AN ABSTRACT OF THE DISSERTATION OF

Yada Chutimanitsakun for the degree of Doctor of Philosophy in Crop Science presented on December 7, 2011.

Title: From the Oregon Wolfe Barley to Fall-sown Food Barley: Markers, Maps, Marker-assisted Selection and Quantitative Trait Loci

Abstract approved:

Patrick M. Hayes

Understanding complex traits is a fundamental challenge in plant genetics and a pre-requisite for molecular breeding. Tools for trait dissection are markers, maps, and quantitative trait locus (QTL) analysis. Marker-assisted selection (MAS) is an application that integrates these tools. In this thesis research, a new sequence-based marker was evaluated, maps were constructed and used, and QTLs were detected using two types of populations. Marker-assisted selection was used to develop a novel class of barley. Restriction-site Associated DNA (RAD), a sequence based-marker technology, allows for simultaneous high-density single nucleotide polymorphism (SNP) discovery and genotyping. We assessed the value of RAD markers for linkage map construction using the Oregon Wolfe Barley (OWB) mapping population. We

compared a RAD-based map to a map generated using Illumina GoldenGate Assay (EST-based SNPs). The RAD markers generated a high quality map with complete genome coverage. We then used the RAD map to locate QTL for agronomic fitness traits. A paper describing this research was published (Chutimanitsakun et al., 2011). Marker-assisted selection was used to rapidly develop fall-sown barley germplasm for human food uses. The target traits were high grain β-glucan, vernalization sensitivity (VS) and low temperature tolerance (LTT). The target loci were *WX* and *VRN-H2*. Marker-assisted selection was effective in fixing target alleles at both loci and waxy starch led to increase in grain β-glucan. Unexpected segregation at *VRN-H1* and *VRN-H3*, revealed by genome-wide association mapping (GW-AM), led to unanticipated phenotypic variation in VS and LTT. We found that GW-AM is an efficient and powerful method for identifying the genome coordinates of genes determining target traits. Precise information is obtained with perfect markers; additional research may be needed when multiple alleles are segregating at target loci and significant associations are with markers in linkage disequilibrium (LD) with the target loci. A paper describing this research will be submitted for publication.

From The Oregon Wolfe Barley to Fall-sown Food Barley: Markers, Maps, Marker-assisted Selection and Quantitative Trait Loci

by
Yada Chutimanitsakun

A DISSERTATION

Submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented December 7, 2011
Commencement June 2012

Doctor of Philosophy dissertation of Yada Chutimanitsakun presented on

December 7, 2011

APPROVED:

Major Professor, representing Crop Science

Head of the Department of Crop and Soil Science

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Yada Chutimanitsakun, Author

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my major professor, Dr. Patrick M. Hayes for his guidance, support, encouragement, and advice throughout the course of my study. I also would like to thank him for the financial support throughout my program.

I further express my appreciation to the other members of my committee who have given of their time and assistance: Dr. Andrew Ross, Dr. Jennifer Kling, Dr. Shaun Townsend and Dr. Ramesh Sagili.

My sincere appreciation is extended to the members of the Oregon State University barley breeding program: Dr. Alfonso Cuesta-Marcos for offering and sharing his knowledge and the expertise to the project. Ann Corey and Tanya Filichkin for their helpful, generosity and advice throughout the research. Scott Fisk, and Natalie Graham for all their help and friendship.

I further extend my gratitude to the members of Floragenex for providing sequenced RAD marker. Special thanks to Caryn Ong who assisted me in β-glucan analysis.

Finally, I would like to give thanks to my family who has been always supportive and believed in me.

CONTRIBUTION OF AUTHORS

Dr. Patrick M. Hayes, initiated, advised, and supervised all aspects of the projects and drafted portions of the thesis. Dr. Alfonso Cuesta-Marcos advised in QTL analysis, performed statistical analysis and drafted portions of the thesis. Rick W. Nipper and Eric A. Johnson conducted RAD sequencing and drafted portions of the first chapter of the thesis. Shiaoman Chao conducted the Illumina GoldenGate genotyping. Ann Corey and Tanya Filichkin assisted with greenhouse, field and laboratory works. Juan Ignacio Rey conducted marker-assisted selection. Dr. Andrew Ross and Caryn Ong helped in quality analysis. Luis Cistué helped in phenotyping. Mathias Kolding managed the field trials at Hermiston, Oregon.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

Figure Page

LIST OF TABLES

FROM THE OREGON WOLFE BARLEY TO FALL-SOWN FOOD BARLEY: MARKERS, MAPS, MARKER-ASSISTED SELECTION AND QUANTITATIVE TRAIT LOCI

GENERAL INTRODUCTION

History and current US importance

Barley is an ancient cereal grain, domesticated in the Fertile Crescent of the Middle East, with importance dating back to 8000 B.C. (Nevo, 1992). Barley is widely grown today and is among the top five crops produced worldwide. According to the Foreign Ag Service (FAS 2009), the US is the seventh largest producer of barley in the world with the majority of its production in the Northern Plains and Pacific Northwest. In 2011, the US barley crop totaled 155.05 million bushels, representing a decline of 14% from 2010. However, the 2011 average price was $5.30 per bushel compared to $3.90 per bushel in 2010. As a result, the US crop was valued at $821.5 million (USDA 2011, available at http://www.nass.usda.gov/Data_and_Statistics/ [verified 16 Nov. 2011]).

Uses

Barley is used in various cultures as a grain for animal feed, brewing, and food. Most commonly, it is used for malting and feed. Only a minor amount of barley is actually used in the production of foods for human consumption. However, in recent years there has been increasing interest in barley as a human food. A health claim that associates consumption of β-glucan from whole grain barley and barley-related products with a reduced risk of coronary heart disease was finalized in 2006 (FDA

2006). The Panel on Dietetic Products, Nutrition and Allergies (NDA) of the European Food Safety Authority (EFSA) supports health benefit endorsement of food containing barley (EFSA 2009). Barley offers an excellent source of soluble fiber β-glucan (5-11%), superior to any other cereal grain (1% in wheat, 3-7% in oat) (Brennan and Cleary, 2005). β-glucan provides health benefits for the human body, especially in terms of reduced risk of heart disease and type II diabetes (Pins and Kaur, 2006). To move the food barley industry forward, it is important to develop high β-glucan varieties for barley growers, processors, and consumers.

Focus on food barley

β-glucan is a non-starch polysaccharide that contains β-D-glucose-like cellulose with one 1β→3 linkage for every three or four 1β→4 linkages. Mixed-linkage (1,3;1,4)-β-D-glucan resides in the cell walls of starchy endosperm and aleurone cells. The genetic basis of β-glucan is remarkably complex and several quantitative trait loci (QTL) have been detected (Emebiri, 2009; Han et al., 1995; Li et al., 2008; Molina-Cano et al., 2007). Two genes, cellulose synthase-like H (*CSLH)* and cellulose synthase-like F (*CSLF*), implicated in mediating (1,3;1,4)-β-D-glucan synthesis, were cloned (Burton et al., 2006; Doblin et al., 2009). Current studies indicate *HvCslF6* is the key determinant controlling the biosynthesis of β-glucan (Taketa et al., 2011). Over-expression of the *HvCslF6* gene in transgenic barley led to increases of more than 80% in β-glucan (Burton et al., 2011). High β-glucan is also reported in cultivars with 100% amylopectin (waxy) starch (Xue et al., 1991). The

positive pleiotropic effects of waxy starch on β-glucan (Xue et al., 1997) offer alternative approaches to breeding for higher β-glucan percentage.

Barley starch granules can be classified as normal, waxy, and high amylose according to the amylose: amylopection ratio. Waxy starch has low or no amylose (0-10%), normal (non-waxy) starch contains approximately 25% amylose and high amylose starch is more than 35% amylose (Matsuki et al., 2008). The waxy trait is controlled by the recessive (*wx*) allele at the granule-bound starch synthase I (*GBSSI*) locus. A deletion in the promoter region of this gene disrupts amylose synthesis (Domon et al., 2002; Patron et al., 2002), causing starch to be waxy and providing a perfect marker for indirect selection for increased β-glucan. Marker-assisted selection (MAS) has the potential to accelerate conventional plant breeding by allowing for selection on the basis of genotype rather than phenotype. In many cases, screening for the target trait(s) can be time-consuming and expensive, particularly for complex traits with low heritability. For example, an enzymatic assay is commonly used for measurement of β-glucan. This method gives the most reliable results as compared to other methods, such as near-infrared reflectance spectroscopy. However, the enzymatic assay is relatively expensive per sample. Recently, Hu and Burton (2008) improved the original protocol to reduce the cost and gain high throughput while maintaining the accuracy of the result. In the long run, selection should be more effective and efficient by using a perfect marker for the *WX* locus.

Fall-sown barley: advantages and genetics

Fall-sown barley is gaining interest to growers in the Pacific Northwest of the US due to its high yield and water use efficiency. Winterhardiness is an important limitation to fall-sown varieties. Winterhardiness is a complex trait controlled by interacting quantitative traits. Low temperature tolerance (LTT) is the main component of winterhardiness, and two other traits are vernalization sensitivity (VS) and photoperiod sensitivity (PPDS) (Hayes et al., 1993). Two major QTL for LTT are Frost Resistance *FR-H1* and *FR-H2*, located on chromosome 5H and ~30 cM apart (Francia et al., 2004; Skinner et al., 2005). *FR-H1* co-segregates with, and is assumed to be a pleiotropic effect of, *VRN-H1* (Dhillon et al., 2010). *FR-H2* is coincident with a cluster of C-repeat binding factor (*CBF)* genes (Skinner et al., 2006). Three loci, *VRN-H1* (5H), *VRN-H2* (4H) and *VRN-H3* (7H), determine VS (the requirement for prolonged exposure to cold to promote flowering). As reviewed by Trevaskis (2010), *VRN-H1* is a promoter of flowering that is activated by cold treatment (vernalization). *VRN-H1* activates expression of Flowering Locus T (*FT*). Flowering Locus T is a candidate for *VRN-H3* that is induced by long days to promote flowering. *VRN-H2* is a repressor of flowering under long days and is down-regulated by vernalization. In the case of VS, without vernalization the expression of *VRN-H1* is low and *VRN-H2* inhibits long day induction of *FT* (*VRN-H3*). With vernalization *VRN-H1* is active and down-regulates *VRN-H2*, allowing expression of *VRN-H3* (flowering). The functional genetic bases of *VRN-H1* and *VRN-H2* are known. Deletions in the first intron of *VRN-H1* cause spring (dominant) allele. The dominant allele at *VRN-H1* permits the plant to flower without vernalization. The size of the deletion in *VRN-H1* corresponds

to different flowering times (Szucs et al., 2007). The complete deletion or point mutation of *VRN-H2* is associated with the recessive (spring) allele, causing plants to flower under long days without vernalization. Little is known about *VRN-H3*. Dominant (spring) alleles at *VRN-H3* lead to very early flowering (Yan et al., 2006) and are found mostly in exotic germplasm (Takahashi and Yasuda, 1971). *PPD-H1* (2H) and *PPD-H2* (1H) are two major genes controlling photoperiod responses. *PPD-H1* promotes flowering under long days while *PPD-H2* prevents flowering under short days. The recessive (*ppd-H1*) allele caused by a mutation in the CCT domain shows delayed flowering under long days, as compared to the dominant *Ppd-H1* allele (Turner et al., 2005).

Vernalization and photoperiod interact with LTT in terms of growth habit classification. Low temperature tolerance, VS and PPDS describe "winter" or fall-sown barley. Vernalization sensitivity delays flowering until cold acclimation is attained (Fowler et al., 1996). Photoperiod sensitivity shows a similar effect in delaying the vegetative to reproductive transition. Photoperiod sensitivity plants will not flower until the day length reaches a critical threshold. However, no photoreceptor candidate genes were mapped to the LTT QTL (Szucs et al., 2006). "Spring" varieties are not VS and have no LTT. Photoperiod sensitivity is not significant because plants are grown under long days. "Facultative" barley is not VS but it is LTT. Some facultative varieties such as Dicktoo show high LTT, comparable to winter types (von Zitzewitz et al., 2011). There is evidence for (Fowler et al., 1996; Limin et al., 2007) and against (von Zitzewitz et al., 2011) the hypothesis that VS is a pre-requisite for

maximum LTT. Likewise, it has not been established that *FR-H1* is a pleiotropic effect of *VRN-H1*. The functional polymorphisms at *VRN-H1* and *VRN-H2* are clear and therefore provide perfect markers for VS, which may lead to LTT.

The Oregon Wolfe Barley population: a resource for genetics and genomics

Barley is not only an economically important crop, but also a model species within *Triticeae,* contributing to the broad utilization in scientific research. Barley is used for genetic and molecular studies as it is a diploid species, with a low chromosome number (2n=14). It is a predominantly self-pollinated, annual plant that can be grown successfully across more varying climatic conditions than any of the other cereals. Wolfe (1972) provided a valuable tool for barley genetics in the form of the dominant and recessive marker stocks. Subsequently, at Oregon State University a doubled haploid (DH) population was derived from the F_1 of the cross of the Wolfe recessive and dominant marker stocks using the *Hordeum bulbosum* method. Because of its high polymorphism frequency and the unique characteristics of each line, the population is an excellent resource for genetics research and instruction (Cistue et al., 2011; Costa et al., 2001). Recently, Cistue et al. (2011) developed new DH lines (*n*=93) from the same cross using the anther culture method. As its population size increases, the Oregon Wolfe Barley (OWB) population will become even more useful for genetic studies.

A detailed understanding of genetic diversity, functionality, and mechanisms of genes is the basis for plant breeding and research. Molecular markers represent a useful tool for the analysis of genetic variation in crop plants. They allow plant

geneticists to construct genetic maps and use them for detecting, mapping, and estimating the effects of QTL. Over the past two decades, advances in technology have led to the introduction of various types of markers, as reviewed by Semagn et al. (2010). Each type of marker has its advantages and shortcomings. The selection of suitable molecular markers will vary, depending on the species and resources available. Among available markers, single nucleotide polymorphism (SNP), a single base change in a DNA sequence, is one of the most common types of genetic variation. Because of their abundance, SNP-based markers have become preferred markers for association and linkage studies (Gupta et al., 2001; Rafalski, 2002).

