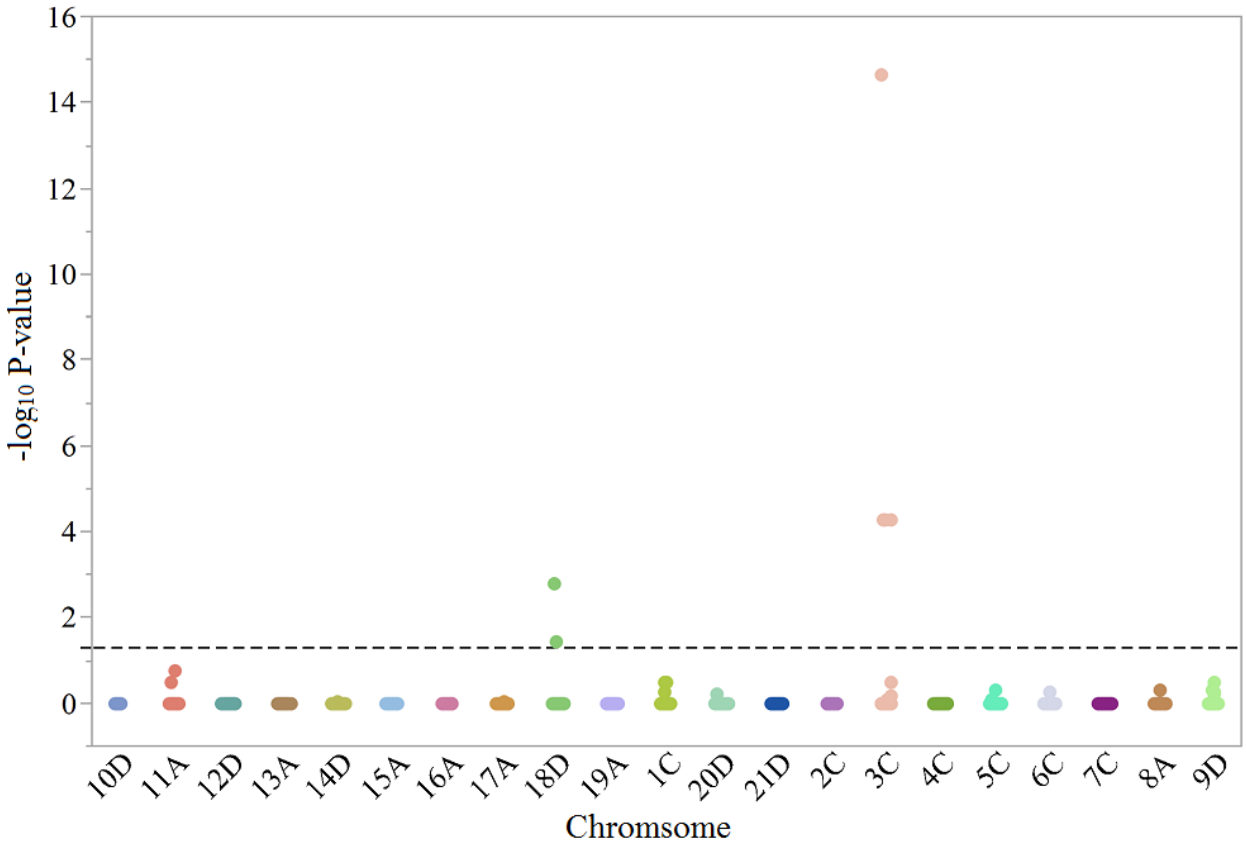

Figure 2.1 Genome-wide association for oat BYDV tolerance. The dotted line represents the significance threshold at $\alpha = 0.05$ ($-\log_{10}$(p-value) ≥1.30), for false discovery rate adjusted P-values.

Table 2.1 Descriptive statistics and broad sense entry-mean heritability for BLUPs for BYDV tolerance in 428 spring oat lines.

	BLUPs					
	Mean	Min	Max	Range	SD	$H^{2\wedge}$
BYDV	4.54	1.33	8.58	7.25	1.85	0.91

$^{\wedge}H^2$ = broad-sense heritability

Table 2.2 SNPs associated with BYDV tolerance in a panel of 428 spring oat lines, chromosomal position, *p*-values and marker effects.

SNP	C	cM	p^a	maf[b]	r^2	Adj. p^c	Effect
GMI_ES22_c20350_257	3C	113.2	6.84	0.40	0.07	4.29	-0.49
GMI_ES22_c20081_313	3C	114.5	17.8	0.32	0.17	14.6	0.82
GMI_DS_LB_10400	3C	115.1	6.91	0.46	0.06	4.29	-0.49
GMI_DS_CC1800_254	3C	115.1	6.93	0.41	0.07	4.29	-0.31
GMI_ES17_c6498_89	18D	132.7	3.82	0.23	0.03	1.45	-0.31
GMI_ES05_c3073_282	18D	147.7	5.25	0.19	0.06	2.8	0.50

[a] *p*-value reported on a $-\log_{10}$ scale; [b]maf = minor allele frequency; [c] False Discovery rate adjusted *p*-value

Table 2.3 Contrasts between haplotypes for BYDV tolerance in oat.

Haplotype	Estimate	Std. Error	F Ratio	prob > F
1 vs 2	-1.34	0.10	228.9	<.0001*
1 vs 3	-2.31	0.08	1183.7	<.0001*
1 vs 4	-3.69	0.08	1986.4	<.0001*
2 vs 3	-0.62	0.12	27.5	<.0001*
2 vs 4	-2.04	0.13	254.7	<.0001*
3 vs 4	-1.40	0.10	213.5	<.0001*

*Significance at $\alpha = 0.05$

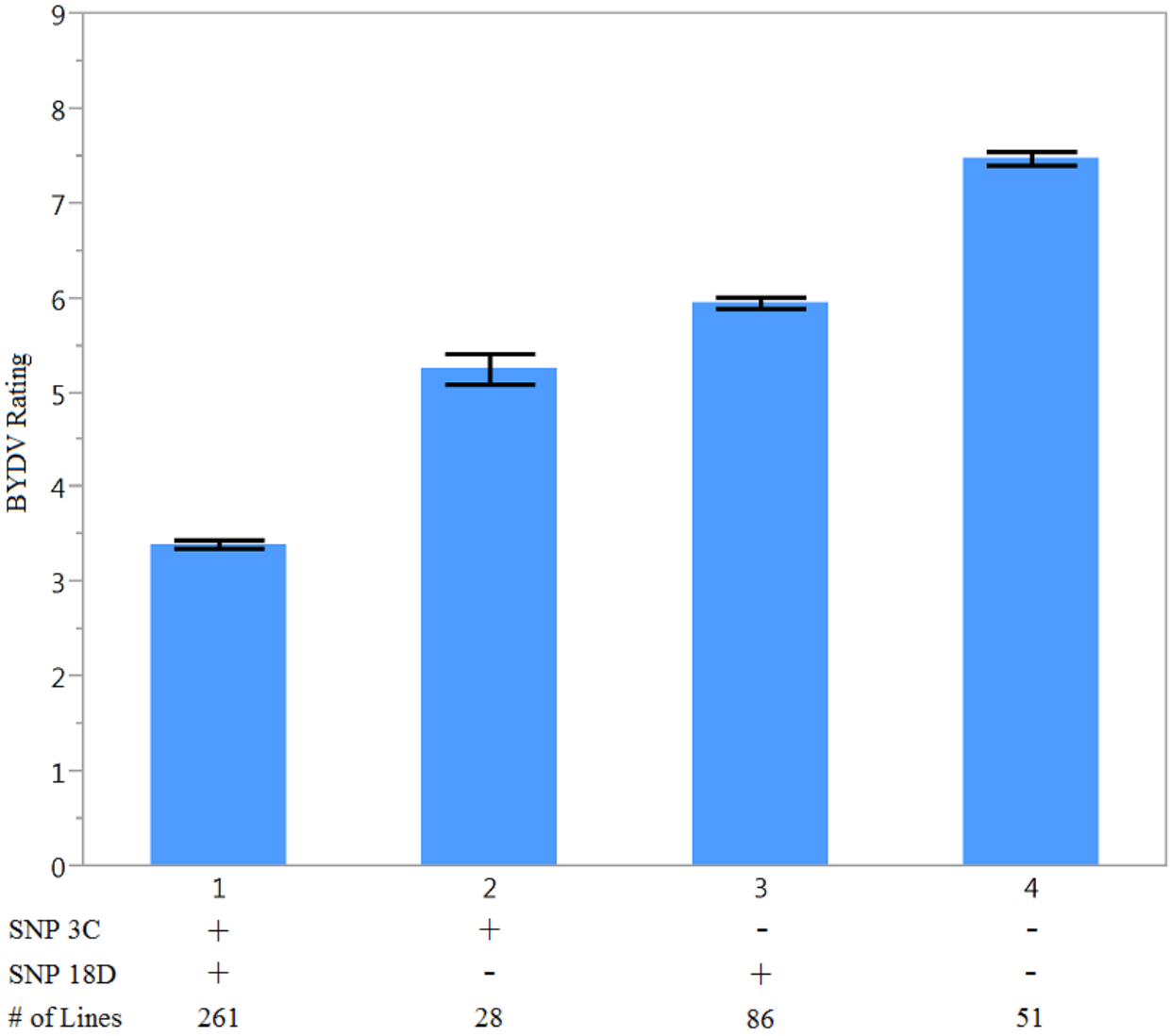

Figure 2.2 Phenotypic mean values for BYDV tolerance for oat haplotypes containing different SNP allele combinations with error bars representing one standard error. " + " signifies a favorable allele. BYDV rating scale is from 0 to 9 where 0 = most tolerant and 9 = most sensitive.

Table 2.4 Barley yellow dwarf virus haplotypes for SNP3C and SNP18D in two sets of near-isogenic lines (Kolb et al. 2006)

NIL	Pedigree	GWAS BLUP	2 Year Mean^	SNP3C	SNP18D
IL2250-18	Clintland 64*5 /IL86-5698	3.14	3.3	+	+
IL2250-14	Clintland 64*5 /IL86-5698	3.37	3.6	+	+
IL2250-3	Clintland 64*5 /IL86-5698	4.95	6.3	-	+
IL2250-15	Clintland 64*5 /IL86-5698	6.98	8.6	-	-
IL2294-3	Clintland 64*5 /IL86-6404	3.37	4.5	+	+
IL2294-8	Clintland 64*5 /IL86-6404	4.73	3.6	+	+
IL2294-1	Clintland 64*5 /IL86-6404	6.31	8.9	+	-
IL2294-2	Clintland 64*5 /IL86-6404	6.98	8.3	-	-

^2 year mean from Kolb et al., 2006

2.8 References

Andon, M.B, and J.W Anderson. 2008. The oatmeal-cholesterol connection: 10 years later. American Journal of Lifestyle Medicine 2: 51-57.

Asoro, F.G., Newell, M.A., Scott, M.P., Beavis W.D., and J.-L. Jannink. 2013. Genome-wide association study for β-glucan concentration in elite North American oat. Crop Science 53: 542.

Barbosa-Neto, J. F., W. Siripoonwiwat, L. S. O'Donoughue, S. M. Gray, D. M. Smith, F. L. Kolb, C. Gourmet, C. M. Brown and M. E. Sorrells. 2000. Chromosomal regions associated with barley yellow dwarf virus resistance in oat. Euphytica. 114:67-76.

Becher, R. EST-derived microsatellites as a rich source of molecular markers for oat. 2007. Plant Breeding 126: 274-278.

Burnett, P. A.,Comeau, A. and C.O. Qualset. 1995. Host plant tolerance or resistance for control of barley yellow dwarf. Barley Yellow Dwarf, 40 Years of Progress. Cleora J. D'Arcy and Peter A. Burnett. 1st ed. St. Paul, Minnesota: APS Press, 321-344. Print.

Carlson, C.S., Eberle, M.A., Rieder, M.J., Yi, Q., Kruglyak, L., and D.A. Nickerson. 2004. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. The American Journal of Human Genetics 74.1: 106-120.

Cooper, J. I., and A.T. Jones. 1983. Responses of plants to viruses: proposals for the use of terms. Phytopathology 73: 127.

D'Arcy, Cleora J, and P. A Burnett. 1995. Barley Yellow Dwarf 40 years of Progress. St. Paul, Minn.: APS Press.

Foresman, Bradley J. 2014. Identification of quantitative trait loci and comparison of selection methods for barley yellow dwarf tolerance in spring oat and winter wheat. M.S. Thesis. The University of Illinois at Urbana-Champaign.

Hewings, A. D., Kolb, F.L., Gregerson, G.R. and E. M. Bauske. 1992. Field research and germplasm evaluation methodology for barley yellow dwarf viruses in cereals. *In*: The Proceedings at the Fourth International Oat Conference. 44-66.

Hewings, Adrianna D. 1995. Purification and virion characterization of barley yellow dwarf viruses. Barley Yellow Dwarf, 40 Years of Progress. Cleora J. D'Arcy and Peter A. Burnett. 1st ed. St. Paul, Minnesota: APS Press. 165-179.

Huang, Y.-Fen., Poland, J.A., Wight, C.P., Jackson, E.W., and T.A. Tinker. 2014. Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. PLoS ONE 9.7: e102448.

Jin, H., L. L. Domier, F. L. Kolb, and C. M. Brown. 1998. Identification of quantitative loci for tolerance to barley yellow dwarf virus in oat. Journal of Phytopathology. 88.5:410-5.

Kolb, F.L. Brown, C.M., Smith, N.J. and L.L. Domier. 2006. Registration of six sets of near-isogenic spring oat germplasm lines differing in tolerance to barley yellow dwarf virus. Crop Science 46.4: 1831.

Kosova, K., Chrpova, J., and V. Sip. 2008. Recent advances in breeding of cereals for resistance to barley yellow dwarf virus - a review. Czech Journal of Genetics and Plant Breeding 44: 1-10.

Li, C. D., B. G. Rossnagel, and G. J. Scoles. 2000. The development of oat microsatellite markers and their use in identifying relationships among *Avena* species and oat cultivars. Theoretical and Applied Genetics 101.8: 1259-1268.

Matthies, I.E., Hintum, T. V., Weise, S., and M.S. Roder. 2012. Population structure revealed by different marker types (SSR or Dart) has an impact on the results of genome-wide association mapping in European barley cultivars. Molecular Breeding 30: 951-966.

McKenzie, R. I. H., P. A. Burnett, C. C. Gill, A. Comeau and P. D. Brown. 1985. Inheritance of tolerance to barley yellow dwarf virus in oat. Euphytica. 34:681-687.

McKirdy, S. J., R. A. C. Jones, and F. W. Nutter. 2002. Quantification of yield losses caused by barley yellow dwarf virus in wheat and oat. Plant Disease 86: 769-773.

Newell, M.A., Asoro, F.G., Scott, M.P., White, P.J., Beavis W.D., and J.-L. Jannink. 2012. Genome-wide association study for oat (*Avena Sativa* L.) β-glucan concentration using germplasm of worldwide origin. Theoretical Applied Genetics 125.8: 1687-1696.

Oliver, R. E. Obert, D.E., Hu, G. Bonman, J.M., O'Leary-Jepsen, E., and E.W. Jackson. 2010. Development of oat-based markers from barley and wheat microsatellites. Genome 53: 458-471.

Oliver, R. E., G. R. Lazo, J. D. Lutz, M. J. Rubenfield, N. A. Tinker, J. M. Anderson, N. H. Wisniewski-Morehead, D. Adhikary, E. N. Jellen, P. J. Maughan, G. L. Brown-Guedira, S. Chao, A. D. Beattie, M. L. Carson, H. W. Rines, D. E. Obert, J. M. Bonman, and E. W. Jackson. 2011. Model SNP development for complex genomes based on hexaploid oat using high-throughput 454 sequencing technology. BMC Genomics. 12: 77.

Oliver, R.E., Tinker, N.A., Lazo, G.R., Chao, S., Jellen, E.N., Carson, M.L., Rines, H.W., Obert, D.E., Lutz, J.D., Shackelford, I., Korol, A.B., Wight, C.P., Gardner, K.M., Hattori, J., Beattie, A.D., Bjornstad, A., Bonman, J.M., Jannink, J.-L., Sorrells, M.E., Brown-Guedira, G.L., Mitchell Fetch, J.W., Harrison, S.A., Howarth, C.J., Ibrahim, A., Kolb, F.L., McMullen, M.S., Murphy, J.P., Ohm, H.W., Rossnagel, B.G., Yan, W. Miclaus, K.J., Hiller, J., Maughan, P.J., Redman Hulse, R.R., Anderson, J.M., Islamovic, E., and E.W. Jackson. 2013. SNP discovery and chromosome anchoring provide the first physically-anchored hexaploid oat map and reveal synteny with model species. PLoS ONE 22;8(3):e58068.

Oswald, J.W, and B.R. Houston. 1951. A New disease of cereals transmissible by aphids. Plant Disease Reporter 35: 471-475.

Pal, N, Sandhu, J.S., Domier, L.L. and F.L. Kolb. 2002. Development and characterization of microsatellite and RFLP-derived PCR markers in oat. Crop Science 42: 912.

Qualset, C.O. 1983. Evaluation and breeding methods for barley yellow dwarf resistance. Barley Yellow Dwarf Workshop. Mexico: CIMMYT, 1983. 72-80.

Stich, B., Mohring, J., Piepho, H.-P., Heckenberger, M., Buckler, E.S., and Melchinger, A.E. 2008. Comparison of mixed-model approaches for association mapping. Genetics 178: 1745-1754.

Tanksley, S. D., Young, N.D., Paterson, A.H., and M.W. Bonierbale. 1989. RFLP mapping in plant breeding: new tools for an old science. Bio/Technology 7.3: 257-264.

Tinker, N. A., A. Kilian, C. P. Wight, K. Heller-Uszynska, P. Wenzl, H. W. Rines, A. Bjornstad, C. J. Howarth, J. Jannick, J. M. Anderson, B. G. Rossnagel, D. D. Stuthman, M. E. Sorrells, E. W. Jackson, S. Tuvesson, F. L. Kolb, O. Olsson, L. C. Federizzi, M. L. Carson, H. W. Ohm, S. J. Molnar, G. J. Scoles, P. E. Eckstein, J. M. Bonman, A. Ceplitis and T. Langdon. 2009. New DArT markers for oat provide enhanced map coverage and global germplasm characterization. BMC Genomics 10:39.

Tinker, N.A., Chao, S., Lazo, G.R., Oliver, R.E., Huangm Y.-F., Poland, J.A., Jellen, E.N., Maughan, P.J., Kilian, A., and E.W. Jackson. 2014. A SNP genotyping array for hexaploid oat. The Plant Genome 7.3.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M, Hornes, A, Frijters, J Pot, J Peleman, and M. Kuiper. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-14.