Several techniques for SNP identification are currently available; most rely on preexisting sequence data (Ganal et al., 2009). In many crops a large number of expressed sequence tags (EST) have been generated and are available in public databases. These public resources, direct sequencing of PCR products targeted at particular candidate genes, and sequence alignments from reference genotypes are straightforward methods for identifying SNPs. However, the lack of DNA sequencing information has limited SNP discovery in non-model species. The human genome sequencing project has given rise to a number of high-throughput, next-generation sequencing (NGS) technologies. The emergence of these technologies has provided valuable information in SNP discovery through comparison of sequence data and opportunities to generate a large number of markers.

Several high-density platforms, such as the Illumina GoldenGate Genotyping Assay, that allow interrogation of 96 to 1536 SNPs in a single reaction are now

available (Gupta et al., 2008). This high-throughput SNP genotyping is being used for genetic analysis in a number of species including barley (Close et al., 2009), wheat (Akhunov et al., 2009), and maize (Yan et al., 2010). In barley, a three custom oligonucleotide pool assay (OPA) containing 1536 SNPs per assay has been developed through a collection of sequence resources (e.g. ESTs and sequenced PCR amplicons). Single nucleotide polymorphisms from three OPAs were selected to generate two production barley oligonucleotide pool assays (BOPA1 and BOPA2) (Close et al., 2009). The available BOPA-SNP-based markers were used by the barley genetics community to facilitate association analysis and gene identification (Roy et al., 2010; von Zitzewitz et al., 2011).

The costs of genotyping and obtaining prior genetic information are limiting factors for many plant breeding programs. Thus, there has been steady improvement in genotyping systems in order to create a cost-effective technology with high data throughput. The recent SNP-based marker called “Restriction-site Associated DNA” (RAD) has been developed for high-density SNP discovery and genotyping by sequencing. Initially, the RAD technique was implemented using microarray for individual and bulk segregant analysis (Miller et al., 2007a; Miller et al., 2007b). The next generation of RAD was modified for use with massive parallel sequencing on the Illumina platform and focused on genome complexity reduction in a form of restriction-fragment-size-based sampling (Baird et al., 2008). The RAD approach produces a large amount of short DNA sequence reads adjacent to restriction sites across the genomes. The target genome is digested with a restriction enzyme, an

adapter containing a “barcode” unique ligated to each sample, and randomly sheared and sequenced. The RAD sequences are then screened for the presence of polymorphisms (reviewed by Davey et al., 2011; Rowe et al., 2011). The main benefit of this method is that it can be used for organisms with reference genomes and for those with no existing genomic data. The RAD technique has proved to be sufficient for genetic mapping in stickleback (Baird et al., 2008), ryegrass (Pfender et al., 2011), and SNP discovery in eggplants (Barchi et al., 2011) and sorghum (Nelson et al., 2011). Therefore, RAD is of interest as a new sequence-based marker for genetic mapping studies.

The continued development of marker types and technologies in order to rapidly assess those markers has resulted in the construction of linkage maps for a wide range of plant species. Having such a map enhances QTL mapping and map-based cloning, and also provides a framework for the use of genetic markers in breeding programs. Although linkage maps have been prepared for many species, the high-resolution maps are limited to major crops such as rice (Yu et al., 2011), and potatoes (van Os et al., 2006). Most barley maps have been generated using earlier generation molecular markers such as restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), and simple sequence repeat (SSR) (Hori et al., 2003). With the availability of NGS, dense maps substantially increased by integrating high-throughput markers with prior reported markers. For example, a 2383 locus map was developed in the OWB mapping population using 1472 SNP and 911 prior markers (Suzcs et al., 2009). Other high-

density linkage maps have been published recently. Sato et al. (2009) constructed an approximately 3000-EST locus map from a cross between Haruna Nijo and H602. Close et al. (2009) built a consensus map containing approximately 3000 SNP loci using four mapping populations. Those maps permit the localization of genes of interest, improve the estimation and rate of successful genetic dissection of QTL, and enable genome-wide association mapping (GW-AM).

Detecting QTLs: from biparental populations to association mapping panels

Linkage and association analysis are two common approaches used for dissecting complex traits. Both methods exploit recombination that can be correlated with phenotypic variation. Biparental mapping is the method traditionally used in linkage analysis. It aims to identify linked markers for indirect selection in breeding as well as to understand the mode of inheritance and genetic architecture of traits (Semagn et al., 2010). Biparental mapping has been successful in identifying QTL that are associated with traits of interest not only in the model plant *Arabidopsis* (Alonso-Blanco et al., 2009), but also in a variety of plant species, including barley (Chen et al., 2009; Pan et al., 1994; Reinheimer et al., 2004; Wang et al., 2010).

Biparental mapping in plants has been based on a cross between two parents that are phenotypically contrasting for the trait(s) of interest. Generally, QTL identified by biparental mapping covers large chromosomal regions because of a limited number of recombination events that occurred during the development of the population. Additional steps may be required to narrow QTL regions for position cloning (Francia et al., 2007). Furthermore since QTL are detected from controlled

crosses, their effects need to be validated in different genetic backgrounds before using in MAS (Romagosa et al., 1999). Several requirements for successful biparental mapping are saturated linkage maps, proper statistical methods for mapping, and a sufficient number of markers. Reducing marker spacing below 10-20 cM does not increase the power of detecting QTL regardless of population size and gene effect (Darvasi et al., 1993). In addition, the size of a population could have a significant impact on the power and precision of QTL mapping. A study on the effect of population size on barley stripe rust, for example, demonstrated that an increase in population size resulted in an increased number of QTL and small sample sizes failed to detect QTL of small effect (Vales et al., 2005). The bias of the estimated position and effects of QTL, as a consequence, will reduce the benefit of applying MAS.

Genome-wide association mapping is a form of association mapping (AM) that surveys common genetic variation in an entire genome to find the associations of a molecular marker with a phenotypic trait (Gupta et al., 2005). Association mapping is increasingly being utilized as a complementary method to biparental mapping. Since AM attempts to use genetic diversity in the existing germplasm or natural population and exploits historical and evolutionary recombination, it provides better mapping resolution, decreases cost, and reveals greater allele numbers (Yu and Buckler, 2006). However, the significant association may sometimes be due to recent occurrences of linkage disequilibrium (LD) (non-random association of alleles at different loci) rather than a close physical linkage between the two loci, particularly for those with

relatively small effect. As a consequence, genomic localization and cloning of genes based on LD may not be successful.

Genome-wide association mapping uses real breeding populations in which genes controlling target traits are segregating, allowing the breeder to map and develop varieties at the same time (Semagn et al., 2010). On the other hand, the AM approach may not assist in the identification of novel alleles. Therefore, additional crosses will be required for the introgression of favorable alleles into selected lines (Tanksley and Nelson, 1996). The marker density required for a whole-genome scan depends on the extent of LD and genome size, hence, it varies among species (Collard et al., 2005). In a species with extensive length of LD, fewer markers are needed, but resolution is lower (Zhao et al., 2007). In contrast, a less extensive level of LD (rapid LD decays) requires many markers, but provides higher a LD mapping resolution (Rafalski, 2002).

Association mapping identifies marker: trait relationships on the basis of LD (Flint-Garcia et al., 2003). Although recombination has a strong impact on LD, there are other factors affecting LD, including population structure and size, admixture, selection, inbreeding, degree of relatedness, and mutation (Gupta et al., 2005). Ignoring these factors can lead to spurious marker: trait association. Cuesta-Marcos et al. (2010) showed that using small arrays in association analysis may create a large number of false-positives. The factors that affected the structure of LD and limited this approach have been discussed elsewhere. Several statistical methods have been proposed to correct for LD caused by population structure and familial relatedness.

These include structured association genomic control, principal component analysis (PCA), and a mixed model approach (reviewed by Myles et al., 2009). Association mapping has proven to have successful application in various crops such as rice (Agrama et al., 2007) and barley (Cuesta-Marcos et al., 2010; von Zitzewitz et al., 2011; Wang et al., 2011), especially after minimizing spurious associations.

The thesis: from biparental mapping using novel markers in a model population to a search for food quality, low temperature tolerance, and vernalization sensitivity quantitative trait loci

The first chapter of this thesis describes the utility of new SNP-based markers, RAD sequences, for linkage mapping in barley. The RAD markers were evaluated for their ability to construct genetic maps and localize genes. The OWB mapping population was used for constructing the RAD map and as a reference for comparison of RAD maps and previous EST-based SNP maps. The application of these maps was demonstrated by biparental mapping. The results of this research were published in BMC Genomics (Chutimanitsakun et al., 2011). The second chapter describes the effectiveness of MAS for waxy and VS traits, and the validation of GW-AM in identifying genes associated with the traits of interest. The perfect markers for waxy (*WX*), vernalization genes (*VRN-H1* and *VRN-H2*) were previously used for MAS in the development of winter waxy barley, based on the assumption that the waxy gene would lead to high β-glucan and VS would lead to maximize LTT for fall-sown barley. A phenotypic selection (PS) population for winter food (non-waxy) barley in a different germplasm base was developed concurrently. The entire germplasm array

(MAS and PS) was used for GW-AM under the auspices of the US Barley Coordinated Agricultural project (CAP).

CONSTRUCTION AND APPLICATION OF A RESTRICTION SITE ASSOCIATED DNA (RAD) LINKAGE MAP IN BARLEY

Yada Chutimanitsakun, Rick W. Nipper, Alfonso Cuesta-Marcos, Luis Cistué, Ann Corey, Tanya Filichkin, Eric A. Johnson, Patrick M. Hayes

BMC Genomics

BioMed Central Ltd
Floor 6, 236 Gray's Inn Road, London, WC1X 8HB, United Kingdom

Published: 4 January 2011

ABSTRACT

Background

Linkage maps are an integral resource for dissection of complex genetic traits in plant and animal species. Canonical map construction follows a well-established workflow: an initial discovery phase where genetic markers are mined from a small pool of individuals, followed by genotyping of selected mapping populations using sets of marker panels. A newly developed sequence-based marker technology, Restriction-site Associated DNA (RAD), enables synchronous single nucleotide polymorphism (SNP) marker discovery and genotyping using massively parallel sequencing. The objective of this research was to assess the utility of RAD markers for linkage map construction, employing barley as a model system. Using the published high density EST-based SNP map in the Oregon Wolfe Barley (OWB) mapping population as a reference, we created a RAD map using a limited set of prior markers to establish linkage group identity, integrated the RAD and prior data, and used both maps for detection of quantitative trait loci (QTL).

Results

Using the RAD protocol in tandem with the Illumina sequence by synthesis platform, a total of 530 SNP markers were identified from initial scans of the OWB parental inbred lines - the "dominant" and "recessive" marker stocks - and scored in a 93 member doubled haploid (DH) mapping population. RAD sequence data from the structured population was converted into allele genotypes from which a genetic map was constructed. The assembled RAD-only map consists of 445 markers with an

average interval length of 5 cM, while an integrated map includes 463 RAD loci and 2383 prior markers. Sequenced RAD markers are distributed across all seven chromosomes, with polymorphic loci emanating from both coding and noncoding regions in the *Hordeum* genome. Total map lengths are comparable and the order of common markers is identical in both maps. The same large-effect QTL for reproductive fitness traits were detected with both maps and the majority of these QTL were coincident with a dwarfing gene (*ZEO)* and the *VRS1* gene, which determines the two-row and six-row germplasm groups of barley.

Conclusions

We demonstrate how sequenced RAD markers can be leveraged to produce high quality linkage maps for detection of single gene loci and QTLs. By combining SNP discovery and genotyping into parallel sequencing events, RAD markers should be a useful molecular breeding tool for a range of crop species. Expected improvements in cost and throughput of second and third-generation sequencing technologies will enable more powerful applications of the sequenced RAD marker system, including improvements in *de novo* genome assembly, development of ultra-high density genetic maps and association mapping.

INTRODUCTION

Plant breeders and geneticists have benefited from the availability of tools for the rapid and cost-effective development of molecular marker-based linkage maps. As predicted by Tanksley et al. (1992), linkage maps have proven to be useful for discovering, dissecting and manipulating the genes that determine simple and complex traits in crop plants. Barley (*Hordeum vulgare*) is a model for plant breeding and genetics because it is diploid (2n = 2x = 14) and has a long history of genetics research. Over the past decade, increasingly dense maps of the barley genome have been constructed using multiple populations and many types of molecular markers (available at http://wheat.pw.usda.gov/GG2/Barley/). Most recently, Szucs et al. (2009) reported an integrated 2383-locus linkage map developed in the Oregon Wolfe Barley (OWB) mapping population based on representative early generation markers (e.g. morphological loci, RFLPs, and SSRs) and single nucleotide polymorphisms (SNPs).

SNP markers have become increasingly important tools for molecular genetic analysis, as single base-pair changes are the most abundant small-scale genetic variation present between related sequences of DNA (Suh and Vijg, 2005). To date, most SNP development efforts in larger, more complex genomes such as barley have focused on "complexity reduction" techniques that aim to sequence a fraction of the genome, such as that represented in EST collections. Once a panel of markers is established from initial SNP discovery, samples from a selected population are then genotyped using oligo-extension or array-based platforms (Fan et al., 2006). Both

these strategies were used for construction of the current barley SNP-based maps (Close et al., 2009; Sato et al., 2009; Szucs et al., 2009).

The emergence of massively-parallel, next-generation sequencing (NGS) platforms capable of producing millions of short (50-100 bp) DNA sequence reads has reduced the costs of DNA sequencing and offers the tantalizing possibility of making direct, genotyping-by-sequencing (GBS) practical (reviewed in Metzker, 2010). Recently, Huang et al. (2009) have elegantly demonstrated how genotyping using NGS data can facilitate the rapid development of linkage maps in domesticated rice, *Oryza sativa*. Despite the attractiveness of this approach and availability of next-generation sequencing platforms, at present, GBS methods retain significant limitations. First, current protocols for synthesis of DNA fragment libraries compatible with high-throughput sequencing platforms are laborious, costly and would be impractical for production efforts involving hundreds of samples (Bentley et al., 2008). Second, sequence-based genotyping is restricted to those species with available, high-quality, pseudomolecule-sized genome assemblies (Huang et al., 2009). While many key economic and scientifically meritorious species will undoubtedly be sequenced as a direct result of the ongoing revolution in NGS technologies, what is required are marker platforms that can provide GBS independent of the status of an assembled genome.