Williams, J., A. R. Kubelik, K. J. Livak, J. A. Rafalski and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research. 18:6531-5.

Yu, J., Pressoir, G., Briggs, W.H., Bi, I.V., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., Kresovich, S., and E.S. Buckler. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics 38: 203-208.

Zhu, S., F.L. Kolb, and H.F. Kaeppler. 2003. Molecular mapping of genomic regions underlying barley yellow dwarf tolerance in cultivated oat. Theoretical and Applied Genetics 106: 1300-1306.

CHAPTER 3

COMPARISONS OF GENOMIC SELECTION AND MARKER-ASSISTED SELECTION MODELS FOR BARLEY YELLOW DWARF VIRUS TOLERANCE IN SPRING OAT (*AVENA SATIVA* L.)

3.1 Abstract

Barley yellow dwarf (BYD) is an important viral disease of cereal grains including spring oat (*Avena sativa* L.) and can have significant impact on yield. Phenotyping for sensitivity to the disease is difficult and therefore the use of molecular markers for screening sensitive lines is extremely desirable for plant breeders. Marker-assisted selection (MAS) and genomic selection (GS) provide alternative breeding strategies that have the potential to improve the breeding progress for BYD. Although similar because they require marker-trait associations, GS and MAS differ in how they estimate breeding values. In this study, GS and MAS models were compared in their accuracy to predict barley yellow dwarf virus tolerance in 428 spring oat lines from North America and Europe. Several GS models were evaluated using 2305 single nucleotide polymorphisms (SNPs) and included models with previously identified or randomly selected markers as fixed effects. Model accuracies were evaluated using five-fold cross evaluation. GS models used ridge regression-best linear unbiased predictor (RR-BLUP) for marker effect estimation while MAS models used ordinary least square (OLS). Moderate to high prediction accuracies (0.5-0.9) were observed across the models. GS models containing fixed effects (GS-GWAS, GS-3C18D) from previously identified QTL performed better than the GS

model with all markers as random effects or the MAS models. Two MAS models (MAS-GWAS and MAS-3C18D) had prediction accuracies higher than the GS model with all markers as random effects. Because of this, GS and MAS both have potential as useful breeding strategies for BYDV tolerance in spring oat.

3.2 Introduction

Marker-assisted selection (MAS) and genomic selection (GS) have become common plant breeding strategies for the selection of superior individuals. Both strategies involve using molecular markers to identify lines with advantageous genotypes; however, this is done in slightly different fashions. In MAS, quantitative trait loci (QTL) are used to select individual lines. These QTL are identified via family mapping or population mapping studies with the goal of using recombination to break up the genome into small fragments that can be correlated with phenotypic variation. Family mapping relies on the development of specifically structured populations, usually from bi-parental crosses (i.e. recombinant inbred lines (RILs)), while population mapping is performed using a collection of lines. Because of this, family mapping is restricted to the recombination events that occur while developing the population, which may represent only a small fraction of the total diversity for the trait. QTL identified from family mapping studies tend to be population specific and not applicable to other populations. Using a collection of lines, population mapping takes advantage of all the recombination events throughout evolutionary history, which results in higher mapping resolution compared to family mapping (Myles et al., 2009). Identification of QTL using family or population mapping relies on a significance level, and only markers that reach this level are used in the analysis. This

usually results in only a small number of markers being used in MAS compared to GS. Polygenic traits or complex traits, which are controlled by several small effect genes, may not be suitable for use in MAS (Bernardo et al., 2008). GS contrasts from MAS in the fact that all markers are used for selection of individuals with a high level of expression of the trait of interest (Meuwissen et al., 2001; Jannink et al., 2010). The idea is to include all of the markers, with the assumptions that the markers are evenly spread across the genome, and all QTL for a trait should be accounted for by at least one marker (Goddard and Hayes, 2007). Theoretically, this should result in GS outperforming MAS for polygenic traits.

Genomic selection models use the markers as the predictor variables (p) and the phenotypes as the response variable (n). With the increasing number of available markers, the numbers of markers often exceed the number of phenotypes (large p small n problem). Because of this, models that depend on shrinkage factors are needed to estimate the marker effects. There are several genomic selection models available including ridge regression best unbiased linear predictor (RR-BLUP), genomic best unbiased linear predictor (G-BLUP), least absolute shrinkage and selection operator (LASSO), elastic-net and several Bayesian models (BayesA, BayesB, BayesC) (Lorenz et al., 2011; Heslot et al, 2012).

The most popular model used is RR-BLUP and was first described by Whittaker et al. (2000). RR-BLUP assumes that the markers are from the same distribution and therefore have a common variance. Because of this, the markers are equally shrunken towards zero (Bernardo and Yu, 2007; Heffner et al., 2009). Although, this is not truly realistic in plants, it has been a successful model in plants and in simulation (Arruda et al., 2015; Lipka et al., 2014; Heffner et al., 2009). RR-BLUP uses the model in equation (i):

$$g(\boldsymbol{x}_i) = \sum_{k=1}^{p} x_{ik}\beta_k \quad \text{(i)}$$

Where x_{ik} signifies the score for SNP k in the individual i, β_k is the effect linked with marker k, and the genetic value of the sum of the p marker effects (Lorenz et al., 2011). β is estimated using a modification to the normal least squares estimators in equation (ii):

$$\hat{\beta} = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{\lambda I})^{-1}\boldsymbol{X}'\boldsymbol{y} \quad \text{(ii)}$$

Where $\boldsymbol{X}$ is an incidence matrix associated with individuals, $\boldsymbol{I}$ is an identity matrix, and $\boldsymbol{y}$ is a vector of estimated breeding values that usually are phenotypes. The $\lambda\boldsymbol{I}$ term is introduced to the ordinary least squares estimator to make $\boldsymbol{X}'\boldsymbol{X}$ nonsingular and reduce the collinearity between predictors.

Using RR-BLUP, major genes that have been identified can be treated as fixed effects (Bernardo, 2014; Zhao et al. 2014). By treating large effect QTL or major genes as fixed in the model, it allows the QTL to have their own variance and does not shrink them to the same degree as the rest of the markers in the model. This idea is very similar to the Bayesian model but does so in a less computationally demanding way.

Certain models are better suited for specific genetic architectures, traits and populations. When a trait is controlled by many small effect loci, RR-BLUP works very well. Under conditions where a few large-effect QTL explain the majority of the genetic variation, BayesB works well (Lorenz et al., 2011). In regards to the genetic control of barley yellow dwarf virus tolerance in spring oat, previous research has shown that a few large effect QTL seem to control the trait, and therefore, BayesB or fixing major QTL in RR-BLUP may be the best option for genomic selection.

Along with model selection, the development of a training population is an important step in the GS process. The training population is used to estimate the model parameters and consists of individuals with phenotypic and genotypic data. The model can then be applied to a

validation population (genotypes only) to produce GEBVs. Accuracy of the prediction models can be measured by the correlation between the GEBV and the phenotypically estimated breeding value (PEBV), $cor(GEBV, PEBV)/\sqrt{h^2}$ (Lorenz et al., 2011).

Recent advancements in marker technology in oat now allow for genome-wide coverage (Wight et al. 2010; Oliver et al., 2010). One advancement was a high density oat single-nucleotide polymorphism (SNP) array containing 6000 SNPs (Oliver et al., 2011). The genotyping assay is performed using an Infinium assay developed by Illumina. These new SNPs along with some previously mentioned markers were published in a physically anchored oat consensus map by Oliver et. al. in 2013. The SNP array has been reported to be extremely reliable and shows potential for being useful in breeding programs. Marker technology has continued to develop and genotyping-by-sequencing (GBS) has recently begun to be used in oat and appears to be how future genotyping will be performed (Huang, 2014).

With the advancements in the number of available markers and the decrease in costs of genotyping, MAS and GS have never been more appealing for use in plants (Patel et al., 2015). A limited number of studies examining the potential for MAS and GS have been reported in oat. Asoro et al. (2011) assessed accuracy of GS for several traits including β-glucan, heading date, plant height, groat percentage and yield over different numbers of markers and different training population sizes. Accuracies improved with increasing training population size and increased number of markers; which is consistent with observations in other plants and simulations (Spindel, 2015; Arruda, 2015; Hickey, 2014). Asoro et al. (2013) compared GS, MAS and pedigree-BLUP selection for β-glucan concentration in oat. On a per cycle basis, the advantage of MAS and GS over phenotypic selection was small; however, the authors concluded that any substantial advantage for GS would come from shortening the breeding cycle. Top-performing

progenies were also identified from GS and MAS compared to phenotypic selection. MAS and GS have also been examined in other small grains such as wheat and rye. MAS and GS models were compared by Heffner et al. (2011) for several agronomic traits in wheat and GS was determined to outperform MAS. In rye, GS out performed MAS for some traits; however, this was not seen in every trait such as plant height and heading time (Wang et al., 2014; Zhao et al., 2014). Arruda et. al. (2015b) compared eight models using MAS and GS for Fusarium head blight resistance (FHB) in wheat and determined GS outperformed MAS for all FHB traits.

Barley yellow dwarf (BYD) causes significant economic losses in small grains. The disease was first described in barley by Oswald and Houston in 1951 and affects all major cereal crops (rice, maize, wheat, oat and rye) as well as other grass species. The disease is caused by a group of phloem limited luteoviruses known as barley yellow dwarf viruses (BYDVs) that are obligately transmitted via aphid vectors (Hewings, 1995). Symptoms on oat depend on cultivar and environment but normally include leaf discoloration, stunted growth and blasting of florets. Economic losses due to BYDVs on oat range from 13-25 kg/ha for each 1% increase in incidence (McKirdy et al., 2002). Methods of control for BYDVs include insecticides to control aphid populations; however, insecticides may only be feasible in highly intensive agricultural systems. The most effective way to control BYD is by planting tolerant cultivars (Burnett et al., 1995).

Two types of resistance to BYDV have been distinguished: virus resistance and field resistance. Field resistance is usually referred to as tolerance (Kosova et al., 2008). Virus resistance refers to low virus titer in infected plants whereas field resistance (tolerance) refers to the reduction of symptoms of infection independent of virus titer. In this paper, resistance will be defined as reduced viral replication in infected plants (Cooper and Jones, 1983). Tolerance

will therefore be defined as the development of mild or negligible symptoms in infected plants. It can also be stated as the ability of plants to yield well under BYDV infection.

From a breeding perspective, comparing different breeding strategies is important. This is specifically true with traits that are difficult to phenotype, such as BYDV. Comparing MAS and GS for BYDV can help determine which strategy is most advantageous. In this study, GS and MAS model accuracies were compared using 428 oat lines and SNPs from a high density oat SNP array. GS models with fixed effects for identified QTL from Chapter 2 were also examined.

3.3 Materials and Methods

Plant material and disease assessment

The panel of lines was developed by the Collaborative Oat Research Enterprise (CORE), a group of research scientists from North America and several other countries worldwide. Oat breeders submitted lines to the panel which includes a high level of diversity. A total of 428 spring oat lines were included in the final panel and phenotyped for BYDV tolerance.

Tolerance for BYDV was evaluated at the University of Illinois Crop Sciences Research and Education Center, Urbana, IL in 2010 and 2011. Two replications of hills with 15 seeds per hill were planted in the BYDV nursery in a randomized complete block design. When the seedlings were in the three leaf stage (approximately 20 days after planting), the hills were inoculated with viruliferous *Rhopalosiphum padi* (L.) carrying the Illinois isolate of BYDV-PAV. After one week the plants were sprayed with insecticide (Cygon 2E™) to kill the aphids.

BYDV tolerance was evaluated after stem elongation was completed, using a scale from 0 to 9 with 0 being assigned to the most tolerant plants and 9 indicating the most sensitive (Qualset, 1983; Hewings et al., 1992).

Phenotypic data analysis

Best linear unbiased predictors (BLUPs) for BYDV tolerance were calculated using a mixed model:

$$Y_{ijk} = \mu + year_i + block(year)_{ij} + line_k + (year\ x\ line)_{ik} + \varepsilon_{ijk} \quad \text{(iv)}$$

Where Y_{ijk} is the observed BYDV phenotype, μ is the overall mean, $year_i$ is the random effect of the *i*th year, $block_j$ is the random effect of the *j*th block within the *i*th year, $line_k$is the random effect of the *k*th line, $year\ x\ line_{ik}$ is the random effect of the interaction between the *i*th year and the *k*th line, and ε_{ijk} is the random error term. BLUPs were used for the phenotypically estimated breeding values (PEBVs).

Genotypic data

The final panel of 428 spring oat lines was genotyped using a high density oat SNP array containing 6000 SNPs. Genotyping was performed at the USDA-ARS Small Grains Genotyping Lab in Fargo, ND using the Infinium assay developed by Illumina (Oliver et al., 2011). The marker number was further reduced by filtering for minor allele frequencies below 0.05 and markers with the proportion of missing genotypes greater than 0.10. Furthermore, by using the LDTagSNP Selection function in JMP Genomics 7 (SAS, Cary, NC) markers showing linkage

disequilibrium (r^2) higher than 0.8 were binned, and a representative SNP from each bin was used (Carlson et al., 2004). This process reduced the redundancy of markers explaining the same information and led to a final number of 2305 SNPs used for the GS models.

A previously developed SNP-based physically anchored consensus map was used for marker locations (Oliver et al., 2013). The map was developed from 390 recombinant inbred lines from six bi-parental populations, which included 985 SNPs and 68 previously-published markers. The final map consisted of 21 chromosomes having a total map distance of 1838.8 cM.

Quantitative trait loci (QTL) information

Six markers on chromosomes 3C and 18Dwere identified as statistically significant in Chapter 2 of this dissertation (Table 3.1). These six markers were used to develop several of the MAS models used for comparison with GS models (GMI_ES22_c20350_257, GMI_ES22_c20081_313, GMI_DS_LB_10400, GMI_DS_CC1800_254, GMI_ES17_c6498_89 and GMI_ES05_c3073_282). Two specific markers (GMI_ES22_c20081_313, "SNP3C" and GMI_ES05_c3073_282, "SNP18D") that were the most significant markers on 3C and 18D (markers used in Chapter 2 haplotype analysis) were used in combination and separately in MAS. Comparisons were also made by fixing these significant markers in GS models.

Model Comparisons

Ten models, made up of six GS models and four MAS models, were compared. The genomic selection models were all based on the RR-BLUP model but had slightly different

marker profiles. Genomic selection models included: all of 2305 SNP markers treated as random effects for GS-All; all SNPs as random with six random SNPs as fixed effects for GS-random; all SNPs as random and the six markers identified from Chapter 2 treated as fixed effects for GS-GWAS; all SNPs as random with SNP3C and SNP18D as fixed effects for GS-3C18D; all SNPs as random with SNP3C as fixed for GS-3C; all SNPs as random with SNP18D as fixed for GS-18D. The MAS models included: six markers identified from chapter 2 for MAS-GWAS; SNP3C and SNP18D for MAS-3C18D; SNP3C for MAS-3C; and SNP18D for MAS-18D.

Prediction accuracy

Five-fold cross validation was used to calculate the prediction accuracy of each model. This approach breaks up the 428 lines into five groups; four of these groups comprising the training population and one group comprising the validation population. For each model, 60 training populations were randomly selected. The five-fold cross validation with sixty iterations resulted in 300 accuracy values. Model prediction accuracy was calculated using equation (v) (Dekkers, 2007; Albrecht et al., 2011):

$$\frac{r(GEBV:PEBV)}{\sqrt{H^2}} \quad \text{(v)}$$

Where *r* is the Pearson's correlation between the GEBVs and the PEBVs in the validation population and H^2 is the broad-sense heritability, calculated from Chapter 2. Prediction accuracies were performed in R (R Development Team) using the "lm()" function for MAS models and the "mixed.solve()" function of the rrBLUP package (Endelman, 2011) for the GS models. Ryan-Einot-Gabriel-Welch test at an alpha level of 0.05 was performed for mean separation for the accuracy of the models using SAS PROC GLM. Principal component analysis

(PCA) and linear regression on the PCs was used to examine the amount of population structure in the data.

3.4 Results

Prediction accuracies were very high for all the GS models for BYDV tolerance. The five-fold cross validated accuracies are presented in Figure 3.1. The highest observed accuracy was in the GS-GWAS (0.892) and GS-3C18D (0.891), which were not significantly different (Table 3.2). GS-All (0.848) can be considered as the base model and when significant markers identified in Chapter 2 were included as fixed effects in the model, they outperformed GS-All (4.94% advantage for GS-GWAS, 4.85% advantage for GS-3C18D, 3.18% advantage for GS-3C, and 1.32% advantage for GS-18D). This was not true when six random markers were included in the model as fixed (GS-random), as there was a 1.32% reduction in prediction accuracy.

The MAS models did not perform as well as the GS models. The top performing MAS models were MAS-GWAS (0.857) and MAS-3C18D (0.856) and were not significantly different from each other. The MAS models that contained only marker, MAS-3C (0.791) and MAS-18D (0.580) had the lowest prediction accuracies of all the models but MAS-3C significantly outperformed MAS-18D. Therefore, as previously mentioned in Chapter 2, the 3C locus seems to be extremely important for prediction of BYDV tolerance.

In general the GS models outperformed the MAS models with a few exceptions. MAS-GWAS and MAS-3C18D outperformed both the GS-All and GS-random models and was also not significantly different from the GS-18D (0.859). Compared to GS-All, MAS-3C18D had the

biggest advantage of the MAS models at 1.1%. MAS-GWAS was the only other model to have an advantage over GS-All at 0.96%. MAS-3C and MAS-18D performed significantly worse than GS-ALL. MAS-3C had a 7.1% reduction in prediction accuracy and MAS-18D had a large reduction of 46.0%.

To examine structure within the population, PCA was performed on the marker data. PCA1 accounted for 8.7% of the variation in the data. The first five PCAs accounted for a total of 23.3% variation in the data and linear regression using four of the PCs revealed an r^2 of 0.4, showing that up to half of the prediction accuracy could be due to population structure.

3.5 Discussion

BYDV tolerance in oat is extremely important due to the devastating nature of the disease. Screening breeding lines for tolerance is difficult, labor intensive, time consuming and requires viruliferous aphids for inoculation. Therefore, a screening platform based on molecular markers could be a beneficial tool for breeders. This is especially important for oat breeders who may not be currently screening their germplasm for BYDV tolerance and would allow them to perform selection of lines with high BYDV tolerance without implementing a phenotyping nursery.

In this study, 2305 SNP markers, along with previously identified QTL from Chapter 2 of this dissertation, were used to create and compare several GS and MAS models for prediction of BYDV tolerance. These models were created to examine specific options for screening lines based on using all the markers available or just a few markers. This is important because some breeding programs may not have the resources available to consistently screen their breeding

lines with a SNP array, but screening breeding lines with a single or a few markers could be feasible. The results of the five-fold cross validation indicated that in general GS models with fixed effects out-performed GS with all markers treated as random effects and MAS models. Nevertheless, even though prediction accuracies of most models had significant differences, the difference between the best GS model (GS-GWAS) and the best MAS model (MAS-3C18D) was only 4.02%. This is important to examine for breeders because even though GS-GWAS was the best performing model, a MAS approach might be a better option for their breeding program and resources and would still result in high prediction accuracies. In contrast, MAS using only one marker (MAS-3C and MAS-18D) were the poorest performing models and would be less effective for selection for BYDV tolerance.

A major conclusion of this study is that fixing known large effect QTL in the genomic selection model does significantly improve the prediction accuracy of GS models. Since a major assumption of RR-BLUP is that all the markers come from the same distribution and have a common variance, fixing specific QTL in the model allows for the QTL to have their own variance. Compared to the basic RR-BLUP model treating all markers as random effects (GS-All), fixing a significant marker that has been identified for BYDV tolerance improved accuracy by between 1.3% and 4.9% depending on which QTL were included. This is consistent with Spindel et al. (2015), as they determined that using information from GWAS to develop the GS model can help improve prediction accuracy in rice. Bernardo (2014) also found that, through simulation experiments, fixing large effect QTL ($R^2 < 10\%$ and $h^2 \geq 0.50$) in a GS model was never disadvantageous to prediction accuracies.

In order to examine the inclusion of fixed effects in the GS models and to ensure that the improved prediction accuracies were not just a function of fixing any random marker in the

model, we included a model using all markers as random but randomly selecting six markers to be fixed (GS-random). Our results revealed that GS-random had a 1.3% reduction in prediction accuracy versus GS-All (basic GS model). This shows that including markers as fixed effects that are not associated with the trait of interest is a detriment to the model and it is better to allow them to be shrunken towards zero via random effects. This is consistent with results from Bernardo (2014), showed that if a QTL did not meet the threshold of $R^2 < 10\%$ and the trait $h^2 \geq 0.50$ it should not be fixed in the model. This can be true for many polygenic traits that consist of many small-effect QTL. Of the 2305 markers used in this study, and specifically the six that were deemed significant from Chapter 2, only one was above the $R^2 < 10\%$; however, four others were very close to the threshold. The extremely high heritability (0.91, from Chapter 2) that was observed in the panel likely helps the prediction accuracy of the six markers even though they explained less than 10% of the variation.

Population structure within the panel of lines has an effect on prediction accuracies. Although, we did not account for population structure in the model, the author acknowledges that within subpopulations in the panel, as much as half of the prediction accuracy could be due to structure. Previous research that identified large effect QTL for BYDV tolerance and the high heritability of the trait, the author is confident that even in the presence of population structure, the marker effects from the loci are being captured in the model and BYDV tolerance can be predicted accurately with GS and MAS.

3.6 Conclusions

Several genomic selection and marker-assisted selection models exhibited high prediction accuracy levels. Using previously identified QTL as fixed effects in the GS models (GS-GWAS, GS-3C18D, GS-3C and GS-18D), resulted in the highest prediction accuracies compared to the basic genomic selection model (GS-All). Two MAS models (MAS-3C18D and MAS-GWAS) had higher prediction accuracies than GS-All. The difference between the best GS model and the two top MAS models, although statistically significant, was minimal and therefore both MAS and GS are viable options for selection for BYDV tolerance in a spring oat breeding program.

3.7 Tables and Figures

Table 3.1 SNPs associated with BYDV tolerance in a panel of 428 spring oat lines, chromosomal position, *p*-values and marker effects.

SNP	C	cM	p^a	maf[b]	r^2	Adj. p^c	Effect
GMI_ES22_c20350_257	3C	113.2	6.84	0.40	0.07	4.29	-0.49
GMI_ES22_c20081_313	3C	114.5	17.8	0.32	0.17	14.6	0.82
GMI_DS_LB_10400	3C	115.1	6.91	0.46	0.06	4.29	-0.49
GMI_DS_CC1800_254	3C	115.1	6.93	0.41	0.07	4.29	-0.31
GMI_ES17_c6498_89	18D	132.7	3.82	0.23	0.03	1.45	-0.31
GMI_ES05_c3073_282	18D	147.7	5.25	0.19	0.06	2.8	0.50

[a] *p*-value reported on a $-\log_{10}$ scale; [b]maf = minor allele frequency; [c] False Discovery rate adjusted *p*-value

Table 3.2 Five-fold cross validated prediction accuracy mean for BYDV tolerance from marker-assisted selection (MAS) and genomic selection (GS) models and accuracy relative to the base GS model (GS-All).

Trait	GS-All	GS-Random	GS-GWAS	GS-3C18D	GS-3C	GS-18D	MAS-GWAS	MAS-3C18D	MAS-3C	MAS-18D
Prediction Accuracy	0.848	0.837	0.892	0.891	0.876	0.859	0.856	0.857	0.792	0.580
Relative to GS-All	0.00%	-1.32%	+4.94%	+4.85%	+3.18%	+1.32%	+0.96%	+1.09%	-7.08%	-46.04%
Mean Separation^	D	E	A	A	B	C	C	C	F	G

^ Ryan-Einot-Gabriel-Welch-q test at $\alpha = 0.05$ level. Models with same letter are not significantly different

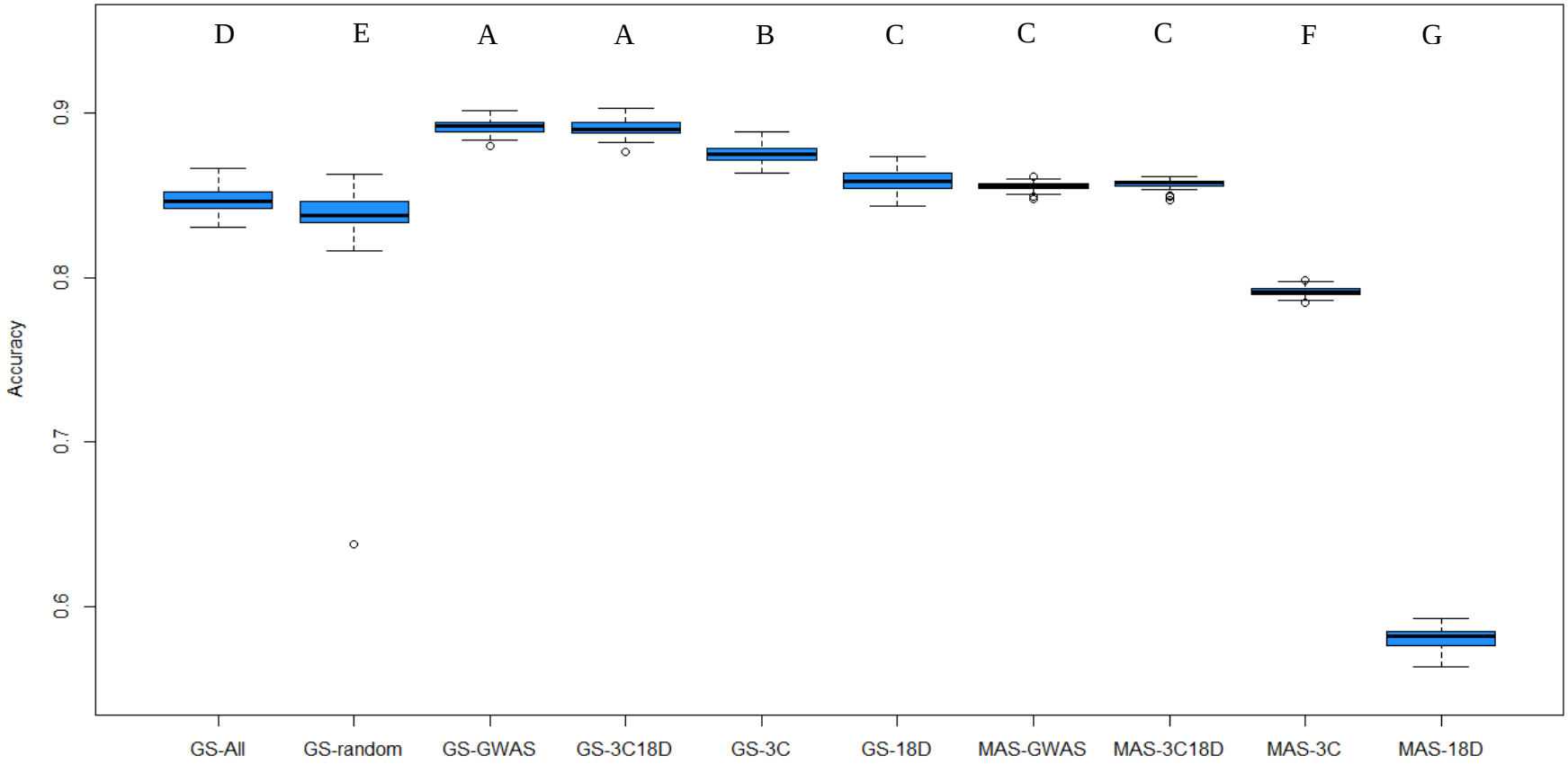

Figure 3.1 Five-fold cross validated prediction accuracies for barley yellow dwarf virus (BYDV) tolerance in spring oat. Genomic selection models performed using ridge regression-best unbiased predictor (RR-BLUP) and MAS models using ordinary least square (OLS) regression. GS-All: genomic selection with 2305 SNPs. GS-random: genomic selection with 2305 SNPs + 6 randomly selected SNPs treated as fixed effects. GS-GWAS: genomic selection with 2305 SNPs + six significant markers from chapter 2 as fixed effects. GS-3C18D: genomic selection with 2305 SNPs + SNP3C and SNP18D treated as fixed effects. GS-3C: genomic selection with 2305 SNPs + SNP3C treated as a fixed effect. GS-18D: genomic selection with 2305 SNPs + SNP18D as a fixed effect. MAS-GWAS: marker-assisted selection with six significant markers from chapter 2 using multiple linear regression. MAS-3C18D: marker-assisted selection with SNP3C and SNP18D using multiple linear regression. MAS-3C: marker-assisted selection with SNP3C using linear regression. MAS-18D: marker-assisted selection with SNP18D using linear regression. Ryan-Einot-Gabriel-Welch-q test at $\alpha = 0.05$. Models with the same letter are not significantly different.

3.8 References

Albrecht, T., V. Wimmer, H.J. Auinger, M. Erbe, C. Knaak, M. Ouzunova, H. Simianer, and C.C. Schon. 2011. Genome-based prediction of testcross values in maize. Theoretical Applied Genetics. 123:339-350.

Arruda, M.P., P.J. Brown, A.E. Lipka, A.M. Krill, C. Thurber and F.L. Kolb. 2015. Genomic selection for predicting Fusarium head blight resistance in a wheat breeding program. Plant Genetics. 8:1-12.

Arruda, Marcio P. 2015. Genomic selection and quantitative trait loci mapping for fusarium head blight resistance in wheat (*Triticum aestivum* L.). Ph.D. Dissertation. The University of Illinois at Urbana-Champaign.

Asoro, F.G., M.A. Newell, W.D. Beavis, M.P. Scott and J.-L. Jannink. 2011. Accuracy and training population design for genomic selection on quantitative traits in elite North American oat. Plant Genome. 4:132-144.

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

Bernardo, R. and J. Yu. 2007. Prospects for genome-wide selection for quantitative traits in maize. Crop Science. 47:1082-1090.

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. Genome-wide selection when major genes are known. Crop Science. 54:68-75.

Burnett, P. A. and C. J. D'Arcy. 1995. Barley Yellow Dwarf: A Brief Introduction. *In*: Barley Yellow Dwarf, 40 Years of Progress. 1-5.

Carlson, C.S., M.A Eberle, M.J. Rieder, Q. Yi, L. Kruglyak, and D.A. Nickerson. 2004. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. American Journal of Human Genetics. 74:106-120.

Cooper, J. I., and A. T. Jones. 1983. Responses of plants to viruses: proposal for the use of terms. Phytopathology. 73:127-128.

Dekkers, J.C.M. 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. Journal of Animal Breeding and Genetics. 124:331–341.

Endelman, J.B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genetics. 4:250-255.

Goddard, M.E., and B.J. Hayes. 2007. Genomic selection. Journal of Animal Breeding and Genetics. 124:323-330.

Heffner, E.L., M.E. Sorrells and J.-L. Jannink. 2009. Genomic selection for crop improvement. Crop Science. 49:1-12.

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

Heslot, N., H.-P. Yang, M.E. Sorrells and J.-L. Jannink. 2012. Genomic selection in plant breeding: a comparison of models. Crop Science. 52:14-160.

Hewings, A. D., F. L. Kolb, G. R. Gregerson and E. M. Bauske. 1992. Field research and germplasm evaluation methodology for barley yellow dwarf viruses in cereals. *In*: The Proceedings at the Fourth International Oat Conference. 44-66.

Hewings, A. D. 1995. Purification and virion characterization of barley yellow dwarf viruses. *In:* Barley Yellow Dwarf, 40 Years of Progress. 165-179.

Hickey, J.M., S.Dreisigacker, J.Crossa, S. Hearne, R. Babu, B.M. Prasanna, M.Grondona, A. Zambelli, V.S. Windhausen, K. Mathews and G. Grojanc. 2014. Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. Crop Science. 54:1476-1488.

Huang, Y.-F., J.A. Poland, C.P. Wight, E.W. Jackson and N.A. Tinker. 2014. Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. PLoS ONE 9(7):e102448. doi: 10.1371/journal.pone.0102448.

Jannink, J.-L., A.J. Lorenz and H. Iwata. 2010. Genomic selection in plant breeding: from theory to practice. Briefings in Functional Genomics. 9:166-177.

Kosova, K., J. Chrpova, and V. Sip. 2008. Recent advances in breeding of cereals for resistance to barley yellow dwarf virus - a review. Czech Journal of Genetics and Plant Breeding. 44.1:1-10.

Lipka, A.E., F. Lu, J.H. Cherney, E.S. Buckler, M.D. Casler, and D.E. Costich. 2014. Accelerating the switchgrass (*Panicum virgatum* L.) breeding cycle using genomic selection approaches. PLoS ONE 9:e112227.

McKirdy, S. J., R. A. C. Jones, and F. W. Nutter Jr. 2002. Quantification of yield losses caused by barley yellow dwarf virus in wheat and oat. Plant Disease. 86:769-73.

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

Myles, S., J. Peiffer, P.J. Brown, E.S. Ersoz, Z. Zhang, D.E. Costich and E.S. Buckler. 2009. Association mapping: Critical considerations shift from genotyping to experimental design. Plant Cell. 21:2194-2202.

Oliver, R. E, D. E. Obert, G. Hu, J. M. Bonman, E. O'Leary-Jepsen, and E. W. Jackson. 2010. Development of oat-based markers from barley and wheat microsatellites. Genome. 53:458-71.

Oliver, R. E., G. R. Lazo, J. D. Lutz, M. J. Rubenfield, N. A. Tinker, J. M. Anderson, N. H. Wisniewski-Morehead, D. Adhikary, E. N. Jellen, P. J. Maughan, G. L. Brown-Guedira, S. Chao, A. D. Beattie, M. L. Carson, H. W. Rines, D. E. Obert, J. M. Bonman, and E. W. Jackson. 2011. Model SNP development for complex genomes based on hexaploid oat using high-throughput 454 sequencing technology. BMC Genomics. 12: 77.

Oliver, R.E., Tinker, N.A., Lazo, G.R., Chao, S., Jellen, E.N., Carson, M.L., Rines, H.W., Obert, D.E., Lutz, J.D., Shackelford, I., Korol, A.B., Wight, C.P., Gardner, K.M., Hattori, J., Beattie, A.D., Bjornstad, A., Bonman, J.M., Jannink, J.-L., Sorrells, M.E., Brown-Guedira, G.L., Mitchell Fetch, J.W., Harrison, S.A., Howarth, C.J., Ibrahim, A., Kolb, F.L., McMullen, M.S., Murphy, J.P., Ohm, H.W., Rossnagel, B.G., Yan, W. Miclaus, K.J., Hiller, J., Maughan, P.J., Redman Hulse, R.R., Anderson, J.M., Islamovic, E., and E.W. Jackson. 2013. SNP discovery and chromosome anchoring provide the first physically-anchored hexaploid oat map and reveal synteny with model species. PLoS ONE 8(3): e58068. doi: 10.1371/journal.pone.0058068.

Oswald, J.W, and B.R. Houston. 1951. A New disease of cereals transmissible by aphids. Plant Disease Reporter 35: 471-475.

Patel, D., M. Zander, J. Dalton-Morgan, and J. Batley. 2015. Advances in plant genotyping: where the future will take us. p. 1-11. In: J. Batley (ed). Plant Genotyping. Springer New York.

Qualset, C.O. 1984. Evaluation and breeding methods for barley yellow dwarf resistance. In: Barley Yellow Dwarf: A Proceedings of the Workshop. 72-80.

Spindel, J., H. Begum, D. Akdemir, P. Virk, B. Collard, E. Redona, G. Atlin, J.-L. Jannink, and S.R. McCouch. 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 Genetics. 11(2): e1004982. doi: 10.1371/journal.pgen.1004982.

Wang, Y., M.F. Mette, T. Miedaner, M. Gottwald, P. Wilde, J.C. Reif, and Y. Zhao. 2014. The accuracy of prediction of genomic selection in elite hybrid rye populations surpasses the accuracy of marker-assisted selection and is equally augmented by multiple field evaluation locations and test years. BMC Genomics. 15:556.

Whittaker, J.C., R. Thompson, and M.C. Denham. 2000. Marker-assisted selection using ridge regression. Genetic Research (Cambridge). 75:249–252.

Wight, C. P., W. Yan, J. M. Fetch, J. Deyl and N. A. Tinker. 2010. A set of new simple sequence repeat and Avenin DNA markers suitable for mapping and fingerprinting studies in oat (*Avena* Spp.). Crop Science. 50:1207-18.

Zhao, Y., M.F. Mette, M. Gowda, C.F.H. Longin, and J.C. Reif. 2014. Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. Heredity 112:638-645.