Restriction-site Associated DNA (RAD) markers detect genetic variation adjacent to restriction enzyme cleavage sites across a target genome (Miller et al., 2007a). The first iteration of RAD markers facilitated cloning of mutants isolated from

genetic screens in classic model systems (Lewis et al., 2007; Miller et al., 2007b). More recent efforts have focused on adapting the RAD technique for use in NGS platforms, specifically the Illumina sequencing-by-synthesis method, to enable individual sequence based genotyping of samples (Baird et al., 2008). The sequenced RAD marker system enjoys two favourable characteristics for high-throughput GBS. As previously mentioned, the RAD method uses restriction enzymes as a complexity reduction strategy to reduce the sequenced portion of the genome anywhere from 0.01% to 10% (Ganal et al., 2009). Furthermore, RAD protocols facilitate the creation of highly multiplexed NGS sequencing formulations containing many tens of samples in a single library, thereby reducing library preparation costs (Baird et al., 2008). While previously published RAD studies have explored NGS of limited numbers of individuals or bulked genotyping of pooled populations, the objective of this research was to determine the feasibility of constructing a RAD marker genetic map in barley. We used the OWB population as a mapping resource in order to directly compare RAD and EST-based SNP maps and to assess the quality and utility of a linkage map built with the two types of data.

MATERIALS AND METHODS

Plant material

The mapping population consists of 93 doubled haploid (DH) lines. The DH lines were produced from the F_1 of the cross of the Wolfe recessive and dominant marker stocks using the *Hordeum bulbosum* method (Costa et al., 2001). In the course of this research we determined that nine sets of DH lines had identical genotypes. Specifically, the following sets of lines are identical: set1 = DH 1,4,27,62; set2 = DH 16,71; set3 = DH 5,18; set4 = DH 31,58; set5 = 35,50; set6 = DH 15, 47 set7 = DH 61, 88; set8 = DH 22,70; set9 = DH 80,77. Retention of one genotype per set (DH 4, 16, 18, 31, 35, 47, 61, 70 and 77) reduces the population size to 82. This report describes mapping and QTL analysis using the OWB population of 82 lines. In order to ascertain the bias introduced by duplicate lines (an unintended consequence of the DH production process), all analyses were also conducted with a population size of n = 93 (available at http://wheat.pw.usda.gov/ggpages/maps/OWB/). Genomic DNA was extracted from young leaf tissue of a single plant representing each DH line, and each of the parents, using DNeasy plant maxi kits (QIAGEN Inc. California, USA).

RAD protocols

OWB genomic DNA from the selected mapping population was digested with the restriction endonuclease *SbfI* and processed into RAD libraries similarly to the method of Baird et al. (2008). Briefly, P0 (parental genotypes) and DH (progeny) genomic DNA (~300 ng; from each sample) was digested for 60 min at 37°C in a 50 μL reaction with 20 units (U) of *SbfI* (New England Biolabs [NEB]). Samples were

heat-inactivated for 20 min at 65°C. 2.0 µL of 100 nM P1 Adapter(s), a modified Solexa© adapter (2006 Illumina, Inc., all rights reserved). *SbfI* P1 adapters each contained a unique multiplex sequence index (barcode) which is read during the first four nucleotides of the Illumina sequence read. 100 P1 nM adaptor were added to each sample along with 1 µL of 10 mM rATP (Promega), 1 µL 10× NEB Buffer 4, 1.0 µL (1000 U) T4 DNA Ligase (high concentration, Enzymatics, Inc), 5 µL H_2O and incubated at room temperature (RT) for 20 min. Samples were again heat-inactivated for 20 min at 65°C, pooled and randomly sheared with a Bioruptor (Diagenode) to an average size of 500 bp. Samples were then run out on a 1.5% agarose (Sigma), 0.5× TBE gel and DNA 300 bp to 700 bp was isolated using a MinElute Gel Extraction Kit (Qiagen). End blunting enzymes (Enzymatics, Inc) were then used to polish the ends of the DNA. Samples were then purified using a Minelute column (Qiagen) and 15 U of Klenow exo$^-$ (Enzymatics) was used to add adenine (Fermentas) overhangs on the 3' end of the DNA at 37°C. After subsequent purification, 1 µL of 10 µM P2 adapter, a divergent modified Solexa© adapter (2006 Illumina, Inc., all rights reserved), was ligated to the obtained DNA fragments at 18°C. Samples were again purified and eluted in 50 µL. The eluate was quantified using a Qubit fluorimeter and 20 ng of this product was used in a PCR amplification with 20 µL Phusion Master Mix (NEB), 5 µL of 10 µM modified Solexa© Amplification primer mix (2006 Illumina, Inc., all rights reserved) and up to 100 µL H_2O. Phusion PCR settings followed product guidelines (NEB) for a total of 18 cycles. Samples were gel purified, excising DNA 300-650 bp, and diluted to 1 nM.

To promote SNP identification in low-copy, gene-rich regions of the barley genome, a species with ~90% retroelement content, selection of a restriction enzyme that does not fragment repetitive-class DNA is desirable. Previous studies have documented epigenetic modification of CpG, CpNpG and CpNpN nucleotides with 5-methylcytosine (5 mC) in retroelement-dense regions of many plant genomes, including triticale (Ma and Gustafson, 2006; Palmer et al., 2003; Rabinowicz et al., 2005). Methylation-sensitive type II restriction endonucleases, which do not cleave 5 mC-modified DNA, can be used to specifically sample the hypomethylated genomic fraction and are commonly used in other restriction-enzyme based genetic marker systems (ReynaLopez et al., 1997). We selected the restriction enzyme *Sbf*I, (5'CCTGCA/GG'3) with a recognition site containing two CpNpG trinucelotide repeats for RAD sequencing of the barley genome.

Illumina Sequencing

The constructed OWB libraries were run on an Illumina Genome Analyzer II at the University of Oregon High Throughput Sequencing Facility. Illumina/Solexa protocols were followed for single read (1 × 36 bp) sequencing chemistry. A total of 20.4 M Illumina reads were obtained from sequencing of the population. Sequences are available at the Sequence Read Archive http://www.ncbi.nlm.nih.gov/Traces/sra/, at accession SRA020593.

Sequence Analysis and SNP Discovery and Genotyping

Internal Floragenex sequence tools and custom PerlScripts were used for processing of raw Illumina/Solexa data. Data from multiple Illumina/Solexa sequence

channels was segregated by the appropriate four nucleotide multiplex identifier (MID) assigned to each sample. All reads were trimmed to 28 nucleotides from the 3' end of genomic sequence to avoid using bases with a high Illumina sequence error rate.

Sequence Alignment and Comparative Genomics

The short-read alignment program Bowtie (Langmead et al., 2009) was used for mapping of polymorphic barley RAD sequence loci to the comprehensive *Hordeum* gene index (HvGI v10.2) database from the Dana-Farber Cancer Institute (Quackenbush et al., 2001). Both tentative consensus (TC) and singleton expressed sequence tags (ESTs) were used in analysis. Briefly, sequences corresponding to all 530 polymorphic RAD loci were aligned against the HvGI assembly. Two criteria were imposed for sequence mapping. First, a maximum of three nucleotide mismatches and no gaps between the RAD sequence and reference were permitted for any alignment. Second, each sequence had to anchor to a single unique position to be scored. For macro-scale syntenic mapping of barley RAD sequences to other grass genomes, we extended the CIP/CALP (Conserved Identity Percentage/Conserved Alignment Percentage) method previously used in Triticale comparative analysis (Salse et al., 2008). 30 bp RAD sequences ordered by the linkage map were aligned against the *Oryza sativa* and *Brachypodium distachyon* chromosome assemblies using relaxed Bowtie alignment parameters. Bowtie is able to tolerate up to three nucleotide mismatches between query and reference, translating to minimum values of 90% and 90% respectively for CIP and CALP.

Linkage mapping

Two linkage maps were constructed. The first map was built with only the RAD data and data for nine morphological markers (Table 1.1). The morphological marker data were reported by Szucs et al. (2009) and were included because they provide anchors for equating linkage groups with six of the seven barley chromosomes. A second map was built using RAD data and all 2383 data points reported by Szucs et al. (2009). Each linkage map was constructed using JoinMap 4 (Van Ooijen, 2006). Linkage groups were identified using minimum LOD values of 5. The Monte Carlo Maximum Likelihood (ML) mapping algorithm was used to determine the orders of markers within each linkage group. Map distances were calculated using the Haldane's mapping function. Maps were drawn using MapChart v2.2 (Voorrips, 2002). Data used for linkage map construction are available at Oregon Wolfe Barley Data and GrainGenes Tools http://wheat.pw.usda.gov/ggpages/maps/OWB/.

Phenotyping

In order to assess the utility of the RAD and RAD + SNP map for quantitative trait locus (QTL) detection, data on phenological and reproductive fitness phenotypes were obtained for the 93 DH lines and the two parents. Individual plants were grown in 13.5 cm pots at the Oregon State University greenhouses (Corvallis, Oregon USA). Supplemental light was used to maintain a 16 h light/24 h photoperiod. Temperatures were maintained at a constant 18 ± 2°C day and night temperature. Each DH and parental line was replicated twice. Eight traits were measured on each plant. The trait abbreviations and definitions are as follows: (1) *Final leaf number* (FLN) was

recorded as the total number of leaves on the main stem of each plant; (2) *Plant height* (PH) was measured as the distance (in cm) from the soil surface to the tip of the tallest inflorescence (spike), exclusive of awns, if present; (3) *Spike number* (SN) was the actual count of the total number of fertile spike on each plant. Three stems with fertile spikes were selected at random from each plant for determining the following traits, and the individual values were averaged: (4) *Spike length* (SL) was measured as the length (in cm) from the first rachis internode to the top of the final fully formed floret, exclusive of awn; (5) *Floret number* (FS) was the count of the number of florets (fertile and sterile) per spike; (6) *Grain number* (GN) was the count of the number of seed-containing florets per spike; (8) *Hundred grain weight* (HGW) was the weight (in g) of 100 grains. Grain yield per plant (GY) was estimated by the function GY = SN*GN*HGW. Phenotype data are available at Oregon Wolfe Barley Data and GrainGenes Tools http://wheat.pw.usda.gov/ggpages/maps/OWB/.

QTL analysis

QTL analyses were performed for each of the nine traits using the RAD-only and RAD + 2383 locus maps as follows: For the RAD-only map, all data included in the linkage map were used. For the RAD + 2383 locus map, a skeleton map was developed using a single marker (selected at random) for an average marker density of 2 cM and a total of 624 markers. The QTL analyses were conducted with QTL Cartographer Version 2.5 (Wang et al., 2005) using Composite Interval Mapping (CIM) (Zeng, 1994). Up to seven cofactors for CIM were chosen, using a forward-selection backward-elimination stepwise regression procedure with a significance

threshold of 0.1. The walk speed was set to 1 cM, and the scan window to 50 cM beyond the markers flanking the interval tested. Experiment-wise significance (α = 0.05) likelihood ratio test (LR) thresholds for QTL identification were determined with 1,000 permutations, and expressed as LOD (LOD = 0.217 LR). Epistatic interactions between QTL were evaluated with the Multiple Interval Mapping (MIM) (Kao et al., 1999) method implemented in Windows QTL Cartographer using Bayesian Information Criteria (BIC-M0). Broad-sense heritability values were estimated using the following formula:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_e^2}{r}}$$

where σ_G^2 represent the genetic variance, σ_e^2 the residual variance and r the number of replicates per genotype.

RESULTS AND DISCUSSION

Genome Analyzer sequence results, SNP Discovery and DH Genotyping

A total of 2,010,583 36-bp sequence reads were obtained for the OWB dominant and recessive inbred genetic stocks (parents of the OWB mapping population), while 27,704,592 sequence reads were obtained for the 93 member DH mapping population. Illumina sequences from the OWB parental lines were first used for identification of SNPs. Putative alleles were mined from the sequence data using several custom PerlScripts and filtering procedures. First, raw 36-bp Illumina sequence reads were partitioned into discrete files using a 5' multiplex identifier (MID) corresponding to each OWB sample and the restriction enzyme site *SbfI* (TGCAGG). Segregated data from each line was then collapsed into putative RAD sequence clusters comprised of a minimum of eight (8×) redundant sequence reads per locus. Sequences not attaining the 8× sequence coverage threshold were excluded from further analysis, as were putative high-copy RAD sequences where the number of sequence occurrences in each cluster was greater than 500 (500×). Homologous RAD clusters from the dominant and recessive lines were then compared using a custom k-mer matching algorithm permitting exact sequence matches (monomorphic loci), single mismatch (one SNP per read) and two nucleotide mismatches (two SNPs per read) per 28 bp sequence. An initial panel of 530 SNPs with fixed genotypes in both parents were identified using these criteria and alleles for each marker were assigned to their respective parental donor.

The putative 530 SNP marker panel was then used to score RAD sequences obtained from each of the DH individuals. As alleles are fixed within each member of the doubled haploid OWB population, we posited sequence genotypes could be accurately determined at low sequence coverage (<5×) (Lander and Waterman, 1988). To further minimize genotyping miscalls due to possible sequencing errors, a minimum of two independent sequence reads were required over any locus to assign any SNP genotype. Putative genotypes developed for individual samples were converted into JoinMap 4 (Van Ooijen, 2006) compatible format using custom PerlScript. Loci lacking sufficient sequence coverage or with conflicting genotype data were coded as missing data.

Linkage map

We used the following criteria to assess the quality of the RAD markers for linkage map construction. First, with the RAD-only map we considered the (i) total number of loci detected, (ii) the percentage of polymorphic loci, (iii) the number of missing allele calls for polymorphic loci, (iv) the percentage of codominant loci, (v) segregation distortion, (vi) the number of significant singletons (vii) linkage map length, and (viii) the number, location, interaction and effect of significant QTL. For purposes of comparison, we used the map reported by Szucs et al. (2009). Subsequently, we added the RAD data to the 2383 locus map and assessed criteria v - vii, above. For criteria viii, however, we used a skeleton map, as described in the Methods. First, we will present results in terms of criteria i - vii; the QTL results will be presented separately.

Of approximately 10,000 RAD sequence clusters interrogated, 530 loci (5.3%) were classified as codominant markers where two distinct alleles were explicitly observed between the OWB parents. A number of dominant-style markers, which are sequences present in one parent but not the other, were also observed within the data but were not used for map construction as dominant markers have reduced genotyping quality. Of the codominant RAD marker class, 67 (13%) were excluded from further analysis due to missing data (≥ 15% missing data points). This left 463 (4.6% of the total) RAD loci, plus the nine morphological markers, for map construction. Twenty-seven RAD markers remained unlinked at LOD 5.0 and the remaining 436 formed seven linkage groups, together with the nine morphological markers. Based on visual assessment of locus orders, there were 22 loci showing apparent double crossover events. Of these, 23 singletons data points were re-coded as missing data for 20 loci where these occurred, except for two loci where distances between flanking markers were large enough to expect recombination. The final map is composed of 436 RAD and nine morphological markers. The total length of the RAD map is 1260 cM. Excluding co-segregating markers, the average marker density is 5 cM. Significant segregation distortion was observed on chromosomes 2H, 3H, 6H, and 7H (Figure 1.1). On chromosomes 2H and 3H the segregation distortion was in favor of the OWB recessive parent allele and on chromosomes 6H and 7H it was in favor of the OWB dominant parent allele. The lengths (in Haldane cM) for each linkage group are shown in Table 1.2.

For construction of the RAD + 2383 locus map, the same 463 RADs selected initially for the RAD-only map were added to the 2383-locus data set reported by Szucs et al. (2009). The 23 singletons were replaced by missing values. The combined map therefore consists of 2846 loci and has a total length of 1286 cM (Table 1.2). Marker orders for the non-RAD markers are consistent between the 2383 and 2846-locus maps. Seventy-eight percent (359) of the RAD markers co-segregate with one or more of the previous markers. There were examples of gap-filling: for example, FGX_OWB00091, mapped to a 17 cM gap on chromosome 7H in Szucs et al. (2009), and incorporation of this marker reduced the distance between the two flanking markers to 10 cM. Segregation distortion was observed at the same positions as in the RAD-only map (Figure 1.1). The lengths (in Haldane cM) for each linkage group are shown in Table 1.1. The same lines identified as identical with the RAD-only data (see Methods, Plant material) were confirmed as identical using the 2383 data points reported by Szucs et al. (2009).

Although a significant number of RAD loci were eliminated based on lack of polymorphism and missing sequence data, the genome scan uncovered over 400 high quality loci that were available for map construction. By way of comparison, there are 722 DArT loci on the Szucs et al. (2009) OWB map, out of the 1,500 loci that were genotyped. The RAD loci are codominant whereas DArTs are dominant markers (Wenzl et al., 2004). In the case of dominant markers, missing data due to error vs. allele absence cannot be distinguished, and this leads to a higher frequency of apparent singletons in map construction. The high quality of the RAD data is further confirmed

by the comparable linkage map lengths for the RAD only, RAD + prior marker, and Szucs et al. (2009) OWB map (Table 1.2). Segregation distortion was observed in all maps at the equivalent positions confirming that this was due to non-random distribution of alleles to haploid progeny and not to scoring errors. The pronounced segregation distortion on 2H is attributable to the *ZEO* locus, with selection against the "dwarfing" alleles of the dominant parent.

The presence of duplicate sets of lines in the OWB population provides an additional test for data quality. The members of each set were not identified as identical in previous iterations of the map (e.g. Costa et al., 2001) due to differences at loci that have been progressively removed from the data set based on quality control criteria. The lines within each subset are identical for the Illumina SNPs and all other loci included in the Szucs et al. (2009) OWB map. That the lines within each set are also identical for all RADs confirms the repeatability of the RAD genotyping assay and that the lines are identical. The most likely explanation for the presence of these identical sets of lines in the population is that multiple haploids were inadvertently advanced from callus regenerated from a single embryo. Removal of the sets of identical lines reduces the mapping population size from 93 to 82. There are no differences in locus order between the n = 93 and n = 82 maps and map lengths are comparable (available at http://wheat.pw.usda.gov/ggpages/maps/OWB/).

EST and genome mapping of RAD sequence markers

The RAD technique develops sequence from regions adjacent to restriction endonuclease digestion sites in a target genome (Baird et al., 2008). To establish if

sequence-based RAD markers from the OWB genetic map would anchor to existing *Hordeum* genomic resources, we used the short-read aligner Bowtie to map RAD sequences onto a barley gene index (Langmead et al., 2009; Quackenbush et al., 2001). Using this database, we successfully identified unique alignment positions for 51 of 436 sequenced RAD loci (11.0%). An additional 22 RAD loci (4.7%) mapped to multiple positions in the gene index. Although the gene index contains approximately 54 Mb of putative coding sequence distributed across 80,723 tentative assemblies, this database spans only a small fraction (~0.1%) of the 5.0 Gb barley haploid genome. As Ty3 and Copia retrotransposon families are believed to inhabit a large portion of the barley genome, we postulated some percentage of RAD sequences might originate from repetitive-class sequences (Ouyang and Buell, 2004). However, several attempts to align the 463 RAD sequence loci to the 1.3 Mb TIGR *Hordeum* repeat database under a variety of thresholds did not reveal any successful alignments. A larger percentage of RAD sequences could be positioned on candidate genes than would expected by random sampling, suggesting that RAD markers are significantly enriched in the gene space. The absence of any alignments to known repetitive sequences also hints that RAD markers are clustered within recombinatorially active regions of the genome.

Comparative Genome Analysis

To examine if assembled grass genomes would serve to anchor other RAD markers, we aligned polymorphic sequences to the 430 Mb *Oryza sativa* and 300 Mb *Brachypodium distachyon* genomes using a modified CIP/CALC method (Salse et al.,

2008; Vogel et al., 2010; Yu et al., 2002). Bowtie alignment results using relaxed parameters indicate that only 16 and 24 of the 463 OWB RAD sequences mapped to either the rice or *Brachypodium* chromosome assemblies, respectively. Despite the small number of orthologous RAD sequences and the short Illumina read of 28 bp, alignments of RAD markers ordered by the genetic map against the finished *Brachypodium* genome (Figure 1.2) agree with macro-scale syntenic relationships established by previous efforts (Vogel et al., 2010). Although this study has relatively few sequence loci available for comparison, our findings suggest that a denser RAD marker scan, using a more frequently cutting restriction enzyme would interrogate more genome sequence and interrogate more sequence for comparative analyses.

Overall, we were able to assign 74 of 463 RAD sequence loci (15.9%) to at least one of the three sequence references, leaving the genomic origin of the remaining barley RAD tags (389 loci, 84.1%) unknown. We postulate the large numbers of RAD sequences placed on the OWB linkage without homology or orthology to known sequences are a result of two factors. First, the lack of a contiguous barley genome, which would allow us to explicitly determine the location of all RAD sequences, restricts our analysis to the small fraction of the haploid genome that has been sequenced. Second, despite established syntenic relationships between the *Oryza, Hordeum* and *Brachypodium* genomes, the inefficient mapping of barley sequenced RAD markers across species is likely a result of the majority of RAD loci emanating from areas of the barley genome which have significantly diverged at the nucleotide

level since the speciation of the Poaceae (Bennetzen and Freeling, 1997; Gale and Devos, 1998).

A cohesive explanation for the results observed in the genetic map and comparative genome analysis is that the majority of RAD loci are linked with, but lie outside gene sequences. In this study, although only 11.0% of RAD sequences align to known barley genes, we report 78% of RAD markers show co-segregation with unigene-EST SNP markers from the Szucs et al. (2009) OWB map. The observed association of RAD markers with known genic-SNPs indicates they are genetically linked, suggesting some physical proximity, though the distances may be on the order of megabases. Additionally, the relative paucity of RAD markers that align to barley genes or other plant genomes indicates that only a small fraction of RAD markers originate from within coding or other conserved sequences. RAD marker development efforts from other grass species for which there is a reference genome show similar distributions of markers across coding and intergenic space (Nipper et al., 2009). When a complete barley genome sequence is available, the sequence identity and location of RAD loci will become clear. In the interim, the current availability of all barley RAD sequences is an advantage over DArTs, where only limited sequence data are publicly available.

QTL mapping

One of the principal applications of linkage maps to crop improvement has been QTL mapping in bi-parental crosses (available at

http://wheat.pw.usda.gov/GG2/Barley/). A principal problem with many QTL mapping efforts is the limited size of the mapping population (Keurentjes et al., 2007; Li et al., 2006; Vales et al., 2005; Zhao et al., 2007). Recognizing that the small size of the OWB population (n = 93 and n = 82 when removing identical lines) will lead to biased estimates of QTL significance, effect, and interaction (Arbelbide et al., 2006; Melchinger et al., 1998; Schon et al., 2004; Utz et al., 2000), we nonetheless proceeded with a QTL analysis of the eight traits, due to the high heritabilities (Table 1.3 and Table 1.4) and our interest in addressing two issues. The OWB population is a widely-used resource for genetic analysis and instruction: reporting the relationships of QTLs with the morphological and phenological characters segregating in the population will further develop this community resource. The RAD markers added to the map reported by Szucs et al. (2009) represent very high quality and novel data and we were interested in determining if their addition would fill gaps in the previous map and thus allow for higher resolution QTL detection.

As shown in Table 1.3, a total of 26 QTLs were found using the higher density map, with a range of one to five QTL for each individual trait. Twenty-six QTLs were also detected with the RAD-only map with a range of two to five QTL for each trait (Table 1.4). Twenty-three QTLs were significant and detected in both maps. Of the three QTL that were significant in the full map, but not the RAD-only map, all showed a trend in the RAD-only map but did not reach the LOD threshold. Three QTL significant in the RAD-only map but not in the full map showed a trend in the full map but did not reach the LOD threshold. Therefore, RADs alone, or in combination with

other markers, are suitable for QTL mapping. This supports the quality of the RAD data, since a key issue for QTL detection is marker quality, given adequate genome coverage (Arbelbide et al., 2006).

The following results highlight findings from the higher density skeleton map (Table 1.3), based on the assumption that by providing the most thorough coverage it optimizes QTL estimates. However, the same large-effect QTL were detected with the RAD-only map (Table 1.4). As shown in Table 1.3, eleven of the twenty-six QTL were associated with four genes: *ZEO-1*, *VRS-1*, *VRN-H1* and *VRN-H2*, and the largest effect QTL for all traits were associated with *ZEO-1*and/or *VRS-1*. The favorable alleles for height, spike length, grain number and grain yield came from the OWB recessive parent (normal height, long spike, and six-row) at *ZEO-1*. The OWB recessive parent also contributed favorable alleles for floret and grain number at *VRS-1*. At this locus, the OWB dominant parent (dwarf height, short spike, and two-row) contributed favorable alleles for spike number and hundred grain weight. Although *VRS-1* and *ZEO-1* were both coincident with yield component QTL, only *ZEO-1* had a significant effect on grain yield. This is probably due to yield component compensation associated with *VRS-1* and negative pleiotropic effects of the *ZEO-1* dwarf allele. This extreme dwarfing allele will not be as immediately useful to agriculture as the *Rth-B1* and *Rht-D1* genes of wheat (Peng et al., 1999). Interestingly, QTLs for final leaf number were coincident with *VRN-H1* and *VRN-H2*. These two genes interact epistatically to determine vernalization sensitivity (von Zitzewitz et al., 2005). The OWB dominant and recessive parents, respectively, have dominant

(winter) and spring (recessive) alleles at *VRN-H2* allele. Therefore, it is of interest that the OWB dominant allele at *VRN-H2* is associated with higher final leaf number, even though there is no binding site in *Vrn-H1* for the repressor encoded by *VRN-H2* since both parents have the same recessive (spring) allele at *VRN-H1* (Szucs et al., 2007). The higher final leaf number QTL allele coincident with *VRN-H1* may be a consequence of regulation of other regions in *VRN-H1* besides *VRN-H2*. There were epistatic QTL interactions for spike length, and grain number but these effects were very small in comparison to the main effects. The QTL we report for the OWB population can be aligned with QTL for other traits assessed in other germplasm via the GrainGenes QTL summary http://wheat.pw.usda.gov/ggpages/maps/OWB/.

Conclusions

In this study we showed that sequenced RAD markers were sufficient to generate a high quality linkage map comparable to current OWB SNP-based maps. The success of linkage map construction supports the reliability of the sequenced RAD markers based on the following criteria i) a small number of singletons ii) consistency with non-RAD marker order iii) segregation distortion between maps in equivalent positions iv) comparable genome coverage and v) comparable map lengths. Construction of this linkage map could serve as a bridge to allow identification of loci associated with traits of interest, thus facilitating gene discovery and manipulation. The consistency of QTL results between RAD and RAD + prior marker maps confirms that sequenced RAD markers will be useful for developing genetic maps and

QTL tagging. Therefore, sequenced RAD markers can contribute to the enrichment of molecular marker resources and have useful applications in molecular breeding.

Ongoing optimization of the RAD marker system will foster more sophisticated analysis in future studies. Selection of nucleases that generate more markers will allow higher density linkage maps to be constructed, while improvements in sequencing chemistries and fragment preparation protocols will permit longer read lengths for comparative genome analysis. Additionally, sequenced RAD markers arrayed in genetic maps would be of significant benefit as a scaffold framework for placement of shotgun sequence reads and *de novo* genome assembly refinement.

REFERENCES

Arbelbide M., Yu J., Bernardo R. (2006) Power of mixed-model QTL mapping from phenotypic, pedigree and marker data in self-pollinated crops. Theoretical and Applied Genetics 112:876-884. DOI: 10.1007/s00122-005-0189-7.