CHAPTER 4

GENOMIC PREDICTION FOR BARLEY YELLOW DWARF VIRUS TOLERANCE IN SPRING OAT (*AVENA SATIVA* L.) FOR CROSS EVALUATION

4.1 Abstract

Significant yield loss is attributed to barley yellow dwarf (BYD), a disease that affects most cereal grains including spring oat (*Avena sativa* L.). Phenotyping for sensitivity to the disease is extremely difficult and requires inoculation with viruliferous aphids. Genomic selection (GS) provides an alternative strategy for identifying tolerant lines and improving breeding for BYDV. GS can also be used to identify potential crosses that have high tolerance levels as well as high genetic variance giving breeders an ability to evaluate crosses before actually making them. In this study, 2138 single nucleotide polymorphisms (SNPs) were used on a panel of 519 spring oat lines for barley yellow dwarf virus tolerance. Of the 519 oat lines, 428 lines had genotypes and phenotypes while 91 of the oat lines were just genotyped. Using the R package "PopVar", several GS models were compared for prediction accuracy. The BayesA model was identified as having the highest prediction accuracy and genomic estimated breeding values (GEBVs) were calculated for the 519 lines using the BayesA model. The top 10% of lines (52 lines) based on GEBVs for BYDV tolerance were selected to perform simulated crosses. A total of 1326 crosses were simulated and the mean, genetic variance and mean of high/low superior progenies were calculated. From the 1326 crosses, 22 crosses were identified as having a balance between a low predicted mean (high tolerance) and high genetic variance.

Because of the high tolerance and high genetic variance, the chance of obtaining transgressive segregants for BYDV tolerance is higher in these crosses. Using GS and simulated crosses gives breeders another tool to improve breeding efficiency for BYDV tolerance and allows for better allocation of time and resources within the breeding program.

4.2 Introduction

Genomic selection (GS) and marker-assisted selection (MAS) have become common plant breeding strategies for the selection of superior individuals. Proposed by Meuwissen et al. (2001) in animal breeding, GS is an extension of MAS and involves predicting breeding values of individuals using molecular markers. With genome-wide markers the assumption is that all QTL for a trait should be accounted for by at least one marker (Goddard and Hayes, 2007). This can provide an advantage over marker-assisted selection for traits that are controlled by many small effect loci. GS employs the use of a training population comprised of genotypes and phenotypes to calculate the model parameters. The model is then applied to a validation population, consisting of individuals for which only genotype information is available, in order to calculate genomic estimated breeding values (GEBVs).

With the increasing number of available markers to breeders, the numbers of markers often exceed the number of phenotypes (large *p* small *n* problem). This can lead to overfitting and multicollinearity in the model. To address this issue, models that depend on shrinkage factors are used to estimate the marker effects. Examples of models that do this include, ridge regression best unbiased linear predictor (RR-BLUP), genomic best unbiased linear predictor (G-

BLUP), least absolute shrinkage and selection operator (LASSO), elastic-net and several Bayesian models (BayesA, BayesB, BayesC) (Lorenz et al., 2011; Heslot et al, 2012).

RR-BLUP is the most popular model and was first described by Whittaker et al. (2000). An assumption in the RR-BLUP model is that the markers are from the same distribution and therefore have a common variance. As a result, the markers are equally shrunken towards zero (Bernardo and Yu, 2007; Heffner et al., 2009). Although, it is recognized that this is not truly realistic in plants, it has been successfully used as a model in plants and in simulations (Arruda et al., 2015; Lipka et al., 2014; Heffner et al., 2009). RR-BLUP uses the model in equation (i):

$$g(\boldsymbol{x}_i) = \sum_{k=1}^{p} x_{ik}\beta_k \quad \text{(i)}$$

Where x_{ik} signifies the score for SNP k in the individual i, β_k is the effect linked with marker k, and the genetic value of the sum of the p marker effects (Lorenz et al., 2011). β is estimated using a modification to the normal least squares estimators in equation (ii):

$$\hat{\beta} = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{\lambda I})^{-1}\boldsymbol{X}'\boldsymbol{y} \quad \text{(ii)}$$

Where $\boldsymbol{X}$ is an incidence matrix associated with individuals, $\boldsymbol{I}$ is an identity matrix, and $\boldsymbol{y}$ is a vector of estimated breeding values that usually are phenotypes. The $\lambda\boldsymbol{I}$ term is introduced to the ordinary least squares estimator to make $\boldsymbol{X}'\boldsymbol{X}$ nonsingular and reduce the collinearity between predictors.

The Bayesian models differ from the RR-BLUP models by allowing different variances for the predictor variables (Meuwissen et al., 2001). This allows each marker to be shrunken towards zero to different degrees. BayesA, BayesB and BayesC are three slightly different types of Bayesian models with BayesA being the most basic model that allows for different variances for each predictor variable. BayesB is similar to BayesA, but each marker is given a probability

that the effect is zero. BayesC uses the data to estimate the probability that each marker effect is zero. The basic Bayesian model (BayesA) is shown in equation (iii)

$$g(\boldsymbol{x}_i) = \sum_{k=1}^{p} x_{ik}\beta_k\gamma_k \qquad \text{(iii)}$$

Where $\gamma_k$is an indicator variable specifying the presence of marker k in the prediction model. All other variables are the same as in equation (i). It is assumed that β_k follows a normal distribution with mean of zero and finite variance. The variance of β_k follows a mixture distribution:

$$var(\beta_k) = 0, \qquad with\ probability\ \pi$$

$$var(\beta_k) \sim X^{-2}(v, S), \qquad with\ probability\ (1-\pi)$$

When $\pi = 0$, BayesB reduces to BayesA.

Certain models are better suited for specific genetic architectures, traits and populations. RR-BLUP works very well when a trait is controlled by many small effect loci. BayesB works well under conditions where a few large-effect QTL explain the majority of the genetic variation (Lorenz et al., 2011). Previous research has shown that a few large effect QTL seem to control barley yellow dwarf virus tolerance in spring oat, therefore, BayesB may be the best option for genomic selection in this case.

The ability to identify superior genotypes using GS is extremely important for breeders, but it is also important to identify the best crosses to make. The ability to predict the genetic variance in bi-parental crosses from genotypic data allows the breeder to make theoretical crosses and before they making the actual crosses with the plants. Recently, a package in R named PopVar was created that allows breeders to simulate bi-parental crosses and output the mean of the progeny, mean of the superior progeny and the genetic variance based on GS models (Mohammadi et al., 2015).

Recent advancements in marker technology in oat now allow for genome-wide coverage (Wight et al. 2010; Oliver et al., 2010). A high density oat single-nucleotide polymorphism (SNP) array containing 6000 SNPs has been developed (Oliver et al., 2011). An Infinium assay developed by Illumina is used to perform the genotyping assay. Oliver et. al. (2013) published in a physically anchored oat consensus map using these new SNPs along with some previously mentioned markers. The SNP array has been reported to be extremely reliable and shows potential for being useful in breeding programs. Genotyping-by-sequencing (GBS) has recently begun to be used in oat and appears to be how future genotyping will be performed (Huang, 2014).

With advancements in the number of available markers and the decreasing costs of genotyping, GS is more appealing for use in plants (Patel et al., 2015). The potential of GS in oat has been reported only in a limited number of studies. Asoro et al. (2011) assessed the accuracy of GS for several traits including β-glucan, heading date, plant height, groat percentage and yield. In their study, accuracies improved with increasing training population size and increased number of markers; which is consistent with observations in other plants and simulations (Spindel, 2015; Arruda et al., 2015a; Hickey, 2014). For β-glucan concentration in oat, Asoro et al. (2013) also compared GS, MAS and pedigree-BLUP selection. On a per cycle basis, the advantage of MAS and GS over phenotypic selection was small; however, the authors concluded that any substantial advantage for GS would come from shortening the breeding cycle through the use of GS. GS and MAS also effectively identified top-performing progenies. MAS and GS have also been examined in other small grains such as wheat and rye. For several agronomic traits in wheat, Heffner et al. (2011) compared MAS and GS models and found that GS outperformed MAS. Arruda et. al. (2015b) compared eight models using MAS and GS for

Fusarium head blight (FHB) resistance in wheat determining GS outperformed MAS for all FHB traits. In rye, GS out performed MAS for some traits; however, this was not seen in every trait such as plant height and heading time (Wang et al., 2014; Zhao et al., 2014).

Using GS to identify superior lines is a powerful tool for breeders to utilize; however, the ability to predict genetic variance and superior progeny from a cross is also extremely important and could help accelerate greater genetic gains. Historically, several methods have been used to attempt to predict genetic variance in a population; these include phenotypic distance, and genetic distance. Tiede et al., (2015) examined several historical methods and several modern methods, including a simulation based method called "PopVar" (Mohammadi et al., 2015). PopVar uses GS models using phenotypes and genome-wide markers to simulate RIL progeny populations for all pairwise crosses between parents. It then estimates marker effects and GEBVs of the progeny. Based on the GEBVs, a mean, genetic variance and mean of the superior progeny are calculated.

Barley yellow dwarf (BYD) was first described in barley by Oswald and Houston in 1951 and affects all major cereal crops (rice, maize, wheat, oat and rye) as well as other grass species. The disease is caused by a group of phloem limited luteoviruses known as barley yellow dwarf viruses (BYDVs) that are obligately transmitted via aphid vectors (Hewings, 1995). Symptoms on oat depend on cultivar and environment but normally include leaf discoloration, stunted growth and blasting of florets. Economic losses due to BYDVs in oat have been reported to range from 13-25 kg/ha for each 1% increase in incidence (McKirdy et al., 2002). Methods of control for BYDVs include insecticides to control aphid populations; however, insecticides may only be feasible in highly intensive agricultural systems. The most effective way to control BYD is by planting tolerant cultivars (Burnett et al., 1995).

Two types of resistance to BYDV have been distinguished: virus resistance and field resistance. Field resistance is usually referred to as tolerance (Kosova et al., 2008). Field resistance (tolerance refers to the reduction of symptoms of infection independent of virus titer. Virus resistance refers to low virus titer in infected plants. In this paper, resistance will be defined as reduced viral replication in infected plants (Cooper and Jones, 1983). Tolerance will therefore be defined as the development of mild or negligible symptoms in infected plants the ability of plants to yield under BYDV infection.

In this study, GS models were compared using PopVar and then the progeny of selected crosses were simulated. Using 1328 molecular markers, BYDV tolerance was predicted for 519 spring oat lines, and 1326 potential crosses were simulated based on the top performing GEBVs. The results of this study will provide breeders with another tool to use in making important decisions in the breeding program. Two major objectives were examined in this study: (1) determine which model has the highest prediction accuracy and calculate GEBVs and (2) simulate the predicted performance of potential crosses.

4.3 Materials and Methods

Plant material and disease assessment

The panel of lines from Chapter 3, developed by the Collaborative Oat Research Enterprise (CORE) was used in this study. In addition to the 428 lines, 91 oat breeding lines used as parents in the University of Illinois oat breeding program were added to the panel. These

91 lines were genotyped but not phenotyped and therefore only used in the validation populations and not the training populations.

BYDV tolerance for each line in the 428 line panel was evaluated at the University of Illinois Crop Sciences Research and Education Center, Urbana, IL in 2010 and 2011. Using a randomized complete block design, two replications were planted in the BYDV nursery with 15 seeds per hill. The hills were inoculated with viruliferous *Rhopalosiphum padi* (L.) carrying the Illinois isolate BYDV-PAV when the seedlings were in the three leaf stage (approximately 20 days after planting). The insecticide (Cygon 2E™) (Active ingrediaent: Dimethoate [0,0-dimethyl-S-(N-methylcarbamoylmethyl) phosphorodithioate]) was sprayed on the plants one week after inoculation to kill the aphids. After stem elongation was completed, BYDV tolerance was evaluated, using a scale from 0 to 9 with 0 being assigned to the most tolerant plants and 9 indicating the most sensitive (Qualset, 1983; Hewings et al., 1992).

Phenotypic data analysis

Best linear unbiased predictors (BLUPs) for BYDV tolerance phenotypes were calculated using a mixed model:

$$Y_{ijk} = \mu + year_i + block(year)_{ij} + line_k + (year\ x\ line)_{ik} + \varepsilon_{ijk} \quad \text{(iv)}$$

Where Y_{ijk} is the observed BYDV phenotype, μ is the overall mean, $year_i$ is the random effect of the *i*th year, $block_j$ is the random effect of the *j*th block within the *i*th year, $line_k$is the random effect of the *k*th line, $year\ x\ line_{ik}$ is the random effect of the interaction between the *i*th year and the *k*th line, and ε_{ijk} is the random error term. BLUPs were used for the phenotypically estimated breeding values (PEBVs).

Genotypic data

The final panel of 519 spring oat lines were genotyped using a high density oat SNP array containing 6000 SNPs. Genotyping was performed at the USDA-ARS Small Grains Genotyping Lab in Fargo, ND using the Infinium assay developed by Illumina (Oliver et al., 2011). The marker number was reduced by filtering for minor allele frequencies below 0.05 and markers with the proportion of missing genotypes greater than 0.10. Furthermore, by using the LDTagSNP Selection function in JMP Genomics 7 (SAS, Cary, NC) markers showing linkage disequilibrium (r^2) higher than 0.8 were binned, and a representative SNP from each bin was used (Carlson et al., 2004). This process was to help reduce the redundancy of markers explaining the same information and led to a final number of 2138 SNPs used for the GS models.

A previously developed SNP-based physically anchored consensus map was used for marker locations (Oliver et al., 2013). The map was developed from 390 recombinant inbred lines from six bi-parental populations, which included 985 SNPs and 68 previously-published markers. The final map consisted of 21 chromosomes having a total map distance of 1838.8 cM.

Model design and breeding values

A five-fold cross validation scheme was used to calculate the prediction accuracy of each model. This approach breaks up 519 lines into five groups; four of these groups comprising the training population and one group comprising the validation population. The 91 oat breeding lines that have been used as parental lines were only used in the validation populations since the

phenotypic data were not included. This procedure was performed using the "x.val" function in the R package PopVar (Mohammadi et al., 2015). This function has the ability to cross validate several GS models at the same time. These models include rrBLUP, BayesA, BayesB, BayesC, Bayesian lasso and Bayesian ridge regression. The model with the highest accuracy was then used to calculate the GEBVs for the panel of lines. The top 10% of the lines (52 lines) based on GEBVs were then used for cross evaluation. All pairwise crosses between these 52 lines were then simulated using the "pop.predict" function of PopVar (Mohammadi et al., 2015). A total of 1326 crosses were evaluating by simulating 25 populations of 250 individuals for each cross. For each cross, a midparent GEBV, predicted mean of the progeny, predicted genetic variance, and the mean of the high/low superior progeny were calculated with the accompanying standard deviations. The mean of the high/low superior progeny is the average of the respective high/low tails of the distribution of the progeny. The tails were considered to be the highest 5% and lowest 5% which is important for the identification of transgressive segregants.

4.4 Results

Several models were compared using 5-fold cross validation in PopVar. The prediction accuracies and standard deviations are presented in Table 4.1. All models exhibited high prediction accuracies with the BayesA model having the highest prediction accuracy of 0.866. Using the BayesA model, GEBVs were predicted for the panel of 519 lines. The top 10% of lines (52 lines) based on GEBVs are presented in Table 4.2, and the complete list of GEBVs is presented in Appendix B. SD031128 was the top performing line with a GEBV of 1.936 while the last line in the top 10% was IL09-2756 with a GEBV of 2.680.

Using the top 10% of lines based on GEBVs, all-pairwise crosses were simulated to produce a mean of the progeny, genetic variance of the progeny and mean of the high/low superior progeny. This resulted in a total of 1326 crosses (Appendix B). The cross SD031128/MN05155, had the lowest predicted mean for BYDV tolerance at 2.01. The highest predicted mean was observed in the cross Buckskin/MN08268 with a mean of 2.71. It is also important to examine the crosses with the greatest genetic variance as well as the lowest mean for the low superior progeny. The cross LAO-1104-028C1/IL11-7150 was predicted to have the highest genetic variance at 0.0451 and the lowest mean of the low superior progeny at 1.77. On the opposite end, the cross IL05-10015/IL09-5737 had the smallest genetic variance at 0.000057 and the highest mean of the low superior progeny at 2.49. Figure 4.1 shows the relationship between the predicted mean of the progeny and the predicted genetic variance. Balancing the predicted mean and genetic variance is important; therefore, 22 lines were selected with low predicted mean for BYDV tolerance and higher genetic variance (shown in the selection box of Figure 4.1 and Table 4.3).

4.5 Discussion

BYDV tolerance in oat is extremely important because the disease can result in significant economic losses in grain yield and grain quality. Evaluation of breeding lines for BYDV tolerance is difficult, labor intensive, time consuming and requires viruliferous aphids for inoculations. Therefore, a screening platform based on molecular markers could be a beneficial tool for breeders. This is especially important for oat breeders who may not be currently evaluating their germplasm for BYDV tolerance and would allow them to select lines with high

BYDV tolerance without implementing a phenotyping program. The ability to predict mean and genetic variance of progeny from simulated crosses using these marker platforms is also an important tool that oat breeders can use to determine which crosses to make.

In this study, using an R package called PopVar, 2138 markers were used on a panel of 519 (428 with genotypes and phenotypes, 91 with genotypes only) oat lines to develop genomic prediction models for BYDV tolerance. A five-fold cross validation scheme was used to calculate the prediction accuracy for a RR-BLUP, BayesA, BayesB, BayesC, BL and BRR models. The BayesA model outperformed the other models in prediction accuracy (Table 4.1) and was used to calculate GEBVs for all of the lines in the panel. Ranking the lines based on their GEBV gives insight into which lines have the marker profiles that theoretically are the best for BYDV tolerance.

One interesting aspect of this study is that 91 oat parental lines were included with genotypes only. Because of this, they were not used in training the GS model and only used in the validation population. Using the large panel of lines with genotypes and phenotypes to train the model, we are able to predict the 91 lines for BYDV tolerance without any knowledge of their phenotypic values. For breeders, information such as this is invaluable. Of course, at some point during the breeding program these lines should be phenotyped to ensure that they convey tolerance and that their genotypes agree with their phenotypic performance, but the ability to screen breeding lines based on genotypes only will prove to be a useful tool.

Based on GEBVs, the top 10% of lines (52 lines, Table 4.2) were selected to move forward with simulation of crosses. All pairwise crosses were simulated between the 52 lines which resulted in a total of 1326 crosses. From each cross, the predicted mean, genetic variance and mean of the high/low superior progeny were calculated. From the 1326 crosses, 22 crosses

were selected and identified as having a low GEBV (high tolerance) and high genetic variance (Figure 4.1). These crosses would provide breeders the best opportunity to have the lowest progeny mean for BYDV tolerance while maintaining genetic variance at a level where transgressive segregants could be attainable.

Tools such as GS and cross prediction give plant breeders another approach to improving genetic gain for specific traits. While BYDV tolerance is only one trait that is desired in an oat variety, the ability to predict tolerance in lines before actually making the real crosses and identification of BYDV tolerant lines based on genotyping is extremely beneficial to time and resource allocations. It must be stressed that these are merely tools to be used in decision-making along with all of the other tools that breeders have at their disposal, but the advantage of having more pieces of information for a line and obtaining that information easily and quickly can have a substantial impact on the breeding program.

4.6 Conclusions

Using the BayesA model and 2138 molecular markers, genomic estimated breeding values were calculated for 519 spring oat lines in order to predict barley yellow dwarf virus tolerance. The top 10% (52 lines) of the lines based on their GEBVs, were then used to predict the progeny of all pairwise (1326) crosses. Of these 1326 crosses, 22 were identified as having low predicted means of the progeny for BYDV tolerance along with high genetic variance. Moving forward, these crosses can be made and evaluated to continue to improve breeding for BYDV tolerance in spring oat.

4.7 Tables and Figures

Table 4.1 Genomic selection model comparison for barley yellow dwarf virus tolerance in spring oat. Models include: ridge regression best unbiased predictor (rrBLUP), Bayes A, Bayes B, Bayes C, Bayesian LASSO (BL), and Bayesian ridge regression (BRR).

Model	rrBLUP	BayesA	BayesB	BayesC	BL	BRR
Average Accuracy	0.812	0.866	0.859	0.812	0.854	0.834
Standard Deviation	0.033	0.038	0.04	0.048	0.048	0.056

Table 4.2 Top 10% of spring oat lines (52 lines) based on genomic estimated breeding values (GEBV) using BayesA model for barley yellow dwarf virus tolerance.

Rank	Line	GEBV	Rank	Line	GEBV
1	SD031128	1.936	**27**	IL00-654	2.503
2	SD110466	2.059	**28**	IL12-10931	2.505
3	MN05155	2.138	**29**	Corral	2.515
4	IL05-9330	2.146	**30**	MN11211	2.520
5	IL12-1919	2.176	**31**	IL09-3929	2.522
6	IL08-9442	2.184	**32**	IL08-2934	2.527
7	IL11-5534	2.284	**33**	ND030078	2.533
8	IL05-10015	2.310	**34**	MN08146	2.545
9	IL09-5737	2.312	**35**	IL07-8721	2.557
10	IL11-6317	2.312	**36**	SD111946	2.567
11	IL12-7416	2.312	**37**	Saber	2.587
12	MN06105	2.317	**38**	IL86-5698 (PI539875)	2.606
13	LAO-1104-028C1	2.325	**39**	IL12-5568	2.609
14	IL12-7401	2.328	**40**	IL11-7655	2.613
15	IL2901 (PI641968)	2.333	**41**	IL12-10307	2.615
16	IL3555 (PI641970)	2.337	**42**	IL11-7150	2.616
17	IL05-3928	2.343	**43**	IL12-7530	2.620
18	IL11-6469	2.362	**44**	IL08-6344	2.629
19	OA1130-1	2.374	**45**	IL2815 (PI641965)	2.641
20	IL05-10069	2.375	**46**	IL12-9020	2.646
21	IL09-5745	2.391	**47**	IL12-6448	2.649
22	IL10-5863	2.395	**48**	IL04-4410	2.653
23	059A1-2-2-4	2.404	**49**	Buckskin	2.655
24	IL04-3664	2.439	**50**	IL06-5433	2.673
25	IL08-10563	2.457	**51**	MN08268	2.678
26	MN06203	2.478	**52**	IL09-2756	2.680

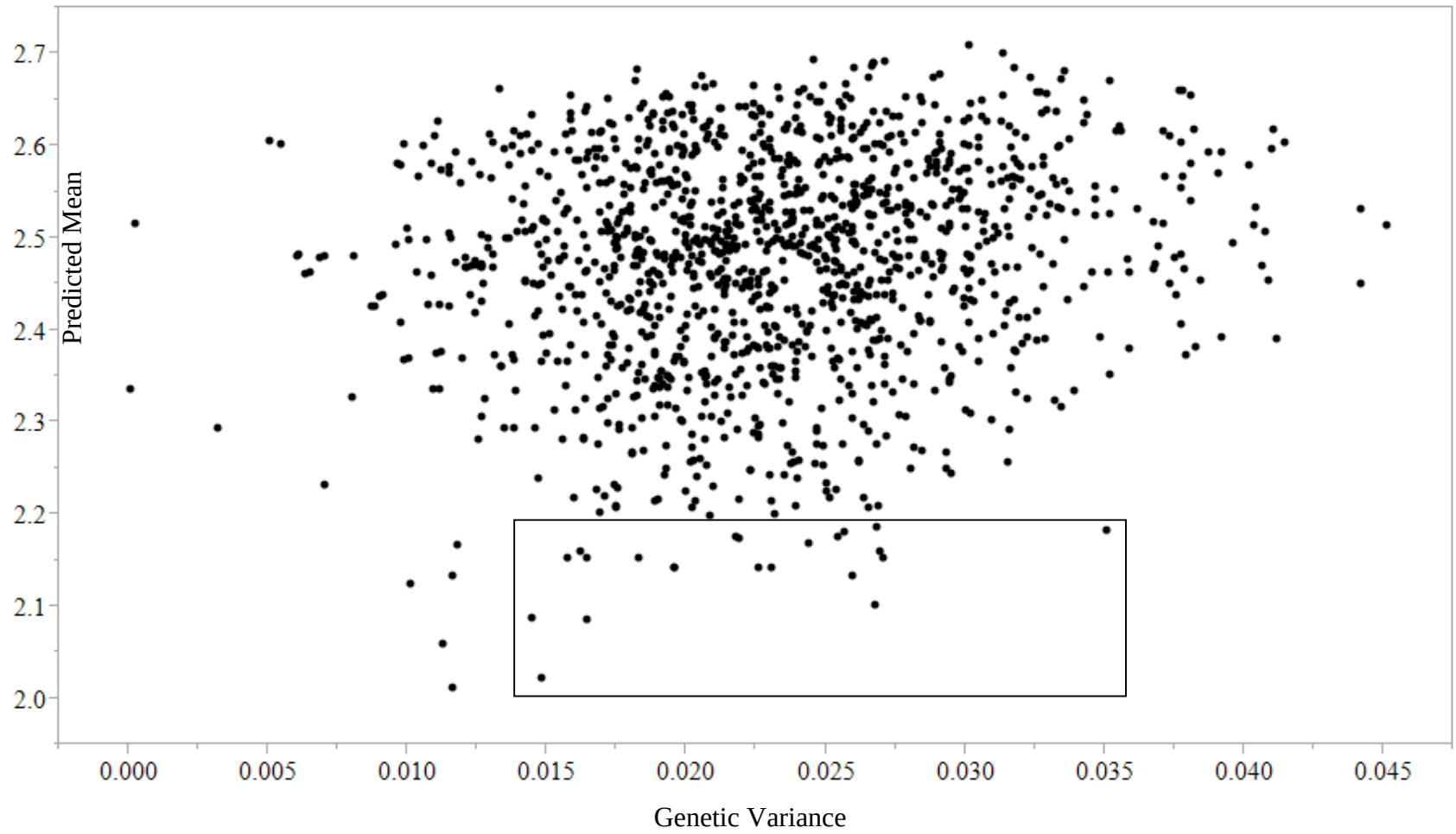

Figure 4.1 Predicted mean and predicted genetic variance of the progeny for all pairwise crosses between 52 spring oat lines for barley yellow dwarf virus tolerance. Crosses within the box consisted of low predicted mean and high genetic variance and are presented in Table 4.3.

Table 4.3 Selected crosses based on predicted cross mean and genetic variance for barley yellow dwarf virus tolerance in spring oat.

Cross	Midparent GEBV	Predicted Mean	Predicted Genetic Variance	Mean of Low 5%	Mean of High 5%
SD031128 / SD110466	2.019	2.023	0.015	1.783	2.272
SD110466 / MN05155	2.087	2.085	0.016	1.827	2.341
SD031128 / IL12-1919	2.085	2.087	0.014	1.847	2.334
SD031128 / IL08-9442	2.098	2.102	0.027	1.775	2.434
SD031128 / IL3555 (PI641970)	2.133	2.134	0.026	1.808	2.454
SD031128 / IL09-5737	2.139	2.141	0.020	1.857	2.420
SD031128 / IL11-6469	2.139	2.141	0.023	1.836	2.445
SD031128 / IL05-10015	2.140	2.142	0.020	1.857	2.422
SD031128 / IL11-6317	2.142	2.143	0.023	1.834	2.448
SD031128 / IL2901 (PI641968)	2.150	2.152	0.027	1.823	2.473
SD031128 / IL12-7416	2.150	2.152	0.018	1.883	2.430
MN05155 / IL12-1919	2.153	2.153	0.016	1.895	2.408
SD031128 / IL12-7401	2.152	2.153	0.016	1.902	2.407
SD110466 / IL12-1919	2.161	2.160	0.016	1.909	2.419
SD031128 / MN06105	2.156	2.160	0.027	1.828	2.492
MN05155 / IL08-9442	2.166	2.168	0.024	1.855	2.483
SD031128 / IL09-5745	2.171	2.174	0.022	1.882	2.469
SD110466 / IL08-9442	2.175	2.175	0.022	1.879	2.475
SD031128 / IL11-5534	2.171	2.176	0.025	1.846	2.497
SD031128 / IL05-3928	2.179	2.182	0.026	1.856	2.497
SD031128 / LAO-1104-028C1	2.183	2.183	0.035	1.801	2.558
SD031128 / 059A1-2-2-4	2.185	2.187	0.027	1.861	2.516

4.8 References

Arruda, M.P., P.J. Brown, A.E. Lipka, A.M. Krill, C. Thurber and F.L. Kolb. 2015. Genomic selection for predicting Fusarium head blight resistance in a wheat breeding program. Plant Genetics. 8:1-12.

Asoro, F.G., M.A. Newell, W.D. Beavis, M.P. Scott and J.-L. Jannink. 2011. Accuracy and training population design for genomic selection on quantitative traits in elite North American oat. Plant Genome. 4:132-144.

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

Bernardo, R. and J. Yu. 2007. Prospects for genome-wide selection for quantitative traits in maize. Crop Science. 47:1082-1090.

Burnett, P. A. and C. J. D'Arcy. 1995. Barley Yellow Dwarf: A Brief Introduction. *In*: Barley Yellow Dwarf, 40 Years of Progress. 1-5.

Carlson, C.S., M.A Eberle, M.J. Rieder, Q. Yi, L. Kruglyak, and D.A. Nickerson. 2004. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. American Journal of Human Genetics. 74:106-120.

Cooper, J. I., and A. T. Jones. 1983. Responses of Plants to Viruses: Proposal for the Use of Terms. Phytopathology. 73:127-128.

Goddard, M.E., and B.J. Hayes. 2007. Genomic selection. Journal of Animal Breeding and Genetics. 124:323-330.

Heffner, E.L., M.E. Sorrells and J.-L. Jannink. 2009. Genomic selection for crop improvement. Crop Science. 49:1-12.

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

Heslot, N., H.-P. Yang, M.E. Sorrells and J.-L. Jannink. 2012. Genomic selection in plant breeding: a comparison of models. Crop Science. 52:14-160.

Hewings, A. D., F. L. Kolb, G. R. Gregerson and E. M. Bauske. 1992. Field research and germplasm evaluation methodology for barley yellow dwarf viruses in cereals. *In*: The Proceedings at the Fourth International Oat Conference. 44-66.

Hewings, A. D. 1995. Purification and virion characterization of barley yellow dwarf viruses. *In:* Barley Yellow Dwarf, 40 Years of Progress. 165-179.

Hickey, J.M., S.Dreisigacker, J.Crossa, S. Hearne, R. Babu, B.M. Prasanna, M.Grondona, A. Zambelli, V.S. Windhausen, K. Mathews and G. Grojanc. 2014. Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. Crop Science. 54:1476-1488.

Huang, Y.-F., J.A. Poland, C.P. Wight, E.W. Jackson and N.A. Tinker. 2014. Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. PLoS ONE 9(7):e102448. doi: 10.1371/journal.pone.0102448.

Kosova, K., J. Chrpova, and V. Sip. 2008. Recent advances in breeding of cereals for resistance to barley yellow dwarf virus - a review. Czech Journal of Genetics and Plant Breeding. 44.1:1-10.

Lipka, A.E., F. Lu, J.H. Cherney, E.S. Buckler, M.D. Casler, and D.E. Costich. 2014. Accelerating the switchgrass (*Panicum virgatum* L.) breeding cycle using genomic selection approaches. PLoS ONE 9:e112227.

McKirdy, S. J., R. A. C. Jones, and F. W. Nutter Jr. 2002. Quantification of yield losses caused by barley yellow dwarf virus in wheat and oat. Plant Disease. 86:769-73.

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

Oliver, R. E, D. E. Obert, G. Hu, J. M. Bonman, E. O'Leary-Jepsen, and E. W. Jackson. 2010. Development of oat-based markers from barley and wheat microsatellites. Genome. 53.6:458-71.

Oliver, R. E., G. R. Lazo, J. D. Lutz, M. J. Rubenfield, N. A. Tinker, J. M. Anderson, N. H. Wisniewski-Morehead, D. Adhikary, E. N. Jellen, P. J. Maughan, G. L. Brown-Guedira, S. Chao, A. D. Beattie, M. L. Carson, H. W. Rines, D. E. Obert, J. M. Bonman, and E. W. Jackson. 2011. Model SNP development for complex genomes based on hexaploid oat using high-throughput 454 sequencing technology. BMC Genomics. 12: 77.

Oliver, R.E., Tinker, N.A., Lazo, G.R., Chao, S., Jellen, E.N., Carson, M.L., Rines, H.W., Obert, D.E., Lutz, J.D., Shackelford, I., Korol, A.B., Wight, C.P., Gardner, K.M., Hattori, J., Beattie, A.D., Bjornstad, A., Bonman, J.M., Jannink, J.-L., Sorrells, M.E., Brown-Guedira, G.L., Mitchell Fetch, J.W., Harrison, S.A., Howarth, C.J., Ibrahim, A., Kolb, F.L., McMullen, M.S., Murphy, J.P., Ohm, H.W., Rossnagel, B.G., Yan, W. Miclaus, K.J., Hiller, J., Maughan, P.J., Redman Hulse, R.R., Anderson, J.M., Islamovic, E., and E.W. Jackson. 2013. SNP discovery and chromosome anchoring provide the first physically-anchored hexaploid oat map and reveal synteny with model species. PLoS ONE 8(3): e58068. doi: 10.1371/journal.pone.0058068.

Patel, D., M. Zander, J. Dalton-Morgan, and J. Batley. 2015. Advances in plant genotyping: where the future will take us. p. 1-11. In: J. Batley (ed). Plant Genotyping. Springer New York.

Qualset, C.O. 1984. Evaluation and breeding methods for barley yellow dwarf resistance. In: Barley Yellow Dwarf: A Proceedings of the Workshop. 72-80.

Spindel, J., H. Begum, D. Akdemir, P. Virk, B. Collard, E. Redona, G. Atlin, J.-L. Jannink, and S.R. McCouch. 2015. Genomic selection and association mapping in rice (*Oryza sativa* L.): 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 Genetics. 11(2): e1004982. doi: 10.1371/journal.pgen.1004982.

Tiede, T., L. Kumar, M. Mohammadi, and K.P. Smith. 2015. Predicting genetic variance in bi-parental populations is more accurate when explicitly modeling the segregation of informative genomewide markers. Molecular Breeding. 25:199

Wang, Y., M.F. Mette, T. Miedaner, M. Gottwald, P. Wilde, J.C. Reif, and Y. Zhao. 2014. The accuracy of prediction of genomic selection in elite hybrid rye populations surpasses the accuracy of marker-assisted selection and is equally augmented by multiple field evaluation locations and test years. BMC Genomics. 15:556.

Whittaker, J.C., R. Thompson, and M.C. Denham. 2000. Marker-assisted selection using ridge regression. Genetic Research (Cambridge). 75:249–252.

Wight, C. P., W. Yan, J. M. Fetch, J. Deyl and N. A. Tinker. 2010. A set of new simple sequence repeat and Avenin DNA markers suitable for mapping and fingerprinting studies in oat (*Avena* Spp.). Crop Science. 50:1207-18.

Zhao, Y., M.F. Mette, M. Gowda, C.F.H. Longin, and J.C. Reif. 2014. Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. Heredity 112: 638-645.

APPENDIX A

Table A.1 Best linear unbiased predictors (BLUPs) and haplotypes for 428 spring oat lines for barley yellow dwarf virus tolerance.