Baird N.A., Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z.A., Selker E.U., Cresko W.A., Johnson E.A. (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. Plos One 3. DOI: 10.1371/journal.pone.0003376.

Barley Boulevard: A Shortcut to Barley Information in GrainGenes and Elsewhere [http://wheat.pw.usda.gov/GG2/Barley/]

Bennetzen J.L., Freeling M. (1997) The unified grass genome: Synergy in synteny. Genome Research 7:301-306.

Bentley D.R., Balasubramanian S., Swerdlow H.P., Smith G.P., Milton J., Brown C.G., Hall K.P., Evers D.J., Barnes C.L., Bignell H.R., Boutell J.M., Bryant J., Carter R.J., Cheetham R.K., Cox A.J., Ellis D.J., Flatbush M.R., Gormley N.A., Humphray S.J., Irving L.J., Karbelashvili M.S., Kirk S.M., Li H., Liu X., Maisinger K.S., Murray L.J., Obradovic B., Ost T., Parkinson M.L., Pratt M.R., Rasolonjatovo I.M.J., Reed M.T., Rigatti R., Rodighiero C., Ross M.T., Sabot A., Sankar S.V., Scally A., Schroth G.P., Smith M.E., Smith V.P., Spiridou A., Torrance P.E., Tzonev S.S., Vermaas E.H., Walter K., Wu X., Zhang L., Alam M.D., Anastasi C., Aniebo I.C., Bailey D.M.D., Bancarz I.R., Banerjee S., Barbour S.G., Baybayan P.A., Benoit V.A., Benson K.F., Bevis C., Black P.J., Boodhun A., Brennan J.S., Bridgham J.A., Brown R.C., Brown A.A., Buermann D.H., Bundu A.A., Burrows J.C., Carter N.P., Castillo N., Catenazzi M.C.E., Chang S., Cooley R.N., Crake N.R., Dada O.O., Diakoumakos K.D., Dominguez-Fernandez B., Earnshaw D.J., Egbujor U.C., Elmore D.W., Etchin S.S., Ewan M.R., Fedurco M., Fraser L.J., Fajardo K.V.F., Furey W.S., George D., Gietzen K.J., Goddard C.P., Golda G.S., Granieri P.A., Green D.E., Gustafson D.L., Hansen N.F., Harnish K., Haudenschild C.D., Heyer N.I., Hims M.M., Ho J.T., Horgan A.M., *et al.* (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53-59. DOI: 10.1038/nature07517.

Close T.J., Bhat P.R., Lonardi S., Wu Y.H., Rostoks N., Ramsay L., Druka A., Stein N., Svensson J.T., Wanamaker S., Bozdag S., Roose M.L., Moscou M.J., Chao S.A.M., Varshney R.K., Szucs P., Sato K., Hayes P.M., Matthews D.E., Kleinhofs A., Muehlbauer G.J., DeYoung J., Marshall D.F., Madishetty K., Fenton R.D., Condamine P., Graner A., Waugh R. (2009) Development and implementation of high-throughput SNP genotyping in barley. Bmc Genomics 10. DOI: 10.1186/1471-2164-10-582.

Costa J.M., Corey A., Hayes P.M., Jobet C., Kleinhofs A., Kopisch-Obusch A., Kramer S.F., Kudrna D., Li M., Riera-Lizarazu O., Sato K., Szucs P., Toojinda T., Vales M.I., Wolfe R.I. (2001) Molecular mapping of the Oregon Wolfe

Barleys: a phenotypically polymorphic doubled-haploid population. Theoretical and Applied Genetics 103:415-424.

Fan J.B., Chee M.S., Gunderson K.L. (2006) Highly parallel genomic assays. Nature Reviews Genetics 7:632-644. DOI: 10.1038/nrg1901.

Gale M.D., Devos K.M. (1998) Comparative genetics in the grasses. Proceedings of the National Academy of Sciences of the United States of America 95:1971-1974.

Ganal M.W., Altmann T., Roeder M.S. (2009) SNP identification in crop plants. Current Opinion in Plant Biology 12:211-217. DOI: 10.1016/j.pbi.2008.12.009.

Huang X.H., Feng Q., Qian Q., Zhao Q., Wang L., Wang A.H., Guan J.P., Fan D.L., Weng Q.J., Huang T., Dong G.J., Sang T., Han B. (2009) High-throughput genotyping by whole-genome resequencing. Genome Research 19:1068-1076. DOI: 10.1101/gr.089516.108.

Kao C.H., Zeng Z.B., Teasdale R.D. (1999) Multiple interval mapping for quantitative trait loci. Genetics 152:1203-1216.

Keurentjes J.J.B., Bentsink L., Alonso-Blanco C., Hanhart C.J., Vries H.B.D., Effgen S., Vreugdenhil D., Koornneef M. (2007) Development of a near-isogenic line population of Arabidopsis thaliana and comparison of mapping power with a recombinant inbred line population. Genetics 175:891-905. DOI: 10.1534/genetics.106.066423.

Lander E., Waterman M. (1988) Genomic mapping by fingerprinting random clones: A mathematical analysis. Genomics 2:231-239.

Langmead B., Trapnell C., Pop M., Salzberg S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology 10. DOI: 10.1186/gb-2009-10-3-r25.

Lewis Z.A., Shiver A.L., Stiffler N., Miner M.R., Johnson E.A., Selker E.U. (2007) High-density detection of restriction-site-associated DNA markers for rapid mapping of mutated loci in neurospora. Genetics 177:1163-1171. DOI: 10.1534/genetics.107.078147.

Li X.M., Quigg R.J., Zhou J., Xu S.Z., Masinde G., Mohan S., Baylink D.J. (2006) A critical evaluation of the effect of population size and phenotypic measurement on QTL detection and localization using a large F2 murine mapping population. Genetics and Molecular Biology 29:166-173.

Ma X.F., Gustafson J.P. (2006) Timing and rate of genome variation in triticale following allopolyploidization. Genome 49:950-958. DOI: 10.1139/g06-078.

Melchinger A.E., Utz H.F., Schon C.C. (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383-403.

Metzker M.L. (2010) Applications of Next-Generation Sequencing Sequencing technologies - the next generation. Nature Reviews Genetics 11:31-46. DOI: 10.1038/nrg2626.

Miller M.R., Atwood T.S., Eames B.F., Eberhart J.K., Yan Y.L., Postlethwait J.H., Johnson E.A. (2007b) RAD marker microarrays enable rapid mapping of zebrafish mutations. Genome Biology 8. DOI: 10.1186/gb-2007-8-6-r105.

Miller M.R., Dunham J.P., Amores A., Cresko W.A., Johnson E.A. (2007a) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Research 17:240-248. DOI: 10.1101/gr.5681207.

Nipper R.W., Atwood T.S., Boone J.Q., Gribbin J.M., Johnson E.A. (2009) SNP Discovert in *Zea Mays* using Sequenced Restriction-site Associated DNA Markers. Acta Hort (ISHS) 859:129-133.

Ouyang S., Buell C.R. (2004) The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. Nucleic Acids Research 32:D360-D363. DOI: 10.1093/nar/gkh099.

Palmer L.E., Rabinowicz P.D., O'Shaughnessy A.L., Balija V.S., Nascimento L.U., Dike S., de la Bastide M., Martienssen R.A., McCombie W.R. (2003) Maize genome sequencing by methylation filtrations. Science 302:2115-2117.

Peng J., Richards D.E., Hartley N.M., Murphy G.P., Devos K.M., Flintham J.E., Beales J., Fish L.J., Worland A.J., Pelica F., Sudhakar D., Christou P., Snape J.W., Gale M.D., Harberd N.P. (1999) 'Green Revolution' genes encode mutant gibberellin response modulators. Nature 400(6741):256-261.

Quackenbush J., Cho J., Lee D., Liang F., Holt I., Karamycheva S., Parvizi B., Pertea G., Sultana R., White J. (2001) The TIGR Gene Indices: analysis of gene transcript sequences in highly sampled eukaryotic species. Nucleic Acids Research 29:159-164.

Rabinowicz P.D., Citek R., Budiman M.A., Nunberg A., Bedell J.A., Lakey N., O'Shaughnessy A.L., Nascimento L.U., McCombie W.R., Martienssen R.A. (2005) Differential methylation of genes and repeats in land plants. Genome Research 15:1431-1440. DOI: 10.1101/gr.4100405.

ReynaLopez G.E., Simpson J., RuizHerrera J. (1997) Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. Molecular & General Genetics 253:703-710.

Salse J., Chague V., Bolot S., Magdelenat G., Huneau C., Pont C., Belcram H., Couloux A., Gardais S., Evrard A., Segurens B., Charles M., Ravel C., Samain S., Charmet G., Boudet N., Chalhoub B. (2008) New insights into the origin of the B genome of hexaploid wheat: Evolutionary relationships at the SPA genomic region with the S genome of the diploid relative Aegilops speltoides. Bmc Genomics 9. DOI: 10.1186/1471-2164-9-555.

Sato K., Nankaku N., Takeda K. (2009) A high-density transcript linkage map of barley derived from a single population. Heredity 103:110-117. DOI: 10.1038/hdy.2009.57.

Schon C.C., Utz H.F., Groh S., Truberg B., Openshaw S., Melchinger A.E. (2004) Quantitative trait locus mapping based on resampling in a vast maize testcross

experiment and its relevance to quantitative genetics for complex traits. Genetics 167:485-498.

Suh Y., Vijg J. (2005) SNP discovery in associating genetic variation with human disease phenotypes. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis 573:41-53. DOI: 10.1016/j.mrfmmm.2005.01.005.

Szucs P., Blake V.C., Bhat P.R., Chao S., Close T.J., Cuesta-Marcos A., Muehlbauer G.J., Ramsay L.V., Waugh R., Hayes P.M. An integrated resource for barley linkage map and malting quality QTL alignment. The Plant Genome 2009, 2:134-140. DOI: 10.3835/plantgenome2008.01.0005.

Szucs P., Skinner J.S., Karsai I., Cuesta-Marcos A., Haggard K.G., Corey A.E., Chen T.H.H., Hayes P.M. (2007) Validation of the VRN-H2/VRN-H1 epistatic model in barley reveals that intron length variation in VRN-H1 may account for a continuum of vernalization sensitivity. Molecular Genetics and Genomics 277:249-261. DOI: 10.1007/s00438-006-0195-8.

Tanksley S.D., Ganal M.W., Prince J.P., Devicente M.C., Bonierbale M.W., Broun P., Fulton T.M., Giovannoni J.J., Grandillo S., Martin G.B., Messeguer R., Miller J.C., Miller L., Paterson A.H., Pineda O., Roder M.S., Wing R.A., Wu W., Young N.D. (1992) High-Density Molecular Linkage Maps of The Tomato And Potato Genomes. Genetics 132:1141-1160.

Utz H.F., Melchinger A.E., Schon C.C. (2000) Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics 154:1839-1849.

Vales M.I., Schon C.C., Capettini F., Chen X.M., Corey A.E., Mather D.E., Mundt C.C., Richardson K.L., Sandoval-Islas J.S., Utz H.F., Hayes P.M. (2005) Effect of population size on the estimation of QTL: a test using resistance to barley stripe rust. Theoretical and Applied Genetics 111:1260-1270. DOI: 10.1007/s00122-005-0043-y.

Van Ooijen J.W. (2006) JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV: Wageningen, Netherlands.

Vogel J.P., Garvin D.F., Mockler T.C., Schmutz J., Rokhsar D., Bevan M.W., Barry K., Lucas S., Harmon-Smoth M., Lail K., Tice H., Grimwood J., McKenzie N., Huo N., Gu Y.Q., Lazo G.R., Anderson O.D., You F.M., Luo M.-C., Dvorak J., Wright J., Febrer M., Idziak D., Hasterok R., Lindquist E., Wang M., Fox S.E., Priest H.D., Filichkin S.A., Givan S.A., Bryant D.W., Chang J.H., Wu H., Wu W., Hsia A.-P., Schnable P.S., Kalyanaraman A., Baarbazuk B., Michael T.P., Hazen S.P., Bragg J.N., Laudencia-Chingcuanco D., Weng Y., Haberer G., Spannagl M., Mayer K., Rattei T., Mitros T., Lee S.-J., Rose J.K.C., Mueller L.A., York T.L., Wicker T., Buchmann J.P., Tanskanen J., Schulman A.H., Gundlach H., de Oliveira A.C., Maia L.d.C., Belknap W., Jiang N., Lai J., Zhu L., Ma J., Sun C., Pritham E., Salse J., Murat F., Abrouk M., Bruggmann R., Messing J., Fahlgren N., Sullivan C.M., Carrington J.C., Chapman E.J., May G.D., Zhai J., Ganssmann M., Gurazada S.G.R., German

M., Meyers B.C., Green P.J., Tyler L., Wu J., Thomson J., Chen S., Scheller H.V., Harholt J., Ulvskov P., Kimbrel J.A., Bartley L.E., Cao P., Jung K.-H., Sharma M.K., Vega-Sanchez M., Ronald P., Dardick C.D., De Bodt S., Verelst W., Inze D., *et al.* (2010) Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature 463:763-768. DOI: 10.1038/nature08747.

Von Zitzewitz J., Szűcs P., Dubcovsky J., Yan L.L., Francia E., Pecchioni N., Casas A., Chen T.H.H., Hayes P.M., Skinner J.S. (2055) Molecular and structural characterization of barley vernalization genes. Plant Molecular Biology 59(3):449-467. DOI: 10.1007/s11103-005-0351-2.

Voorrips R.E. (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. Journal of Heredity 93:77-78.

Wang S., Basten C.J., Gaffney P., Zeng Z.B. (2005) Windows QTL Cartographer 2.5 user manual. [http://www.kyazma.nl/docs/JM4manual.pdf].

Wenzl P., Carling J., Kudrna D., Jaccoud D., Huttner E., Kleinhofs A., Kilian A. (2004) Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proceedings of the National Academy of Sciences of the United States of America 101:9915-9920. DOI: 10.1073/pnas.0401076101.