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
Chinese 4	AAFC-Ottawa	Chinese 4	4.73	-	-	--
OA1058-4	AAFC-Ottawa	AC Aylmer/06897	6.53	-	-	--
Cantal	AAFC-Ottawa	QO190.2/QO189.5	7.66	-	-	--
Vao-48	AAFC-Ottawa	VAO-1/OA516-2	7.66	-	-	--
OA1272-1	AAFC-Ottawa	07877-3-13-4/07878-3-7-1	3.82	-	+	-+
OA1250-2	AAFC-Ottawa	Souris/07836-2-5-2-1a-2	4.95	-	+	-+
OA1250-1	AAFC-Ottawa	Souris/07836-2-5-2-1a-2	5.40	-	+	-+
VAO -58	AAFC-Ottawa	AC Gwen/Gehl	5.40	-	+	-+
Nice	AAFC-Ottawa	Sylva//PGR8648/Woodstock/3/Donegal	5.86	-	+	-+
Navaro	AAFC-Ottawa	Navaro	6.08	-	+	-+
OA1226-4	AAFC-Ottawa	07836-2-9-3-1a-2/ND960620	6.31	-	+	-+
VAO-51	AAFC-Ottawa	2002VB-33/Gehl	6.31	-	+	-+
OA1207-1	AAFC-Ottawa	95Ab10854/07776-3-9-1-1a	6.76	-	+	-+
Sylva	AAFC-Ottawa	Shaw/Ogle	6.76	-	+	-+
OA1130-1	AAFC-Ottawa	06914-1-1/06919-5-1	1.34	+	+	++
OA1260-1	AAFC-Ottawa	OA1070-2/07834-3-1-1-1a-1	1.79	+	+	++
OA1262-1	AAFC-Ottawa	07593-1-1-2/07989-10-1-8	2.24	+	+	++
OA1253-1	AAFC-Ottawa	97Ab7979/07189-10-1-5	2.47	+	+	++
OA1180-5	AAFC-Ottawa	Ida /06909-2-26	2.69	+	+	++
OA1189-4	AAFC-Ottawa	OA1025-1/ 07800-2-8-1-1a	2.69	+	+	++
OA1256-1	AAFC-Ottawa	Souris/07189-10-8-6	2.92	+	+	++
04P07B-GT3D	AAFC-Ottawa	OT2022/Leggett	3.14	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
OA1174-3	AAFC-Ottawa	Ida/06909-3-94	3.14	+	+	++
OA1196-3	AAFC-Ottawa	07800-2-8-8-1a/OA1025-1	3.14	+	+	++
OA1251-1	AAFC-Ottawa	Souris/07774-3-9-1-1a-1	3.14	+	+	++
OA1263-2	AAFC-Ottawa	07878-3-6-5/07870-3-5-4	3.14	+	+	++
OA1180-4	AAFC-Ottawa	Ida/OA1019-6	3.37	+	+	++
OA1189-1	AAFC-Ottawa	OA1025-1/ 07800-2-8-1-1a	3.37	+	+	++
OA1232-2	AAFC-Ottawa	06973-5-10-15-2/ND960620	3.37	+	+	++
OA1176-1	AAFC-Ottawa	Ida/06909-1-19	3.43	+	+	++
OA1232-5	AAFC-Ottawa	06973-5-10-15-2 / ND960620	3.60	+	+	++
OA1268-3	AAFC-Ottawa	07868-1-71-1/07867-1-14-1	3.71	+	+	++
Brown 1409-164	AAFC-Ottawa	Brown 1409-164	3.82	+	+	++
OA1242-5	AAFC-Ottawa	06909-3-94/Irish	4.05	+	+	++
OA1226-1	AAFC-Ottawa	07836-2-9-3-1a-2/ND960620	4.27	+	+	++
OA1228-1	AAFC-Ottawa	W99A747/07836-2-5-2-1a-4	4.27	+	+	++
OA1266-1	AAFC-Ottawa	07869-2-9-4/07868-2-1-1	4.27	+	+	++
OA1197-1	AAFC-Ottawa	07800-2-8-5-1a/Sherwood	4.50	+	+	++
VAO-44	AAFC-Ottawa	2002VB-33/Gehl	4.50	+	+	++
OA1204-2	AAFC-Ottawa	MN-00222/07836-2-9-2-1a	4.73	+	+	++
OA1248-1	AAFC-Ottawa	07836-2-9-2-1a-1/07823-2-4-3-1a-1	4.73	+	+	++
Sherwood	AAFC-Ottawa	AC Aylmer/Goslin	4.73	+	+	++
OA1202-1	AAFC-Ottawa	07783-3-2-1-1a/SA99297	5.40	+	+	++
OA1234-1	AAFC-Ottawa	OT288/Triple Crown	5.40	+	+	++
LAO-1136-014	AAFC-Winnipeg	AC Morgan//Elvy/CDC Boyer	6.53	-	-	--
LAO-1134-022	AAFC-Winnipeg	AC Morgan/AC Mustang	6.98	-	-	--
LAO-1134-045	AAFC-Winnipeg	AC Morgan/AC Mustang	7.33	-	-	--
LAO-1135-046	AAFC-Winnipeg	AC Morgan/OT394	2.24	+	-	+-
00P01-A11A4	AAFC-Winnipeg	AC Ronald//ND931475/AC Assiniboia	3.14	+	-	+-

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
LAO-1136-024	AAFC-Winnipeg	AC Morgan//Elvy/CDC Boyer	3.43	+	-	+-
LAO-1135-015	AAFC-Winnipeg	AC Morgan/OT394	3.82	+	-	+-
LAO-1136-056	AAFC-Winnipeg	AC Morgan//Elvy/CDC Boyer	4.73	+	-	+-
00P28-AN01B1	AAFC-Winnipeg	JC1624#3/HiFi#3	3.37	-	+	-+
02G22-GM1E	AAFC-Winnipeg	AC Mustang/Baragan//01HN3021	5.86	-	+	-+
02G31-NU6D	AAFC-Winnipeg	W97254/Pinnacle//Assiniboia/Omskij	1.56	+	+	++
LAO-1104-028C1	AAFC-Winnipeg	ND9508252-75/CR245-Dw	1.56	+	+	++
01G08-AG3E	AAFC-Winnipeg	Omskij/Ronald	2.02	+	+	++
00P06-HD1D	AAFC-Winnipeg	AC Ronald/00UMOPN (USDA) 05	2.24	+	+	++
02G31-NL7A	AAFC-Winnipeg	W97254/Pinnacle//Assiniboia/Omskij	2.24	+	+	++
01G08-AA5C	AAFC-Winnipeg	Omskij/Ronald	2.69	+	+	++
01G04-CC3E	AAFC-Winnipeg	Assiniboia/Omskij	3.14	+	+	++
01G04-CC5B	AAFC-Winnipeg	Assiniboia/Omskij	3.14	+	+	++
04P07B-GY5E	AAFC-Winnipeg	OT2022/Leggett	3.14	+	+	++
04P08-DN5D	AAFC-Winnipeg	OT2022/01RAT23	3.14	+	+	++
LAO-882-036	AAFC-Winnipeg	CDC Dancer//Paul/AC Kaufmann	3.14	+	+	++
LAO-883-010	AAFC-Winnipeg	OT7000//Paul/AC Kaufmann	3.14	+	+	++
02G13-AR1C	AAFC-Winnipeg	Baragan/Pinnacle//Ronald	3.37	+	+	++
02G31-NN4B	AAFC-Winnipeg	W97254/Pinnacle//Assiniboia/Omskij	3.37	+	+	++
03P26A-BV3A	AAFC-Winnipeg	01RAT26/CDC Sol-Fi	3.37	+	+	++
04P07A-BK3A	AAFC-Winnipeg	Leggett/OT2022	3.37	+	+	++
04P07B-FR3B	AAFC-Winnipeg	OT2022/Leggett	3.37	+	+	++
04P07A-AA4D	AAFC-Winnipeg	Leggett/OT2022	3.60	+	+	++
04P07B-GN1C	AAFC-Winnipeg	OT2022/Leggett	3.60	+	+	++
04P07B-HC2A	AAFC-Winnipeg	OT2022/Leggett	3.60	+	+	++
Summit (00P08-BD1A)	AAFC-Winnipeg	AC Ronald/OT299	3.60	+	+	++
LAO-900-042	AAFC-Winnipeg	Paul/Kaufmann//Ronald	3.82	+	+	++
Leggett	AAFC-Winnipeg	OT294/Pc94	4.05	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
04P08-CZ3C	AAFC-Winnipeg	OT2022/01RAT23	4.05	+	+	++
02G13-BR1C	AAFC-Winnipeg	Baragan/Pinnacle//Ronald	4.50	+	+	++
04P08-DE5A	AAFC-Winnipeg	OT2022/01RAT23	4.50	+	+	++
02G22-LK1A	AAFC-Winnipeg	AC Mustang/Baragan//01HN3021	4.73	+	+	++
04P06A-CY5D	AAFC-Winnipeg	Leggett/01RAT23	4.73	+	+	++
04P08-EJ3B	AAFC-Winnipeg	OT2022/01RAT23	4.73	+	+	++
07P08-A04	AAFC-Winnipeg	ND9508252-75-5//03P22A-BM1/Pc97	4.73	+	+	++
LAO-1099-011A	AAFC-Winnipeg	CR245-Dw//Paul/AC Kaufmann	5.18	+	+	++
LAO-1099-011C	AAFC-Winnipeg	CR245-Dw//Paul/AC Kaufmann	5.40	+	+	++
OT3050	CDC-Saskatchewan	OT399/Freddy	5.86	-	-	--
OT3047	CDC-Saskatchewan	OT393/OA019-1	6.53	-	-	--
SA070270	CDC-Saskatchewan	Ajay/OT399+C156	6.53	-	-	--
OT3046	CDC-Saskatchewan	OT392/OT557	6.76	-	-	--
CDC Orrin	CDC-Saskatchewan	OT349/J775-1	7.66	-	-	--
Triactor	CDC-Saskatchewan	SW9619019/Stork	7.89	-	-	--
SA070906	CDC-Saskatchewan	CDC Sol-Fi/OT399	3.14	+	-	+-
SA070513	CDC-Saskatchewan	SA01714/99-511Cn161	3.45	+	-	+-
Derby	CDC-Saskatchewan	Calibre/Cascade	5.86	+	-	+-
OT399	CDC-Saskatchewan	OT368/Gem	4.05	-	+	-+
SA070576	CDC-Saskatchewan	OT559/CDC ProFi	4.73	-	+	-+
SA071760	CDC-Saskatchewan	CDC Weaver/Vista	5.18	-	+	-+
OT3037	CDC-Saskatchewan	AC AssS42/OT385	5.63	-	+	-+
SA070367	CDC-Saskatchewan	TAM2002/CDC ProFi	5.63	-	+	-+
SA070469	CDC-Saskatchewan	OT399/ABSP14-6	5.63	-	+	-+
OT3045	CDC-Saskatchewan	OT3002/OT394	5.83	-	+	-+
SA070089	CDC-Saskatchewan	CDC Dancer/CDC ProFi	5.86	-	+	-+
SA071369	CDC-Saskatchewan	SO02249/OT3013	5.86	-	+	-+

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
SA071405	CDC-Saskatchewan	CDC Weaver/Betania	5.86	-	+	-+
SA071616	CDC-Saskatchewan	SO03224/Betania	5.86	-	+	-+
CDC Minstrel	CDC-Saskatchewan	OT293/CDC Dancer	6.08	-	+	-+
OT3040	CDC-Saskatchewan	OT3002/OT394	6.08	-	+	-+
SA070655	CDC-Saskatchewan	OT399/CDC Dancer	6.08	-	+	-+
SA070712	CDC-Saskatchewan	Leggett/CDC Orrin	6.17	-	+	-+
SA070860	CDC-Saskatchewan	CDC Weaver/98Ab6491	6.31	-	+	-+
OT3039	CDC-Saskatchewan	OT3004/CDC-01-499-04-227	6.53	-	+	-+
SA070972	CDC-Saskatchewan	94-116Cn4/1/CDC Weaver	6.53	-	+	-+
SA070845	CDC-Saskatchewan	96Ab8597/CDC Weaver	6.72	-	+	-+
OT3028	CDC-Saskatchewan	AC Goslin/SA96400	7.21	-	+	-+
OT3036	CDC-Saskatchewan	OT396/HiFi	7.21	-	+	-+
OT3048	CDC-Saskatchewan	CDC Sol-Fi/HiFi	2.92	+	+	++
SA051172	CDC-Saskatchewan	CDC Sol-Fi/HiFi	2.92	+	+	++
SA070631	CDC-Saskatchewan	Jordan/SA01594	3.14	+	+	++
Bw4903	CDC-Saskatchewan	INTA, Barrow Station, Tres Arroyos, Argentina	3.37	+	+	++
SA061148	CDC-Saskatchewan	SA98741-11/CDC ProFi	3.60	+	+	++
SA070444	CDC-Saskatchewan	OT3007/SA99940	3.82	+	+	++
SA060830	CDC-Saskatchewan	OT399/OT2030	4.27	+	+	++
SA070592	CDC-Saskatchewan	Jordan/CDC Orrin	4.50	+	+	++
SA060605	CDC-Saskatchewan	X7535-14/OT3002	4.59	+	+	++
Bw103	CDC-Saskatchewan	INTA, Barrow Station, Tres Arroyos, Argentina	4.73	+	+	++
CDC Weaver	CDC-Saskatchewan	OT369/OT2007	4.95	+	+	++
SA070781	CDC-Saskatchewan	CDC Weaver/OT572	4.95	+	+	++
Bw1103	CDC-Saskatchewan	INTA, Barrow Station, Tres Arroyos, Argentina	5.18	+	+	++
OT7053	CDC-Saskatchewan	CR245-Dw//Paul/Kaufmann	5.63	+	+	++
Ave117.2	CDC-Saskatchewan	INIA, Carillanca Station, Temuco, Chile	5.86	+	+	++
Firth	Germany	CR 3/418/ Flamingsvita	6.31	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
Maida	North Dakota State University	ND873126/ASSINIBOIA	4.73	+	-	+-
ND060432	North Dakota State University	ND970216/ND000750	4.73	+	-	+-
ND060418	North Dakota State University	ND931318/ND931314	4.05	-	+	-+
ND051037	North Dakota State University	ND000306/ND980370	4.73	-	+	-+
ND051467	North Dakota State University	ND001336/HiFi-19	5.18	-	+	-+
ND070813	North Dakota State University	ND000824/ND010848	5.18	-	+	-+
ND060570	North Dakota State University	ND990148/ND991293	5.21	-	+	-+
ND061614	North Dakota State University	ND011604/ND981065	5.40	-	+	-+
ND060235	North Dakota State University	HiFi-93/ND981442	5.86	-	+	-+
ND060487	North Dakota State University	ND981845/HiFi-9	6.31	-	+	-+
Bond	North Dakota State University	Avena sterilis /Golden Rain	6.45	-	+	-+
ND060464	North Dakota State University	ND981065/ND011600	6.45	-	+	-+
ND070388	North Dakota State University	ND011054/ND030612	6.76	-	+	-+
ND030078	North Dakota State University	YOUNGS/ND980479	1.56	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
ND000861	North Dakota State University	ND950524/HiFi	2.02	+	+	++
ND060223	North Dakota State University	HiFi-9/ND980370	2.02	+	+	++
ND050506	North Dakota State University	ND981906/RONALD	2.24	+	+	++
ND061868	North Dakota State University	HiFi/IAN979-5-1-22	2.24	+	+	++
ND050578	North Dakota State University	ND990232/ND970216	2.47	+	+	++
ND060182	North Dakota State University	CDC Dancer/ND011608	2.47	+	+	++
ND050490	North Dakota State University	ND981903/ND990232	2.69	+	+	++
ND051306	North Dakota State University	ND000811/ND980671	2.69	+	+	++
ND051513	North Dakota State University	ND001444/HiFi-51	2.69	+	+	++
ND060249	North Dakota State University	HiFi-117/HiFi SR1	2.69	+	+	++
ND060342	North Dakota State University	Killdeer/ND000931	2.69	+	+	++
ND071063	North Dakota State University	ND011616/ND9508252-7-2 Drov/Jud BG	2.69	+	+	++
ND040196	North Dakota State University	ND931314/Souris	2.73	+	+	++
ND051312	North Dakota State University	ND000811/ND980671	3.14	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
ND060449	North Dakota State University	ND980671/Souris	3.14	+	+	++
ND060897	North Dakota State University	ND000824/ND010278	3.14	+	+	++
ND061975	North Dakota State University	Otana//M2609/Otana	3.14	+	+	++
ND060111	North Dakota State University	96-503 Cn32/ND990232	3.37	+	+	++
ND060507	North Dakota State University	ND981887/ND011288	3.37	+	+	++
ND061097	North Dakota State University	ND001018/Souris	3.37	+	+	++
ND081924	North Dakota State University	Triple Crown/ ND030220	3.37	+	+	++
ND060925	North Dakota State University	ND000865/ND991056	3.43	+	+	++
ND001397	North Dakota State University	BG27/BG25	3.60	+	+	++
ND051069	North Dakota State University	ND000490/ND000916	3.60	+	+	++
ND061590	North Dakota State University	ND011598/ND000223	3.60	+	+	++
ND071521	North Dakota State University	ND031348/ND951394	3.60	+	+	++
ND040492	North Dakota State University	ND970216/Souris	3.71	+	+	++
ND051236	North Dakota State University	ND000798/ND980370	3.82	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
ND060652	North Dakota State University	ND990355/ND000802	3.82	+	+	++
ND071694	North Dakota State University	ND873126/Morton	3.82	+	+	++
ND072258	North Dakota State University	ND990118/ND030288	3.82	+	+	++
ND050017	North Dakota State University	HiFi-11/ND001407	4.05	+	+	++
ND080724	North Dakota State University	Drover/MN00207	4.27	+	+	++
ND061813	North Dakota State University	Tri.Crown/ND010074	4.50	+	+	++
ND020290 (PO 808)	North Dakota State University	MN97112/ND971454	6.31	+	+	++
Kolbu	Norway	unknown/unknown	6.08	-	-	--
Olram	Norway	Sv692013/POL	6.08	-	-	--
Odal	Norway	Guldregn /Foredlad Dalahavre	6.53	-	-	--
Moholt	Norway	Voll /Palu	6.76	-	-	--
Nudist	Norway	Nudist	7.21	-	-	--
ARDENTE	Norway	KANTON/BELINDA	7.44	-	-	--
Kapp	Norway	Grakall /Tador	7.44	-	-	--
Grenader	Norway	Hedmarkshavre/Hedmarkshavre	7.66	-	-	--
Lena	Norway	Sang /Unisignum	7.66	-	-	--
Ringsaker	Norway	Ringsaker	7.89	-	-	--
LIPOPLUS	Norway	BELINDA/MATILDA	8.11	-	-	--
Biri	Norway	Grakall/Lena	8.11	-	-	--
Gere	Norway	Gere	8.11	-	-	--
Nes	Norway	Nes	8.11	-	-	--

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
CIRCLE	Norway	Sv 92158/Sv 923793	8.57	-	-	--
POL	Norway	Bambu/Norum 206	4.24	+	-	+-
Grane	Norway	unknown/unknown	6.31	+	-	+-
GN04399	Norway	GN04399	4.95	-	+	-+
Clintland 60	Purdue University	Clintland/unnamed_9214	6.08	-	-	--
Tyler	Purdue University	Clintland 60 /unnamed_4021	7.21	-	-	--
Clintland 64	Purdue University	unnamed_6497/unnamed_9222	7.40	-	-	--
Tippecanoe	Purdue University	Clntland 60/unnamed_4021	7.44	-	-	--
Clinton 59	Purdue University	unnamed_6925/Bond	7.66	-	-	--
Porter	Purdue University	P623A1-1-9-1/Stout	4.05	+	-	+-
Anthony	Purdue University	White Tartar/Victory	5.18	+	-	+-
Putnam 61	Purdue University	Putnam /unnamed_8806	5.18	+	-	+-
Clintland	Purdue University	Clinton 59 /unnamed_5257	6.76	+	-	+-
Allen	Purdue University	unnamed_4878 /unnamed_6500	7.21	+	-	+-
Clintford	Purdue University	Purdue 5124 A6-4/Milford	4.73	-	+	-+
Columbia	Purdue University	offtype plant selection from Fulghum/offtype plant selection from Fulghum	6.76	-	+	-+
059A1-2-2-4	Purdue University	973A38-4-4-5-4-1/P9741A41-4-6	1.79	+	+	++
001A1-24-2-4-1-3	Purdue University	P8669C2-6-4-16-7-6/94163A2-4-2-2-5	2.69	+	+	++
0216A1-1-55	Purdue University	OO1A1-4/9413RB1-2-32-1.-2	2.69	+	+	++
Excel	Purdue University	P9741A41-4-6/P9741A41-4-6	2.69	+	+	++
026A1-88-2-2	Purdue University	001A1/P971A10-4-6	2.92	+	+	++
0514A1-16-3	Purdue University	P9741A41-4-6-86/9741A41-4-6-88	2.92	+	+	++
8669C2-4-6-16-33	Purdue University	IL79-4924/P7869D1-5-3-4	2.92	+	+	++
0222A1-21-7-5-1-1	Purdue University	SA00207-2 / Robust	3.14	+	+	++
0513A1-18-5	Purdue University	9741A41-4-6-32/9741A41-4-6-88	3.14	+	+	++
0528A1-1	Purdue University	Woodburn-l/9741A41-4-6-32//9741A41-4-6-32	3.14	+	+	++
Woodburn	Purdue University	P8674B1/4/Classic/3/P9337A2/8674B1//WI X6141-2/909A23	3.14	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
021A1-78-1-5	Purdue University	5978A29-13-2-1,-2/P971A10-4-6	3.37	+	+	++
027A1-87-8-1	Purdue University	SA99297-1/P971A10-4-6	3.37	+	+	++
055A1-3-5-3	Purdue University	9741A41-4-6-88/ND95171E	3.37	+	+	++
Classic	Purdue University	Ogle /unnamned_9285	3.37	+	+	++
0219A1-84-4-4-4-4	Purdue University	OO1A1-5/Robust	3.60	+	+	++
0541A1-1	Purdue University	9741A41-4-6-32/98101A1-1-4-3//9876Gl-2-1-5	3.60	+	+	++
IN09201	Purdue University	IN09201	3.60	+	+	++
Robust	Purdue University	P973A38-9-3/P973A38-9-3	3.60	+	+	++
971A9-7-4-11	Purdue University	8674B1-2-4-2 /4/Classic/3/9337A2-2-4/8674B1-2-4-2//WI X6141-2/P909A23-1	3.82	+	+	++
9876C1-2-1-5-2-4-1	Purdue University	8669C2-6-4-16-1/ACC3028_A.STERILIS//95211RD1-4	4.05	+	+	++
053B1-95	Purdue University	9741A41-4-6-32/ND991293	4.50	+	+	++
AVENY	Sweden	BELINDA/SW 951865	7.21	-	-	--
SW VAASA	Sweden	Sv 841034/Sv 83626	7.21	-	-	--
BARRA	Sweden	SELMA KM	7.44	-	-	--
SW INGEBORG	Sweden	Sv 88359/SILVANO	7.44	-	-	--
CILLA	Sweden	A 83180/STIL	7.89	-	-	--
FREJA	Sweden	Vg 75842/DULA	7.89	-	-	--
GUNHILD	Sweden	Sv 97707/SANNA	8.11	-	-	--
SW KERSTIN	Sweden	PETRA/BALETT	8.57	-	-	--
Lennon	University of Aberystwyth	MF9018-11801/13174Cn3/6n	6.08	+	-	+-
Zuton	University of Aberystwyth	12799Cn4/6n /Bullion	6.26	+	-	+-
Racoon	University of Aberystwyth	Krypton/91-221Cn4	6.98	+	-	+-
Bullion	University of Aberystwyth	08974CnI/1/08944Cn1	6.76	-	+	-+

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
IL2250-15 (PI641978)	University of Illinois at Urbana-Champaign	Clintland 64*5/IL86-5698	6.98	-	-	--
IL2294-2 (PI641997)	University of Illinois at Urbana-Champaign	Clintland 64*5/IL86-6404	6.98	-	-	--
IL2294-1 (PI641996)	University of Illinois at Urbana-Champaign	Clintland 64*5/IL86-6404	6.31	+	-	+-
IL05-8515	University of Illinois at Urbana-Champaign	IL00-4858/IL2838-1	3.37	-	+	-+
IL05-1705	University of Illinois at Urbana-Champaign	OA1021-1/IL00-4858	4.05	-	+	-+
IL2250-3 (PI641976)	University of Illinois at Urbana-Champaign	Clintland 64*5/IL86-5698	4.95	-	+	-+
IL00-654	University of Illinois at Urbana-Champaign	Brawn/IL95-8346	1.56	+	+	++
IL2901 (PI641968)	University of Illinois at Urbana-Champaign	IL86-5698/IL86-1156//Ogle/IL86-6404	1.56	+	+	++
IL02-8658	University of Illinois at Urbana-Champaign	Tack/Spurs	1.79	+	+	++
IL05-10015	University of Illinois at Urbana-Champaign	IL00-8279/IL00-8622	1.79	+	+	++
IL05-9330	University of Illinois at Urbana-Champaign	IL00-8007/Spurs	1.79	+	+	++
IL86-6404 (PI539874)	University of Illinois at Urbana-Champaign	IL74-5234/IL75-5662//IL81-1454	1.79	+	+	++
IL05-10069	University of Illinois at Urbana-Champaign	IL00-8439/IL98-10145	2.02	+	+	++
IL05-3928	University of Illinois at Urbana-Champaign	IL98-2344/IL00-8279	2.02	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
IL3555 (PI641970)	University of Illinois at Urbana-Champaign	IL86-5698/IL86-1156//Ogle/IL86-6404	2.02	+	+	++
IL86-5698 (PI539875)	University of Illinois at Urbana-Champaign	IL74-5234/IL75-5662//IL81-1454	2.02	+	+	++
IL03-7936	University of Illinois at Urbana-Champaign	IL96-16806/IL98-18614	2.24	+	+	++
IL04-2727	University of Illinois at Urbana-Champaign	Sesqui/IL3538	2.24	+	+	++
IL04-3664	University of Illinois at Urbana-Champaign	Spurs/IL96-11037	2.24	+	+	++
IL04-4410	University of Illinois at Urbana-Champaign	IL96-3151/IL98-14767	2.24	+	+	++
IL05-3806	University of Illinois at Urbana-Champaign	IL98-2344/Buckskin	2.47	+	+	++
IL02-5630	University of Illinois at Urbana-Champaign	IL95-1555/IL97-18116	2.69	+	+	++
IL03-2658	University of Illinois at Urbana-Champaign	IL94-784/IL94-3961	2.69	+	+	++
IL2838 (PI641966)	University of Illinois at Urbana-Champaign	IL86-5698/IL86-1156//Ogle/IL86-6404	2.69	+	+	++
IL05-3337	University of Illinois at Urbana-Champaign	IL96-10351/OA1021-1	2.92	+	+	++
IL05-6223	University of Illinois at Urbana-Champaign	IL99-1515/OA1021-1	2.92	+	+	++
IL2815 (PI641965)	University of Illinois at Urbana-Champaign	IL86-5698/IL86-1156//Ogle/IL86-6404	2.92	+	+	++
IL02-8011	University of Illinois at Urbana-Champaign	Tack/IL94-3961	3.14	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
IL05-11942	University of Illinois at Urbana-Champaign	P971A9-7-4/IL00-8279	3.14	+	+	++
IL06-1161	University of Illinois at Urbana-Champaign	P971A10-4-6/IL00-7931	3.14	+	+	++
IL2250-18 (PI641979)	University of Illinois at Urbana-Champaign	Clintland 64*5/IL86-5698	3.14	+	+	++
IL75-5665	University of Illinois at Urbana-Champaign	Coker 227//Clintford/Portal	3.14	+	+	++
IL98-10145	University of Illinois at Urbana-Champaign	IL88-854/IL90-7147	3.14	+	+	++
IL02-10836	University of Illinois at Urbana-Champaign	IL97-19238/SD 97852	3.37	+	+	++
IL2294-3 (PI641998)	University of Illinois at Urbana-Champaign	Clintland 64*5/IL86-6404	3.37	+	+	++
IL2250-14 (PI641977)	University of Illinois at Urbana-Champaign	Clintland 64*5/IL86-5698	3.37	+	+	++
IL75-5743	University of Illinois at Urbana-Champaign	Coker 227//CI5068/CI 8074	3.37	+	+	++
IL05-9931	University of Illinois at Urbana-Champaign	IL00-8279/IL00-8622	3.82	+	+	++
IL05-9948	University of Illinois at Urbana-Champaign	IL00-8279/IL00-8622	3.82	+	+	++
IL06-3258	University of Illinois at Urbana-Champaign	Buckskin/Winona	3.82	+	+	++
IL06-3751	University of Illinois at Urbana-Champaign	IL00-205/Buckskin	4.05	+	+	++
IL86-4189	University of Illinois at Urbana-Champaign	Lang/IL75-5662//IL79-1776	4.05	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
IL06-5456	University of Illinois at Urbana-Champaign	IL00-4827/Buckskin	4.27	+	+	++
IL2294-8 (PI641999)	University of Illinois at Urbana-Champaign	Clintland 64*5/IL86-6404	4.73	+	+	++
IL05-1778	University of Illinois at Urbana-Champaign	OA1021-1/IL00-8279	4.95	+	+	++
MN08139	University of Minnesota	Morton/SA02880	5.63	+	-	+-
MN08238	University of Minnesota	ND010786//SA01511/Leonard	5.63	+	-	+-
Andrew	University of Minnesota	Bond/Rainbow	6.08	+	-	+-
MN08242	University of Minnesota	SO03244//MN862055/MN00206	4.27	-	+	-+
MN06239	University of Minnesota	Sask01T-602-05-06/MN01117	5.18	-	+	-+
MN08132	University of Minnesota	Sask01T-602-05-06/MN00206//MN03115	5.40	-	+	-+
MN08254	University of Minnesota	MN02155/3wSQC3-15	6.08	-	+	-+
MN08251	University of Minnesota	SA04305/Loyal	6.76	-	+	-+
MN07210	University of Minnesota	Sesqui*2/Bettong//MN02108	2.24	+	+	++
MN08222	University of Minnesota	ACAssiniboia/S42/OT394=SO04600	2.24	+	+	++
MN08268	University of Minnesota	MN03119/MN03205	2.24	+	+	++
MN08146	University of Minnesota	MN02218/3wSQC3-10	2.47	+	+	++
MN06108	University of Minnesota	SD986600/ND981502	2.69	+	+	++
MN08266	University of Minnesota	MN02231/MN03205	2.69	+	+	++
MN08106	University of Minnesota	Sesqui/WIX7571-1//Kame	2.92	+	+	++
MN08253	University of Minnesota	MN02218/3wSQC3-10	2.92	+	+	++
MN08260	University of Minnesota	MN01117/WIX7571-1	2.92	+	+	++
MN08160	University of Minnesota	IL99-8803/Sesqui	3.14	+	+	++
MN08217	University of Minnesota	OT2008/OT394=SA060201	3.14	+	+	++
MN08230	University of Minnesota	W98241/CDCORRIN//ND010426	3.14	+	+	++
MN06125	University of Minnesota	IL99-8803/MN01135	3.37	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
MN07203	University of Minnesota	Morton/IL95-1241	3.60	+	+	++
MN08130	University of Minnesota	IL99-912/MN00206//Loyal	3.60	+	+	++
MN06108	University of Minnesota	SD986600/ND981502	2.69	+	+	++
MN08266	University of Minnesota	MN02231/MN03205	2.69	+	+	++
MN08106	University of Minnesota	Sesqui/WIX7571-1//Kame	2.92	+	+	++
MN08253	University of Minnesota	MN02218/3wSQC3-10	2.92	+	+	++
MN08260	University of Minnesota	MN01117/WIX7571-1	2.92	+	+	++
MN08160	University of Minnesota	IL99-8803/Sesqui	3.14	+	+	++
MN08217	University of Minnesota	OT2008/OT394=SA060201	3.14	+	+	++
MN08230	University of Minnesota	W98241/CDCORRIN//ND010426	3.14	+	+	++
MN06125	University of Minnesota	IL99-8803/MN01135	3.37	+	+	++
MN07203	University of Minnesota	Morton/IL95-1241	3.60	+	+	++
MN08130	University of Minnesota	IL99-912/MN00206//Loyal	3.60	+	+	++
MN08212	University of Minnesota	OT2022/OT3006=SA061027	3.60	+	+	++
MN08150	University of Minnesota	MN02231/MN03205	3.82	+	+	++
MN08252	University of Minnesota	SA04913/MN02225	3.82	+	+	++
MN08262	University of Minnesota	ACAss/S42/OT394	4.05	+	+	++
MN06109	University of Minnesota	SD986600/ND981502	4.27	+	+	++
MN08123	University of Minnesota	Sesqui/WIX7571-1//Kame	4.27	+	+	++
MN08131	University of Minnesota	Morton//IL99-912/MN00206	4.27	+	+	++
MN08243	University of Minnesota	Sesqui/WIX7571-1//SA03668	4.50	+	+	++
MN08225	University of Minnesota	3wSCC2-15/MN03205	4.73	+	+	++
MN08211	University of Minnesota	ACAss/S42/SA01717=SA060726	4.95	+	+	++
MN08234	University of Minnesota	Morton//IL99-912/MN00206	4.95	+	+	++
MN08270	University of Minnesota	MN02231/SO04390	4.95	+	+	++
MN08129	University of Minnesota	Kame//IL99-912/MN00206	5.18	+	+	++
MN08155	University of Minnesota	UFRGS952521/X7464-4	5.18	+	+	++
MN08124	University of Minnesota	Sesqui/WIX7571-1//ND010426	5.40	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
MN08134	University of Minnesota	SA03668/MN02231	5.40	+	+	++
MN08137	University of Minnesota	SA03668/Morton	5.40	+	+	++
X9487-4	University of Wisconsin	X8402-2/X9191-5	4.27	+	-	+-
X9474-2	University of Wisconsin	X8177-1/X8341-3	5.18	-	+	-+
X9492-1	University of Wisconsin	X8416-2/X8175-3	6.08	-	+	-+
X9082-1	University of Wisconsin	X7822-3/X7066-5	2.24	+	+	++
X9386-1	University of Wisconsin	X8188-5/X8163-2	2.47	+	+	++
X9195-6	University of Wisconsin	X8208-5/ND961161	2.69	+	+	++
X9396-1	University of Wisconsin	X8191-8/X8179-1	2.69	+	+	++
X9396-4	University of Wisconsin	X8191-8/X8179-1	2.69	+	+	++
X9410-1	University of Wisconsin	X8342-1/X8179-1	2.69	+	+	++
X9487-1	University of Wisconsin	X8402-2/X9191-5	2.69	+	+	++
X8787-1	University of Wisconsin	X7535-9/X7395-4	2.92	+	+	++
X9200-4	University of Wisconsin	OA981-9/X8208-6	2.92	+	+	++
X9258-5	University of Wisconsin	X8175-2/DRUMLIN	2.92	+	+	++
X9290-2	University of Wisconsin	X8208-6/DRUMLIN	2.92	+	+	++
X8826-1	University of Wisconsin	IL92-7186/X6356-2	3.14	+	+	++
X9270-4	University of Wisconsin	X8191-2/X7571-1	3.14	+	+	++
X9375-1	University of Wisconsin	X8170-3/X8179-1	3.14	+	+	++
X8903-2	University of Wisconsin	X7728-3/X7509-1	3.25	+	+	++
X8791-1	University of Wisconsin	X7571-1/X6984-3	3.37	+	+	++
X9287-2	University of Wisconsin	X8208-6/X7464-4	3.37	+	+	++
X9414-1	University of Wisconsin	X8342-6/X8254-5	3.37	+	+	++
X9498-1	University of Wisconsin	X8470-6/X8163-2	3.37	+	+	++
X9221-6	University of Wisconsin	X8163-1/ND961161	3.60	+	+	++
X9221-8	University of Wisconsin	X8163-1/ND961161	3.60	+	+	++
X9253-1	University of Wisconsin	X8163-1/X8179-1	3.60	+	+	++
X9368-1	University of Wisconsin	X8163-2/X8342-6	3.60	+	+	++

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
X9392-1	University of Wisconsin	X8191-5/X8463-3	3.60	+	+	++
X9395-2	University of Wisconsin	X8191-8/X7994-2	3.60	+	+	++
X9509-3	University of Wisconsin	X8377-1/X8179-1	3.82	+	+	++
X9150-1	University of Wisconsin	X7994-2/Gem	4.05	+	+	++
X9195-2	University of Wisconsin	X8208-5/ND961161	4.05	+	+	++
X9410-2	University of Wisconsin	X8342-1/X8179-1	4.05	+	+	++
X9435-1	University of Wisconsin	X8463-3/X8179-1	4.05	+	+	++
X9500-6	University of Wisconsin	X8470-6/X8177-1	4.05	+	+	++
X9507-1	University of Wisconsin	X8210-2/X8377-1	4.05	+	+	++
X8995-4	University of Wisconsin	OA966-1/Vista	4.27	+	+	++
X9285-1	University of Wisconsin	X8313-2/X7766-1	4.27	+	+	++
X9500-2	University of Wisconsin	X8470-6/X8177-1	4.27	+	+	++
X9507-3	University of Wisconsin	X8210-2/X8377-1	4.27	+	+	++
X9509-4	University of Wisconsin	X8377-1/X8179-1	4.41	+	+	++
X9422-1	University of Wisconsin	X8177-1/X8254-5	4.50	+	+	++
X9439-1	University of Wisconsin	ND9508252/X8341-3	4.50	+	+	++
X9449-1	University of Wisconsin	OA021-2/X8341-3	4.50	+	+	++
X9503-1	University of Wisconsin	X7994-2/X8470-6	4.50	+	+	++
X9421-3	University of Wisconsin	X8177-1/X7535-14	4.73	+	+	++
X9384-2	University of Wisconsin	X8184-2/X8463-3	4.95	+	+	++
X9192-5	University of Wisconsin	X8191-2/ND961161	5.18	+	+	++
HA05AB38-22	USDA-ARS Aberdeen	97Ab8510/90Ab1322	6.31	-	-	--
99Ab10937	USDA-ARS Aberdeen	MN83207/74Ab2608	6.53	-	-	--
HA05AB29-39	USDA-ARS Aberdeen	AbSP 9-2/MN 94112	6.76	-	-	--
HA08-03X31-1	USDA-ARS Aberdeen	99Ab11899/TAMO7-3	6.76	-	-	--
99Ab10987	USDA-ARS Aberdeen	90Ab1322/Monida	6.98	-	-	--
99Ab11227	USDA-ARS Aberdeen	90Ab1322/SP3-2	6.98	-	-	--
HA05AB41-38	USDA-ARS Aberdeen	84Ab358/95Ab12729	6.98	-	-	--