Yu J., Hu S.N., Wang J., Wong G.K.S., Li S.G., Liu B., Deng Y.J., Dai L., Zhou Y., Zhang X.Q., Cao M.L., Liu J., Sun J.D., Tang J.B., Chen Y.J., Huang X.B., Lin W., Ye C., Tong W., Cong L.J., Geng J.N., Han Y.J., Li L., Li W., Hu G.Q., Huang X.G., Li W.J., Li J., Liu Z.W., Liu J.P., Qi Q.H., Liu J.S., Li T., Wang X.G., Lu H., Wu T.T., Zhu M., Ni P.X., Han H., Dong W., Ren X.Y., Feng X.L., Cui P., Li X.R., Wang H., Xu X., Zhai W.X., Xu Z., Zhang J.S., He S.J., Zhang J.G., Xu J.C., Zhang K.L., Zheng X.W., Dong J.H., Zeng W.Y., Tao L., Ye J., Tan J., Ren X.D., Chen X.W., He J., Liu D.F., Tian W., Tian C.G., Xia H.G., Bao Q.Y., Li G., Gao H., Cao T., Zhao W.M., Li P., Chen W., Wang X.D., Zhang Y., Hu J.F., Liu S., Yang J., Zhang G.Y., Xiong Y.Q., Li Z.J., Mao L., Zhou C.S., Zhu Z., Chen R.S., Hao B.L., Zheng W.M., Chen S.Y., Guo W., Li G.J., Liu S.Q., Tao M., Zhu L.H., Yuan L.P., Yang H.M. (2002) A draft sequence of the rice genome (Oryza sativa L. ssp indica). Science 296:79-92.

Zeng Z.B. (1994) Precision Mapping of Quantitative Trait Loci. Genetics 136:1457-1468.

Zhao H.H., Fernando R.L., Dekkers J.C.M. (2007) Power and precision of alternate methods for linkage disequilibrium mapping of quantitative trait loci. Genetics 175:1975-1986. DOI: 10.1534/genetics.106.066480.

Table 1.1. Anchor markers for RAD-only map construction. Nine morphological marker loci used for linkage map construction, together with RAD loci, in the Oregon Wolfe Barley mapping population, showing chromosome assignments, phenotypes, And genes (if known).

Locus	Gene	Chromosome	Phenotype	Gene
VRS-1	*HvHox1*	2H	Two-row inflorescence (*Vrs1Vrs1*)/six-row inflorescence (*vrs1vrs1*)	Kao et al. (1999) GenBank:[AB489122.1]
ZEO-1	NA	2H	Dwarf plant with compact head (*Zeo1*)/normal height and head length (*zeo1*)	
ALM	NA	3H	Green lemma and nodes (*Alm*) /albino lemma and nodes (*alm*)	
HSH	NA	4H	Hairs on lower leaf sheaths (*Hsh*) / lack of hair on lower leaf sheaths (*hsh*)	
SRH	NA	5H	Long hairs on rachilla (*Srh* /short hairs on rachilla (*srh*)	
ROB	NA	6H	Green lemma and nodes (*Rob*) /orange lemma and nodes (*rob*)	
WX	*GBSS-I*	7H	Wild type endosperm starch (*Wx*)/ waxy endosperm starch (*wx*)	Komatsuda et al. (2007) GenBank:[AF486518.1]
NUD	NA	7H	Hulled seed (*Nud*) /hulless seed (*nud*)	Patron et al. (2002) GenBank:[AP009567]
LKS2	NA	7H	Long awn (*Lks2*)/ short awn (*lks2*)	

Figure 1.1. Segregation distortion on chromosome 2H linkage maps in the Oregon Wolfe Barley mapping population. The results of mapping with two different data sets are shown in A) the OWB-2383 map + 463 RAD loci, B) the 436 RAD and morphological marker loci and C) the OWB-2383 map. The X axis represents map distance in cM and the Y axis represent -*log* of the χ^2 *p*-value for segregation distortion. A positive value means distortion in favor of OWB-D whereas a negative value means distortion in favor of OWB-R. Dashed lines represent significance thresholds at 0.05. Marker positions are represented as perpendicular lines to the X axis.

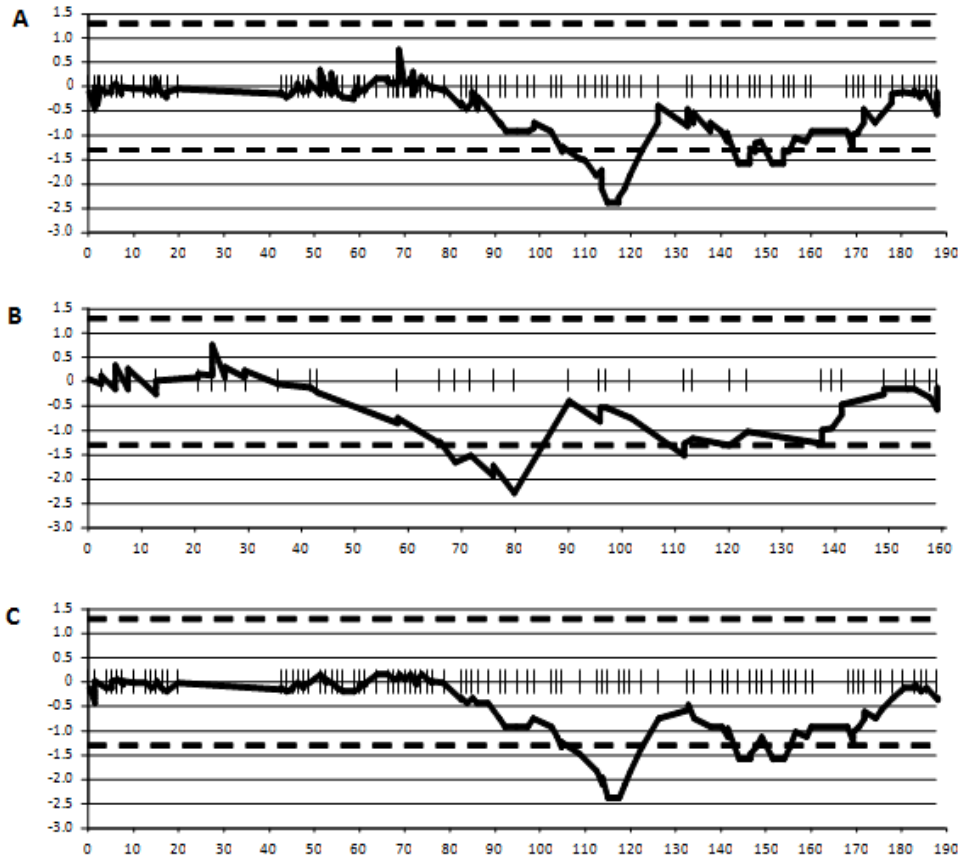


Table 1.2. Summary of chromosome length in three linkage maps. Lengths (in Haldane cM) for linkage groups corresponding to barley chromosomes 1H – 7H of the Oregon Wolfe Barley mapping population.

Linkage map	Chromosome							Total
	1H	2H	3H	4H	5H	6H	7H	
OWB-2383*	158	188	208	127	238	163	206	1288
OWB-2383 + 463RAD	158	188	208	127	238	163	204	1286
RAD only	175	158	228	123	226	134	216	1260

*The OWB-2383 map was reported by Szucs et al. (2009)

Figure 1.2. Macro-scale syntenic relationships between barley and *Brachypodium* revealed with sequenced RAD markers. RAD sequences anchored by linkage analysis are distributed across the seven *Hordeum* linkage groups. Alignments to orthologous sequence loci in *Brachypodium* are shown. Solid lines denote relationships supported by EST sequence comparison. Two dashed lines indicate sequence alignments that do not coincide with expected chromosomal relationships.

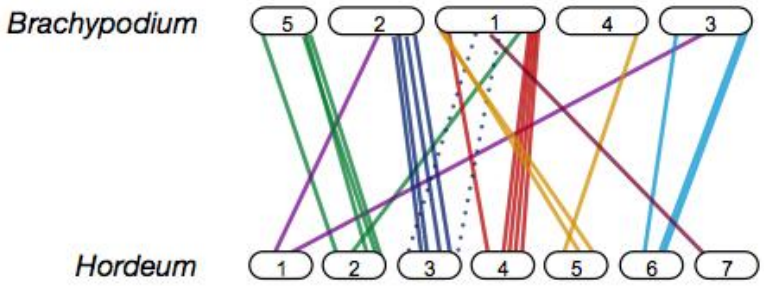


Table 1.3. Summary of QTL based on a skeleton map derived from the 2846 loci dataset. The QTLs in italics and underlined showed a trend in the RAD-only map but did not reach the LOD threshold.

Trait and QTL number	Chrom.	QTL Peak position	2-LOD confidence interval	Morphological and/or cloned genes within 2 LOD conf. interval	LOD	Additive effect	R^2	LOD Threshold	MIM R^2	H^2
Final leaf number								3.0	0.47	0.87
1	1H	158	(154-158)		6.9	-0.9	0.17			
2	4H	118	(117-120)	*VRN-H2*	4.6	0.7	0.10			
3	5H	157	(154-161)	*VRN-H1*	3.5	-0.7	0.09			
4	7H	29	(26-36)		3.6	0.7	0.08			
Plant Height								3.1	0.82	0.94
1	1H	131	(120-136)		7.1	-6.7	0.06			
2	2H	156	(155-158)	*ZEO1*	32.1	-22.8	0.67			
3	3H	51	(43-60)		6.1	6.3	0.05			
4	6H	100	(74-104)		6.0	5.7	0.05			
Spike number								2.9	0.50	0.63
1	2H	99	(95-102)	*VRS1*	13.3	3.5	0.38			
2	*5H*	*43*			*2.4*	*-1.3*	*0.05*			
3	*6H*	*80*			*2.7*	*1.3*	*0.06*			
Spike length								3.1	0.93	0.98
1	1H	158	(152-158)		6.7	-0.5	0.00			
2	2H	156	(155-159)	*ZEO1*	45.8	-3.0	0.81			
3	3H	20	(13-23)		5.0	-0.5	0.04			
4	5H	138	(137-144)		3.8	-0.3	0.02			
5	6H	95	(93-100)		4.5	0.4	0.02			
1x2						0.5	0.05			
Floret number								3.0	0.87	0.97
1	1H	156	(150-158)		9.4	-5.8	0.06			
2	2H	99	(97-101)	*VRS1*	40.3	-21.6	0.79			
3	*3H*	*36*			*2.9*	*-2.9*	*0.01*			
4	4H	120	(112-127)	*VRN-H2*	3.6	3.3	0.02			
5	6H	91	(73-103)		4.6	3.7	0.02			
Grain number								3.0	0.74	0.92
1	1H	151	(142-156)		3.9	-4.9	0.01			
2	2H	99	(99-101)	*VRS1*	20.1	-12.3	0.46			
3	2H	156	(150-163)	*ZEO1*	10.4	-7.3	0.18			
1x2						2.9	0.04			
2x3						3.4	0.05			
Hundred grain weight								2.9	0.66	0.78
1	2H	99	(97-101)	*VRS1*	19.9	0.5	0.53			
2	6H	60	(54-61)		3.5	-0.1	0.04			
3	7H	103	(96-111)		4.2	0.2	0.05			
Grain yield								3.1	0.29	0.49
1	1H	157	(153-158)		4.4	-2.9	0.11			
2	2H	160	(150-165)	*ZEO1*	8.1	-4,1	0.23			

Table 1.4. Summary of QTL based on RAD-only map constructed with 436 RAD and nine morphological loci. The QTLs in italic and underlined showed a trend in the full map but did not reach the LOD threshold.

Trait and QTL number	Chrom.	QTL Peak position	2-LOD confidence interval	Morphological and/or cloned genes within 2 LOD conf. interval	LOD	Additive effect	R^2	LOD Threshold	MIM R^2	H^2
Final leaf number								2.8	0.41	0.87
1	1H	175	(170-175)		6.2	-0.9	0.18			
2	4H	123	(120-123)	*VRN-H2*	2.9	0.6	0.08			
3	5H	154	(148-156)	*VRN-H1*	2.8	-0.6	0.07			
4	*7H*	*27*			2.7	0.6	0.07			
Plant Height								2.8	0.78	0.94
1	1H	144	(133-158)		6.5	-6.5	0.06			
2	2H	122	(119-128)	*ZEO1*	31.0	-21.7	0.64			
3	3H	57	(51-73)		5.6	6.0	0.05			
4	6H	112	(103-121)		5.7	5.9	0.05			
Spike number								2.7	0.56	0.63
1	2H	57	(50-64)	*VRS1*	11.4	3.6	0.40			
2	5H	9	(2-24)		3.8	-1.8	0.06			
3	6H	91	(85-98)		3.6	1.2	0.05			
1x2						-1.5	0.06			
Spike length								2.9	0.92	0.98
1	1H	175	(158-175)		4.9	-0.5	0.00			
2	2H	122	(120-127)	*ZEO1*	42.9	-3.1	0.82			
3	3H	20	(7-34)		5.1	-0.5	0.04			
4	*5H*	*105*			2.4	-0.2	0.01			
5	*6H*	*106*			2.7	0.3	0.02			
1x2						0.4	0.04			
Floret number								2.8	0.89	0.97
1	1H	175	(157-1175)		7.5	-5.3	0.05			
2	2H	57	(54-61)	*VRS1*	38.6	-21.4	0.77			
3	3H	39	(30-45)		3.2	-3.2	0.02			
4	4H	123	(117-123)		3.1	3.2	0.02			
5	6H	103	(66-120)		4.8	3.9	0.03			
Grain number								3.0	0.75	0.92
1	1H	169	(162-174)		3.8	-5.0	0.01			
2	2H	56	(53-62)	*VRS1*	18.7	-12.4	0.47			
3	2H	122	(116-130)	*ZEO1*	10.3	-7.2	0.18			
1x2						2.9	0.04			
2x3						3.3	0.05			
Hundred grain weight								2.8	0.66	0.78
1	2H	57	(54-62)	*VRS1*	19.8	0.5	0.54			
2	6H	66	(56-70)		3.5	-0.1	0.04			
3	7H	104	(86-130)		3.9	0.2	0.05			
Gain yield								2.8	0.30	0.49
1	1H	175	(170-175)		3.5	-2.8	0.11			
2	2H	122	(112-131)	*ZEO1*	6.2	-3.7	0.19			

FROM MARKER ASSISTED SELECTION TO GENOME-WIDE ASSOCIATION MAPPING: COUPLING GERMPLASM IMPROVEMENT AND GENE DISCOVERY IN FALL-SOWN FOOD BARLEY

Yada Chutimanitsakun, Alfonso Cuesta-Marcos, Shiaoman Chao, Ann Corey, Tanya Filichkin, Mathias Kolding, Caryn Ong, Juan Ignacio Rey, Andrew Ross, and Patrick M. Hayes

The Plant Genome

Crop Science Society of America
5585 Guilford Road, Madison, WI 53711-5801, USA

(To be submitted)

ABSTRACT

Barley (*Hordeum vulgare*) is an important component of heart-healthy whole grain diets because it supplies β-glucan. All current barley varieties with high β-glucan are spring habit, and these varieties have waxy starch. Fall-sown varieties have agronomic advantages but must have low temperature tolerance (LTT). Vernalization sensitivity (VS) is associated with higher levels of LTT. To rapidly develop fall-sown varieties with LTT and higher grain β-glucan, we used marker-assisted selection (MAS) at the *WX* and *VRN-H2* loci. The MAS-derived lines, together with unrelated non-waxy germplasm developed via phenotypic selection (PS) were used for genome-wide association mapping (GW-AM). The panel was phenotyped for grain β-glucan, LTT and VS. It was genotyped with 3072 single nucleotide polymorphisms (SNPs) and allele-specific primers. Marker-assisted selection fixed target alleles at both loci but only one of the target phenotypes (higher β-glucan percentage) was achieved. Continued segregation for VS and LTT is attributable to (i) incomplete information about *VRN-H1* at the outset of the project and (ii) unexpected allelic variation at *VRN-H3* with a large effect on VS and LTT. This new information will be essential in developing fall-sown food barley varieties. Additional research will be necessary to establish the functional basis of the associations. Genome-wide association mapping can be a more efficient QTL detection tool than biparental QTL mapping because it can be applied directly to any array of breeding germplasm segregating for the trait of interest. However, unless perfect markers based on functional polymorphisms of

determinant genes are used, follow-up research may be necessary to characterize favorable alleles.

INTRODUCTION

Barley was once a staple food in many parts of the world, but consumption declined as wheat (*Triticum aestivum*) and rice (*Oryza sativa*) enjoyed greater favor. There is renewed interest in barley as a food crop due to its nutritional properties, the most familiar and well-characterized of which is the soluble fiber β-glucan. The benefits of barley β-glucan consumption include lowering total and LDL cholesterol, blood sugar and the risk of cardiovascular disease and type II diabetes (Ames and Rhymer, 2008). The United States Food and Drug Administration (US-FDA) approved a health claim for barley in 2006 (FDA 2006). However, the amount of barley used in foods in the US has not changed appreciably over the past five years (USDA Feed Grains Database, available at http://www.ers.usda.gov/Data/FeedGrains/ [verified 16 Nov. 2011]). Causes include limited consumer awareness, a reluctance of food processors to launch new products, and limited availability of suitable varieties. One approach to variety development is based on maximizing β-glucan percentage. Recently, the USDA-ARS announced the development of barley germplasm with the potential to produce grain containing nearly 16% β-glucan (available at http://www.capitalpress.com/content/JOIdahoBarleyCommission-102811 [verified 16 Nov. 2011]). Barley and oats (*Avena sativa*) typically range from 3-11% β-glucan (Brennan and Cleary, 2005). An alternative approach is to develop varieties with normal, or slightly elevated, β-glucan content for use in whole-grain products. This strategy is based on the FDA health claim that specifies consumption of 0.75g of soluble fiber per serving. Two slices of bread based on flour containing 50% barley at

5% β-glucan would meet this requirement. Ultimately, producers, processors, and consumers will all benefit from a portfolio of barley varieties with different food quality attributes.

Effective breeding for β-glucan requires an understanding of its chemical and genetic basis. β-glucan is a soluble fiber and non-starch polysaccharide composed of mixed-linkage (1,3;1,4)-β-D-glucan. It is a major component of cell walls of the endosperm and aleurone. Although the chemical structure of β-glucan and its localization are well characterized, the complete gene network involved in synthesis is as yet uncharacterized. Therefore, MAS for β-glucan content *per se* is currently not feasible. An alternative approach is to capitalize on the positive pleiotropic effects of waxy starch on β-glucan (Xue et al., 1997). In the homozygous recessive condition alleles at the *Waxy* (*WX*) locus encoded by granule-bound starch synthase I (*GBSSI*) alter the amylose: amylopectin ratio. Higher amylopectin content in *wxwx* genotypes is due to a deletion in the promoter of *GBSSI* (Domon et al., 2002; Patron et al., 2002). Waxy starches are suitable for the extractive/ fractionation model in which the barley kernel is "mined" for β-glucan, starch, and protein, as waxy starches are used as thickeners in processed foods (Hung et al., 2006). Waxy starch can complicate whole-grain baking due to rapid water absorption (Tester and Morrison, 1990) and waxy starch may have a higher glycemic index (Akerberg et al., 1998). The promoter deletion in *GBSSI* provides a perfect marker for indirect selection for increased β-glucan.

A perfect marker for indirect selection for β-glucan, via starch type, will allow for rapid development of novel varieties. To date, there has been limited selection for fall-sown varieties with higher levels of β-glucan. Fall-sown barley is a relatively novel class that has advantages over spring-sown barley because it conserves scarce water resources and increases yield. In many target areas, including the Pacific Northwest of the US, winterhardiness is an essential attribute for fall-sown barley. Winterhardiness is a complex trait determined by a plant's ability to survive the many possible stresses encountered during the winter. Foremost among these attributes is LTT and two associated characteristics are VS and photoperiod sensitivity (PPDS) (Hayes et al., 1993). The roles of VS and PPDS in LTT relate to the timing of the vegetative to reproductive transition. Low temperature tolerance requires acclimation and maximum LTT is achieved in the vegetative stage. Once the plant transitions to reproductive growth, there is a rapid loss of LTT (Fowler et al., 1996). Vernalization sensitivity delays the transition until a sufficient input of low temperature is received (~6 weeks at 6^{o}C under controlled environment conditions) and PPDS delays the transition until daylength is sufficient (>10 h under controlled environment conditions) (Mahfoozi et al., 2000).

The genetic bases of LTT, VS, and PPDS in barley were recently reviewed (Galiba et al., 2009; Kosova et al., 2008; von Zitzewitz et al., 2011). Vernalization sensitivity is determined by three loci located on three different chromosomes: *VRN-H1* (5H), *VRN-H2* (4H) and *VRN-H3* (7H). As summarized by Trevaskis (2010) in a VS genotype under field conditions, once vernalization has occurred and daylength is

increasing *VRN-H1* down-regulates *VRN-H2*, which allows expression of *VRN-H3* and in return *VRN-H3* activates expression of *VRN-H1*. The result is a transition from the vegetative to the reproductive phase. *VRN-H1* therefore plays a central role in the vegetative to reproductive transition. Functional allelic variation at *VRN-H1* is reported to be due to deletions of varying lengths in a "critical region" within the first intron of *HvBM5A*. The "critical region" is hypothesized to contain a binding site for a repressor encoded by *VRN-H2* (Fu et al., 2005). Vernalization sensitivity genotypes are homozygous recessive and have an intact critical region. There are likely to be other regulators of *VRN-H1* besides *VRN-H2* (Dhillon et al., 2010; Dubcovsky et al., 2006; Trevaskis et al., 2006).

VRN-H2 is encoded by one or more of three zinc finger-CCT domain (*ZCCT-H*) transcription factors: *ZCCT-Ha*, *ZCCT-Hb* and *ZCCT-Hc* (Dubcovsky et al., 2005; Yan et al., 2004). The functional *Vrn-H2* allele (required for VS) encodes a dominant repressor of *VRN-H1* that is down-regulated by vernalization and short days (Yan et al., 2006). An alternative model suggested that *VRN-H2* inhibits flowering under long days by repressing *HvFT1* (*VRN-H3*). This would occur through direct mechanism with *HvFT1* or indirectly through interactions with other components of the photoperiod pathways such as *PPD-H1* (Trevaskis et al., 2007). Recessive alleles at *VRN-H2* are due to deletion of one or more *ZCCT* gene family members (Dubcovsky et al., 2005; Hemming et al., 2008). Homozygous recessive *VRN-H2* genotypes are not VS, regardless of their allele configuration at *VRN-H1* and/or *VRN-H3*. *VRN-H3* is the least understood of the vernalization genes. Takahashi and Yasuda, (1971) first

described *VRN-H3* and reported that most barley accessions are homozygous recessive (the allele state required for VS, if the accession is homozygous recessive at *VRN-H1* and has at least one dominant allele at *VRN-H2*). These authors reported that the rare dominant allele at *VRN-H3* leads to very early flowering (under long-day conditions). Takahashi and Yasuda's conclusions, based on phenotypic data obtained from an elegantly designed and executed series of crosses, are corroborated by recent gene expression studies (Yan et al., 2006). What remains to be determined is the functional basis of allelic variation at *HvFT1*. Yan et al. (2006) reported that allelic variation was due to a deletion in the first intron. Subsequently, Cuesta-Marcos et al. (2010) showed that polymorphisms at this intron position did not account for differences in growth habit and that nucleotide variation in the promoter was a more likely candidate. In summary, for VS, a perfect marker is available for *VRN-H2*; the deletion in *VRN-H1* provides, theoretically, a perfect marker, but the large number of possible deletions challenges rigorous characterization; and only polymorphisms of possible functional significance are known for *VRN-H3*.

Genes with key roles in photoperiod response are *PPD-H1* (2H) and *PPD-H2* (1H). *PPD-H1* is a principal inducer of flowering under long days, whereas *PPD-H2* affects flowering under short days. The dominant allele at *PPD-H1* promotes earlier flowering under long days and the recessive allele is conferred by a CCT domain mutation (Turner et al., 2005). At *PPD-H2*, functional alleles of *HvFT3* are dominant and confer insensitivity to short-day conditions (e.g. capability to flower). The recessive (short-day sensitive) allele is due to gene deletion or truncation (Faure et al.,

2007; Kikuchi et al., 2009). Perfect markers based on functional polymorphisms are therefore available for both long-day and short-day photoperiod sensitivity.

Genetic analysis of LTT has identified two major QTLs: *FR-H1* and *FR-H2*. These two loci are ~30 cM apart on chromosome 5H (Francia et al., 2004; Skinner et al., 2005). In barley, *FR-H1* co-segregates with *VRN-H1* (Stockinger et al., 2006) and coincidence is likely due to a pleiotropic effect of *VRN-H1* rather than to the effect of an unknown gene or genes (Dhillon et al., 2010). Therefore, indirect selection for *VRN-H1* is expected to be effective for *FR-H1* and therefore LTT. A cluster of C-repeat binding factor (*CBF)* genes coincides with *FR-H2* (Francia et al., 2007; Skinner et al., 2006). A causal relationship between one or more of these *CBF* genes and LTT has not been established. Therefore, tightly linked and/or associated markers for LTT are available at *FR-H2*, but not a perfect marker based on a functional polymorphism.

Growth habit - winter, spring, and facultative - is defined by LTT, VS and PPDS. Winter growth habit types are LTT, VS and vary in short-day PPDS. Spring growth habit types have minimal LTT and are not VS. They may have short-day PPDS without agronomic impact since they are grown under long days. Facultative refers to genotypes that are LTT, are not VS, and should have short-day PPDS in order to delay the vegetative to reproductive transition until long day conditions prevail. Facultative types are therefore equivalent to winter types, except that they have the *VRN-H2* deletion (recessive allele). A perfect marker for facultative growth habit is therefore available – the *VRN-H2* deletion. However, scoring a deletion has the drawback that the target allele is a null. Therefore, MAS will be more reliable if based on the

deletion plus tightly linked codominant markers, or preferably, a multi-locus haplotype.

In this project, we used prior information on the *WX*, *VRN-H1*, and *VRN-H2* loci to select parents for a MAS project to rapidly develop winter growth habit barley varieties (with the assumption that we would maximize LTT) with waxy starch (with the assumption that we would increase β-glucan percentage in the grain). The parental survey was conducted with allele-specific primers. In this way, we identified the winter varieties Luca and Strider as donors of the dominant allele at *VRN-H2* and LTT alleles at *FR-H1* and *FR-H2*. The spring varieties Waxbar and Merlin were identified as having target recessive alleles at *WX*, recessive alleles at *VRN-H2* and, according to the allele-specific primers used, a recessive allele at *VRN-H1*. We then initiated a MAS program, using these parents. Concurrently, in a different germplasm base, we had initiated PS for normal starch and the hull-less character using winter and facultative parents. The resulting lines, all at the F_4 or later generations, were available for GW-AM under the auspices of the U.S. Barley Coordinated Agricultural Project (CAP).

The goal of the Barley CAP is to genotype and phenotype germplasm in breeding programs for gene discovery and utilization in barley development. The interest in GW-AM was to increase the efficiency and relevance of QTL detection. A number of recent reports attest to its utility, and have revealed its drawbacks, using barley as a model (Cockram et al., 2008; Cuesta-Marcos et al., 2010; Hamblin et al., 2010; Roy et al., 2010; von Zitzewitz et al., 2011; Wang et al., 2010; Waugh et al.,

2009; Wang et al., 2011). We hypothesized that the MAS and PS germplasm would allow us to test the utility of GW-AM for detecting the genes we had targeted for MAS (validation) and would allow us to detect new genes determining the traits of interest (discovery). In this report we focus on three traits of primary importance to the development of fall-sown food barley: LTT, VS, and β-glucan percentage in the grain.