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
98Ab7265	USDA-ARS Aberdeen	Corbit/88Ab3073	6.31	+	-	+-
HA08-03X49-1	USDA-ARS Aberdeen	CDC Dancer/Monida	5.18	-	+	-+
97Ab7767	USDA-ARS Aberdeen	IL81-2570/83Ab3250	5.40	-	+	-+
99Ab10971	USDA-ARS Aberdeen	90Ab1322/Monida	5.40	-	+	-+
02Ab6655	USDA-ARS Aberdeen	95Ab1284/Powell	5.63	-	+	-+
95Ab12770	USDA-ARS Aberdeen	86Ab1867/87Ab5597	5.63	-	+	-+
HA05AB42-20	USDA-ARS Aberdeen	P909A23-1/AbSP9-2	5.63	-	+	-+
99Ab11136	USDA-ARS Aberdeen	90Ab1322/91Ab2	5.86	-	+	-+
99Ab11787	USDA-ARS Aberdeen	91Ab2/SP9-7	5.86	-	+	-+
HA05AB36-33	USDA-ARS Aberdeen	96Ab8796/95Ab12743	5.86	-	+	-+
HA05AB16-31	USDA-ARS Aberdeen	Powell/96Ab9074	5.90	-	+	-+
00Ab7085	USDA-ARS Aberdeen	90Ab 1322/Ab Sp9-2	6.08	-	+	-+
HA05AB10-47	USDA-ARS Aberdeen	95Ab1284/Powell	6.08	-	+	-+
HA05AB10-51	USDA-ARS Aberdeen	95Ab1284/Powell	6.08	-	+	-+
HA05AB21-7	USDA-ARS Aberdeen	90Ab1322/ND930122	6.08	-	+	-+
HA05AB22-9	USDA-ARS Aberdeen	94Ab5543/ND 930122	6.08	-	+	-+
HA08-03X09-1	USDA-ARS Aberdeen	98Ab6346/TAMO6-4	6.08	-	+	-+
00Ab6963	USDA-ARS Aberdeen	82Ab 248/Ab Sp9-2	6.31	-	+	-+
00Ab7006	USDA-ARS Aberdeen	90Ab 1322/Ab Sp9-2	6.31	-	+	-+
00Ab6711	USDA-ARS Aberdeen	83Ab 3250/Ab Sp9-2	6.53	-	+	-+
02Ab5836	USDA-ARS Aberdeen	95AB1284/Powell	6.53	-	+	-+
02Ab6078	USDA-ARS Aberdeen	95Ab1284/Powell	6.53	-	+	-+
99Ab11098	USDA-ARS Aberdeen	90Ab1322/91Ab2	6.53	-	+	-+
HA05AB35-16	USDA-ARS Aberdeen	95Ab12743/90Ab1322	6.53	-	+	-+
HA05AB9-52	USDA-ARS Aberdeen	91Ab502/ND930122	6.53	-	+	-+
HA05AB9-32	USDA-ARS Aberdeen	91Ab502/ND930122	6.71	-	+	-+
00Ab6112	USDA-ARS Aberdeen	94Ab 6921/Ajay	6.76	-	+	-+
02HO-209	USDA-ARS Aberdeen	IA91001-2/Powell//Powell	6.76	-	+	-+

Table A.1 (continued)

Name	Cooperator	Pedigree	BYDV BLUP	SNP 3C	SNP 18D	Haplotype
HA05AB20-1	USDA-ARS Aberdeen	82Ab248/ND930122	6.76	-	+	-+
95Ab13050	USDA-ARS Aberdeen	84Ab825/86Ab5259	7.44	-	+	-+
HA05AB29-17	USDA-ARS Aberdeen	AbSP 9-2/MN 94112	2.92	+	+	++
HA05AB34-48	USDA-ARS Aberdeen	90Ab1620/95Ab12743	2.92	+	+	++
HA05AB53-40	USDA-ARS Aberdeen	94Ab5326/95Ab10854	3.37	+	+	++
02HO-139	USDA-ARS Aberdeen	Maverick/IA91324-2	4.27	+	+	++
97Ab7761	USDA-ARS Aberdeen	IL81-2570/83Ab3250	4.27	+	+	++
HA05AB38-39	USDA-ARS Aberdeen	97Ab8510/90Ab1322	4.27	+	+	++
HA08-03X45-1	USDA-ARS Aberdeen	Ajay/TAMO5-4	4.27	+	+	++
00Ab8118	USDA-ARS Aberdeen	90Ab 1322/Derby	4.68	+	+	++
99Ab11391	USDA-ARS Aberdeen	87Ab6153/SP3-9	6.31	+	+	++
Pl263412-1	USDA-ARS Minnesota	selection from Red Algerian No.31	6.53	-	-	--
CI1712-5	USDA-ARS Minnesota	Siberian selection (Ottawa)	6.76	-	-	--
PI266887-1	USDA-ARS Minnesota	San Jose (Portugal)	8.11	-	-	--
PI260616-1	USDA-ARS Minnesota	Amarela 5282 FB 32 (Brazil)	5.63	+	-	+-
CI8000-4	USDA-ARS Minnesota	AR 2-31-20 (Arkansas)	6.76	+	+	++

APPENDIX B

Table B.1 Genomic estimated breeding values (GEBVs) for 518 spring oat lines using BayesA model and 2138 single nucleotide polymorphisms.

Line	GEBV	Line	GEBV	Line	GEBV	Line	GEBV
SD031128	1.936	MN06203	2.478	MN08268	2.678	IL12-7556	2.830
SD110466	2.059	IL00-654	2.503	IL09-2756	2.680	IL11-6852	2.838
MN05155	2.138	IL12-10931	2.505	IL07-8720	2.681	MN08260	2.839
IL05-9330	2.146	Corral	2.515	IL2838 (PI641966)	2.693	IL02-8663	2.840
IL12-1919	2.176	MN11211	2.520	IL06-1161	2.702	IL12-7403	2.848
IL08-9442	2.184	IL09-3929	2.522	IL11-2739	2.703	OA1260-1	2.849
IL11-5534	2.284	IL08-2934	2.527	IL12-12805	2.712	IL02-8658	2.850
IL05-10015	2.310	ND030078	2.533	Tack	2.724	OA1180-5	2.862
IL09-5737	2.312	MN08146	2.545	OA1262-1	2.727	IL08-7031	2.873
IL11-6317	2.312	IL07-8721	2.557	00P06-HD1D	2.729	X9487-1	2.887
IL12-7416	2.312	SD111946	2.567	02G31-NU6D	2.749	X9200-4	2.891
MN06105	2.317	Saber	2.587	0528A1-1	2.759	SD111779	2.893
LAO-1104-028C1	2.325	IL86-5698 (PI539875)	2.606	IL12-1266	2.762	ND051513	2.902
IL12-7401	2.328	IL12-5568	2.609	SD111939	2.784	03P26A-BV3A	2.904
IL2901 (PI641968)	2.333	IL11-7655	2.613	01G08-AG3E	2.788	IL04-2727	2.907
IL3555 (PI641970)	2.337	IL12-10307	2.615	MN06101	2.793	MN08266	2.911
IL05-3928	2.343	IL11-7150	2.616	IL12-5538	2.796	ND071063	2.915
IL11-6469	2.362	IL12-7530	2.620	MN11140	2.796	MN07210	2.924
OA1130-1	2.374	IL08-6344	2.629	ND060223	2.802	X9258-5	2.954
IL05-10069	2.375	IL2815 (PI641965)	2.641	Blaze	2.803	ND051306	2.957
IL09-5745	2.391	IL12-9020	2.646	IL05-3806	2.819	IL03-7936	2.963
IL10-5863	2.395	IL12-6448	2.649	01G08-AA5C	2.826	IL11-5748	2.968
059A1-2-2-4	2.404	IL04-4410	2.653	0513A1-18-5	2.827	X8903-2	2.972
IL04-3664	2.439	Buckskin	2.655	IL12-12427	2.827	ND060182	2.979
IL08-10563	2.457	IL06-5433	2.673	ND050490	2.828	ND000861	2.980

Table B.1 (continued)

Line	GEBV	Line	GEBV	Line	GEBV	Line	GEBV
ND050506	2.989	X9082-1	3.104	ND040196	3.217	0216A1-1-55	3.329
Excel	2.995	ND060925	3.105	IL05-3337	3.221	MN08230	3.334
IL12-8561	2.997	021A1-78-1-5	3.110	04P07B-GT3D	3.223	IL06-8153	3.341
02G31-NL7A	3.002	LAO-883-010	3.117	04P08-DN5D	3.228	X9290-2	3.342
BT1021-1-1	3.002	OA1256-1	3.125	MN08106	3.231	MN08217	3.346
ND060897	3.005	MN08222	3.127	ND060342	3.232	ND060111	3.347
ND050578	3.018	IL12-7074	3.137	IL09-5508	3.237	027A1-87-8-1	3.351
MN06120	3.020	IL03-2658	3.141	8669C2-4-6-16-33	3.238	OA1232-2	3.352
SA051172	3.025	IL75-1056 (Hazel)	3.144	OT3048	3.251	X9287-2	3.356
026A1-88-2-2	3.026	IL12-5536	3.152	X9395-2	3.251	OA1189-1	3.369
IL08-9452	3.032	IL05-6223	3.153	ND060449	3.257	OA1196-3	3.378
SD111753	3.034	IL09-2968	3.159	IL05-11942	3.267	X9195-6	3.388
IL75-5665	3.048	IL12-3846	3.162	Summit (00P08-BD1A)	3.268	MN06125	3.390
0514A1-16-3	3.049	X9410-1	3.170	ND061590	3.276	IL05-9948	3.392
IL86-6404 (PI539874)	3.052	OA1251-1	3.170	04P07B-HC2A	3.286	01G04-CC5B	3.392
ND061868	3.060	OA1189-4	3.187	971A9-7-4-11	3.287	P0714A1-29-2	3.392
IL02-10836	3.065	IL12-5532	3.187	055A1-3-5-3	3.293	01G04-CC3E	3.393
ND060249	3.069	SA070631	3.194	IL09-1126	3.300	X8791-1	3.409
IL10-9867	3.073	ND061097	3.194	MN06108	3.311	LAO-882-036	3.415
IL10-9872	3.075	X9410-2	3.197	IL02-5630	3.311	0222A1-21-7-5-1-1	3.417
X9396-1	3.078	02G13-AR1C	3.198	Woodburn	3.318	IL12-1632	3.421
IL12-5508	3.084	IL05-9931	3.200	IL08-2010	3.322	MN07203	3.426
ND060507	3.092	ND051312	3.204	IL98-10145	3.327	OA1263-2	3.426
X9396-4	3.095	MN08253	3.209	OA1253-1	3.328	X8826-1	3.431
02G31-NN4B	3.103	X9386-1	3.215	MN08160	3.328	ND051236	3.434

Table B.1 (continued)

Line	GEBV	Line	GEBV	Line	GEBV	Line	GEBV
Spurs	3.435	IL11-7728	3.546	OA1242-5	3.683	IL06-3751	3.859
X9368-1	3.446	MN08262	3.557	04P08-CZ3C	3.686	SA070513	3.860
IL02-8011	3.447	OA1174-3	3.561	X9195-2	3.693	X9507-3	3.866
X8787-1	3.450	IL10-2656	3.561	X9414-1	3.697	OA1226-1	3.881
Robust	3.452	OA1176-1	3.564	0541A1-1	3.730	OA1266-1	3.899
0219A1-84-4-4-4-4	3.453	BT1020-1-1	3.564	Leggett	3.739	MN08225	3.906
MN08150	3.461	IL250-14 (PI641977)	3.580	OA1232-5	3.753	LAO-1135-046	3.908
X9270-4	3.463	ND060652	3.583	X9285-1	3.758	LAO-1135-015	3.930
X9221-6	3.471	IL11-2353	3.587	MN08252	3.767	X9500-6	3.939
04P07B-GY5E	3.485	X9392-1	3.592	9876C1-2-1-5-2-4-1	3.787	HA05AB53-40	3.944
ND001397	3.495	Bw4903	3.600	IL06-5456	3.790	97Ab7761	3.947
IL11-5728	3.498	ND040492	3.601	X8995-4	3.792	MN08123	3.968
IL86-4189	3.502	IL75-5743	3.607	001A1-24-2-4-1-3	3.797	Classic	3.990
X9221-8	3.505	OA1197-1	3.610	04P07A-AA4D	3.805	SA060830	3.997
ND072258	3.507	ND051069	3.611	04P07B-GN1C	3.810	02HO-139	4.005
X9498-1	3.509	IL12-7370	3.612	MN08130	3.816	SA061148	4.018
ND050017	3.511	ND061975	3.624	07P08-A04	3.825	IL09-6937	4.035
MN06109	3.514	OA1180-4	3.636	X9435-1	3.830	04P06A-CY5D	4.043
X9253-1	3.524	ND071521	3.643	ND081924	3.830	X9192-5	4.043
04P07A-BK3A	3.526	MN08212	3.653	IL2294-3 (PI641998)	3.836	IL06-3258	4.046
HA05AB29-17	3.526	X9439-1	3.657	IL11-6459	3.840	OA1228-1	4.052
IN09201	3.527	HA05AB34-48	3.661	00P01-A11A4	3.841	X9509-3	4.055
OA1268-3	3.528	X9150-1	3.664	ND071694	3.845	053B1-95	4.072
IL2250-18 (PI641979)	3.538	X9375-1	3.681	SA070444	3.851	X9449-1	4.078
04P07B-FR3B	3.544	X9503-1	3.682	LAO-900-042	3.854	OA1248-1	4.090

Table B.1 (continued)

Line	GEBV	Line	GEBV	Line	GEBV	Line	GEBV
04P08-DE5A	4.102	IL12-7166	4.370	MN08242	4.812	Lennon	5.259
Brown 1409-164	4.119	OA1204-2	4.385	Bw1103	4.827	ND061614	5.274
IL11-6458	4.140	MN08211	4.402	ND051037	4.827	X9474-2	5.298
ND080724	4.143	ND061813	4.440	MN08137	4.828	ND070813	5.321
SA070592	4.171	CDC Weaver	4.469	Bw103	4.840	PI260616-1	5.368
HA05AB38-39	4.174	X9500-2	4.487	LAO-1099-011A	4.843	Anthony	5.432
04P08-EJ3B	4.178	MN08124	4.491	ND020290 (PO 808)	4.847	MN06239	5.434
X9384-2	4.185	SA070781	4.499	OT7053	4.848	CDC Minstrel	5.461
SA070906	4.189	OA1272-1	4.503	LAO-1099-011C	4.862	Grane	5.466
IL2294-8 (PI641999)	4.207	HA08-03X45-1	4.503	POL	4.868	ND060487	5.466
X9507-1	4.227	IL12-7430	4.505	IL08-9435	4.947	SA070089	5.474
IL05-1778	4.240	Porter	4.516	99Ab11391	4.949	HA05AB21-7	5.475
IL05-8515	4.258	LAO-1136-056	4.528	MN08129	4.956	MN08132	5.488
02G22-LK1A	4.263	X9421-3	4.557	SD111972	5.007	SA070469	5.500
VAO-44	4.267	IL09-6934	4.597	MN08238	5.014	OA1250-2	5.517
Sherwood	4.269	X9487-4	4.624	OA1202-1	5.034	Clintland	5.528
00Ab8118	4.287	MN08234	4.631	ND060418	5.034	ND060570	5.532
00P28-AN01B1	4.310	MN08134	4.642	Ave117.2	5.042	OA1250-1	5.535
X9422-1	4.315	IL05-1705	4.653	OA1234-1	5.054	SA070367	5.541
LAO-1136-024	4.320	SA060605	4.658	SA071760	5.078	SA070655	5.571
MN08243	4.322	IL09-2838	4.680	IL250-3 (PI641976)	5.095	ND060464	5.587
MN08270	4.325	MN08155	4.721	ND051467	5.107	95Ab12770	5.594
02G13-BR1C	4.326	Maida	4.770	SA070576	5.115	MN08254	5.617
X9509-4	4.338	Firth	4.777	Putnam 61	5.224	97Ab7767	5.619
ND060432	4.341	MN08131	4.781	OT399	5.240	Nice	5.626

Table B.1 (continued)

Line	GEBV	Line	GEBV	Line	GEBV	Line	GEBV
OT3045	5.627	IL12-12821	5.992	02Ab6078	6.293	Columbia	6.681
Clintford	5.635	99Ab10971	6.010	Kolbu	6.312	OA1207-1	6.702
HA08-03X49-1	5.639	HA08-03X09-1	6.019	MN08251	6.345	00Ab6963	6.712
IL2294-1 (PI641996)	5.639	Zuton	6.020	OT3046	6.358	CI1712-5	6.714
Derby	5.671	HA05AB35-16	6.044	02Ab5836	6.358	02HO-209	6.736
98Ab7265	5.680	SA070860	6.048	99Ab11098	6.369	IL2250-15 (PI641978)	6.795
HA05AB16-31	5.702	HA05AB22-9	6.051	95Ab13050	6.418	OT3047	6.835
CI8000-4	5.713	VAO-51	6.059	OT3028	6.421	Clintland 64	6.911
X9492-1	5.742	OA1226-4	6.061	SA070270	6.433	OA1058-4	6.918
SA071616	5.743	02Ab6655	6.066	SA070972	6.441	Grenader	6.923
MN08139	5.744	ND070388	6.097	Pl263412-1	6.470	99Ab10987	6.957
SA071369	5.748	Bullion	6.112	SA070845	6.490	HA08-03X31-1	7.035
HA05AB42-20	5.798	HA05AB9-32	6.117	00Ab7006	6.508	PI266887-1	7.071
ND060235	5.800	00Ab7085	6.137	HA05AB29-39	6.521	Kapp	7.122
Navaro	5.835	Allen	6.158	00Ab6711	6.526	Tippecanoe	7.131
Adrew	5.840	GN04399	6.165	HA05AB38-22	6.537	99Ab11227	7.143
OT3040	5.881	OT3037	6.182	LAO-1136-014	6.546	SW VAASA	7.153
02G22-GM1E	5.893	Racoon	6.207	HA05AB10-47	6.559	Nudist	7.172
VAO -58	5.899	HA05AB36-33	6.208	Clintland 60	6.567	Moholt	7.172
HA05AB10-51	5.912	HA05AB9-52	6.241	OT3036	6.573	CDC Orrin	7.184
SA071405	5.919	Bond	6.242	Sylva	6.630	LAO-1134-022	7.203
SA070712	5.948	99Ab11136	6.262	LAO-1134-045	6.636	HA05AB41-38	7.225
HA05AB20-1	5.949	OT3050	6.263	99Ab10937	6.646	BARRA	7.260
IL12-7208	5.961	OT3039	6.272	Olram	6.654	CILLA	7.277
Chinese 4	5.980	99Ab11787	6.284	00Ab6112	6.670	IL2294-2 (PI641997)	7.299

Table B.1 (continued)

Line	GEBV
AVENY	7.330
Clinton 59	7.354
Tyler	7.470
ARDENTE	7.499
Vao-48	7.516
Cantal	7.529
Lena	7.582
Odal	7.583
SW INGEBORG	7.616
Biri	7.673
FREJA	7.757
GUNHILD	7.830
Gere	7.831
Nes	7.833
LIPOPLUS	7.899
CIRCLE	7.912
Ringsaker	7.938
Triactor	8.070
SW KERSTIN	8.077