MATERIALS AND METHODS

Germplasm Development

The germplasm development process is shown in Fig. 2.1. The full germplasm array (*n* = 96) represents the final and fourth set of germplasm developed by the Oregon State University (OSU) barley program for the Barley CAP project in 2009 and is identified as CAP IV. Entries are numbered consecutively based on this designation, e.g. "09OR-01, 09OR-02", etc. All genotype and phenotype data are available at the *Triticeae* Toolbox (T3) (http://triticeaetoolbox.org [verified 6 Nov. 2011]). The CAP IV array represents a snapshot of advanced generation food barley germplasm that was available in the OSU program as of 2009. Per Barley CAP guidelines, all germplasm was at the F_4 generation or higher. CAP IV consists of two types of germplasm, based on derivation and starch type. MAS group (*n* = 64; 09OR-01 - 09OR-64) was developed in the BC_1F_1 and BC_1F_2 generations for the *waxy* (recessive) allele at *GBSSI* and the winter allele (dominant) at *VRN-H2*. The MAS lines were genotyped at the BC_1F_4 using allele-specific primers for validation. The MAS group consists of 54 two-row and 10 six-row barley populations. There were four parents of the MAS lines: Luca (two-row, normal starch, hulled, and winter growth habit), Merlin and Waxbar (two-row, waxy starch, hull-less, and spring/facultative growth habit), and Strider (six-row, normal starch, hulled, winter growth habit). Merlin and Waxbar are noted as facultative/spring growth habit because at the time the germplasm was developed all available information indicated that they were facultative. As described in the Results and Discussion, these two accessions are

now described as spring growth habit. The PS group (*n* = 30; 09OR-65 - 09OR-94) was developed by pedigree selection for the hull-less trait and adaptation to Pacific Northwest conditions. All parents and progeny are six-row types and have normal starch. Four parents developed the PS lines: Strider (described above), Doyce (hull-less, winter growth habit), Maja (hulled, facultative growth habit) and Legacy (hulled, spring growth habit). Merlin and Waxbar were included as 09OR-95 and 09OR-96, respectively.

Low temperature tolerance, grain β-glucan percentage and starch composition

Beginning in the fall of 2007, field evaluations were conducted over multiple years and locations for both the MAS and PS groups. Test sites included Corvallis, Oregon (CVO), Pendleton, Oregon (PEN), Hermiston, Oregon (HER), and Aberdeen, Idaho (ABER). In this report we focus on results from trials planted in the fall of 2009 at CVO and HER. At CVO and HER, the experimental design was a partially balanced lattice with three replications. Due to a planting error, one replication was dropped from the HER dataset. Plot sizes were 7 m^2 and consisted of six rows. The seeding rate was 100 kg ha^{-1}. Plots were seeded with a Hege cone planter. Fertility and weed management was in accordance with local practice. Corvallis is a high rainfall site. There were 1091 mm of total precipitation between planting in 2009 and harvest in 2010. Hermiston is a dryland site. There were 233 mm of total precipitation between planting and harvest. Irrigation was applied in accordance with local practice via a center-pivot system. Plots were harvested with a Wintersteiger Nursery Master combine. Data were recorded for a range of agronomic and disease resistance traits, all

of which are available at http://triticeaetoolbox.org [verified 4 Nov. 2011]. In this report we focus on two traits from the field trials: grain β-glucan percentage and LTT. Low temperature tolerance was determined based on the visual assessment of the number of surviving plants when plots resumed growth after exposure to low temperature during winter. Luca, Merlin, Waxbar, Strider, Charles (a winter malting check selected by the American Malting Barley Association (AMBA)) and Eight-twelve (the AMBA winter hardiness check) were included in the field trials. Legacy and Doyce were not included because of their hyper-susceptibility to barley stripe rust (incited by *Puccinia striiformis* fsp. *hordei*). This disease is endemic at CVO.

Seed samples from each line from the first replication of each trial were ground in a Cleanmill. The resulting flour was used to determine the mixed-linkage β-glucan percentage following the method of Hu and Burton (2008) using Megazyme kits (Megazyme International Ireland Ltd.). Briefly, samples of 80-120 mg of flour were mixed with ethanol and sodium phosphate buffer. The solutions were then incubated with lichenase for 1 h, after which the reaction was stopped with sodium acetate buffer. After adding β-glucosidase and GOPOD reagent, absorbance at 510 nm was measured using a spectrophotometer. β-glucan percentage was expressed on a dry-weight basis based on the McCleary method (McCleary and Codd, 1991). Based on samples from CVO, the Rapid Visco Analyser (RVA) was used to categorize waxy and non-waxy starches as a function of paste viscosity and a peak time (Newport Scientific, Pty. Ltd., Warriewood, NSW, Australia). Tests were run according to a shortened profile of Crosbie et al. (2002). The viscosity parameters obtained were

pasting temperature, peak viscosity, peak time, trough, breakdown and final viscosity. The peak time (time required to reach peak viscosity) is plotted against peak viscosity (maximum viscosity during heating stage) to determine starch type.

Vernalization sensitivity and flowering time

Vernalization sensitivity was measured under greenhouse (GH) conditions using unvernalized and vernalized treatments. For both treatments, there were two replications of each genotype. For the unvernalized treatment plants were sown directly to 1.65 L pots. Plants were grown in a GH maintained at 18 ± 1.5°C day and night, with a photoperiod of 16 h light/24 h provided by supplemental lighting with 400 W high pressure sodium light fixtures. For the vernalized treatment seeds were sown in moist soil and immediately transferred to a growth chamber set to 6°C. They were kept in a dark chamber for seven weeks before transplanting to the same GH conditions as the unvernalized treatment. Days to flowering were recorded as the number of days from emergence (unvernalized) or transplanting (vernalized) until anthesis. Anthesis was assessed as the moment when awns were first visible above the flag leaf sheath. The experiment was terminated 120 days after planting the unvernalized treatment. Plants that had not flowered were assigned a days to flowering value of 150.

Genotyping

The CAP IV set was genotyped with 3072 SNP markers using Illumina GoldenGate Bead Array technology. The SNPs were designed from EST sequences and PCR amplicons, and were organized into two Oligonucleotide Pool Assays

(OPAs) known as Barley OPA1 and OPA2, as previously described in Close et al. (2009). In addition, the CAP IV set was genotyped for the *VRN-H1*, *VRN-H2*, *VRN-H3*, *PPD-H1*, and *PPD-H2* loci using allele-specific assays, as described in von Zitzewitz et al. (2011), and for *GBSSI* using allele specific primers described in Domon et al. (2002). Later, we developed an additional primer set of HvFT1.13F (5'CACCACGTCCCAAGAGTTTTC3') and HvFT1.15R (5'GCGTACAACATCCA CAGTCC3') and product digestion with *AciI* to differentiate alleles at *VRN-H3*.

Association analysis

Genome-wide association mapping was conducted with TASSEL V.3 (Bradbury et al., 2007) using three datasets i) the entire CAP IV set, ii) the MAS population only, and iii) the PS population only. In addition to 09OR-01 - 09OR-96, Luca, Strider and Eight-twelve were included in the analysis for all three datasets. Genotypic data of all three were obtained from the Barley OPA1 and 2 genotyping of Oregon CAP 2006 (available at http://triticeaetoolbox.org [verified 5 Nov. 2011]). A linear mixed model was used to test for association between markers and traits of interest. Minor allele frequencies (MAF) lower than 0.05 were excluded. For the full CAP IV set, the population structure was determined using STRUCTURE 2.2 (Pritchard et al., 2000) with the linkage model described by Falush et al. (2003).The estimated logarithm of the probability lnPr(X|K) was determined between K = 2-5 with 25,000 burn in and 25,000 MCMC iterations. The most probable number of subpopulations was selected, and the program was run again with 500,000 MCMC iterations. Principle component analysis (PCA) was performed to determine

population structure for the MAS dataset using TASSEL V.3. The kinship (K) matrix was generated with TASSEL V.3 and was used to correct for relatedness of individuals. Only the K matrix was included in the model for the PS dataset. For each GW-AM analysis, the cut-off *p*-value that controls the false discovery rate at α=0.05 was calculated using SAS PROC MULTTEST (SAS v9.2) and the method of Benjamini and Hochberg (1995).

Two-way interactions between significant markers were tested with SAS PROC MIXED (SAS v9.2) using the restricted maximum likelihood method. The Q matrix was included as a covariate and the K matrix was the variance-covariance matrix of the random term (genotype). Using the same model we also estimated the effects of individual markers. In order to determine the degree of linkage disequilibrium (LD) on chromosome 5H in the vicinity of *FR-H1*, an LD heat map (Shin et al., 2006) was generated with markers having a *p*-value below 0.05 for LTT at HER. The LD plot was created with the snp.plotter package implementation for the R software environment (Luna and Nicodemus, 2007).

RESULTS AND DISCUSSION

Allele specific assays in the BC_1F_4 and F_4 generations

Lines in the MAS group were genotyped in the BC_1F_4 generation to confirm the presence of target alleles at the *VRN-H2* locus (dominant alleles) and *WX* locus (recessive alleles). Of the 64 MAS-derived lines, all were homozygous dominant at *VRN-H2* and 63 were homozygous recessive at *WX*. 09OR-16 is homozygous dominant for *WX*, indicating normal starch. This is due to an error in genotyping, selection, and/or generation advance because records indicate that at the BC_1F_2 the antecedent of 09OR-16 was homozygous *wxwx*. The PS group was genotyped at the F_4 generation for *VRN-H1*, *VRN-H2*, and *WX*. All progeny are *WxWx* (normal starch) and have winter alleles (according to the primers used) at *VRN-H1*. At *VRN-H2*, 2 of 30 lines (09OR-65 and 09OR-68) are homozygous for the *VRN-H2* deletion (spring allele). The facultative (Maja) and spring (Legacy) parents both have the *VRN-H2* deletion.

Validation of MAS for *wxwx* based on starch viscosity

Using samples from CVO, the RVA analysis identified two principal groups that correspond to the waxy (MAS) and normal starch (PS) germplasm (Fig. 2.2). The peak times of the waxy starch group were lower than those of the normal starch group, in agreement with the results of Kim et al. (2003) who reported that the rapid swelling of waxy starch is due to the high amylopectin content. There were two lines (09OR-18 and 09OR-21) with intermediate peak times. Both are homozygous *wxwx*, indicating that they have a waxy starch but have lower amylopectin/higher amylose content.

Homozygous *wxwx* lines with intermediate starch properties were also reported by Box et al. (2003). 09OR-16 was in the normal starch group, confirming that this line is indeed *WxWx*. It is faster to measure grain β-glucan content than starch viscosity, and the former is the principal trait currently targeted by the barley foods industry. However, the RVA is useful for identifying lines, such as 09OR-18 and 09OR-21 with novel starch properties that may be useful for specific food processing applications.

Genome-wide association analysis

The GW-AM was performed first on the full CAP IV dataset (n = 96) plus checks (n = 3; Luca, Strider, and Eight-twelve). The results of the structure analysis revealed four subpopulations (K = 4). Then, when appropriate, we performed separate analyses on the two groups: MAS and PS. For the MAS group we accounted for structure using the first three principal components of the PCA analysis. The MAS group consists of n = 64 lines and 5 checks (Luca, Strider, Eight-twelve, Waxbar, and Merlin). For the PS group there was no population structure detected, thus no correction was needed. The PS group consists of 30 lines and the same 5 checks as for MAS.

Grain β-glucan percentage

The phenotypic distributions for grain β-glucan at CVO and HER were similar, bimodal, and indicative of phenotypic transgressive segregation (Fig. 2.3). Average grain β-glucan values were significantly (p<0.0001) higher at HER than at CVO. At HER and CVO respectively, grain β-glucan percentage ranged from 5.0-7.0% and 4.1-6.3% in *wxwx* types. For *WxWx* types, the grain β-glucan percentages ranged from

3.5–5.0% and 3.0-4.5%. The effect of environment on grain β-glucan has been reported (Zhang et al., 2002), with hot and dry conditions during grain filling favoring higher β-glucan (Ehrenbergerova et al., 2008). Neither moisture nor temperature can account for the differences we observed: available moisture for plant growth was comparable at HER and CVO due to the use of supplemental irrigation at the former site. Average daytime temperatures during grain filling (January-May) were 8.3^{o}C and 8.7^{o}C for HER and CVO, respectively.

On average, grain β-glucan percentage in *wxwx* types was significantly (p<0.0001) higher than in normal starch types. This result confirms that selection for *wxwx* alleles will increase grain β-glucan percentage, as was previously reported by Xue et al. (1997). The bimodal distributions indicate segregation of alleles at a single locus and the transgressive segregation within each of the bimodal classes could be indicative of the effects of minor genes and/or experimental error. As shown in the GW-AM scans (Fig. 2.4 and supplemental Fig. 1), there were significant associations of markers on chromosome 7H, with grain β-glucan at both locations. These significant markers include a perfect marker within the *GBSSI* gene and SNPs in LD with *GBSSI*. For the MAS lines, ~4 cM from *GBSSI*, there was a strong association (p = 0.001) for marker 2_0227. However, the p value of this marker is larger than the cut-off p-value that controls the false discovery rate at 0.05. The significance of this marker is due to its LD with *GBSSI* and its MAF is greater than the 5% cut-off. There were nine individuals (the two non-waxy parents, 09OR-16 and six other *wxwx* CAP lines) within one of the two possible allele classes for this SNP. The failure to detect

other significant associations driving the phenotypic variation within the *WxWx* and *wxwx* classes is likely due to small population size. Although a direct calculation of error rates is not possible in these analyses due to a lack of replication, the highly significant difference between the *WxWx* and *wxwx* classes at both HER and CVO indicates a high degree of repeatability of the β-glucan assay. Furthermore, in other experiments involving this germplasm the standard errors for grain β-glucan in *WxWx* and *wxwx* genotypes were 0.2% and 0.1%, respectively (data not shown).

Low temperature tolerance in the CAP IV panel

All CAP IV entries, including Merlin and Waxbar, showed complete winter survival at CVO. The lowest minimum temperature was 2.3^{o}C and occurred in December 2009 (Supplemental Table 1). Therefore, LTT was evaluated only at HER where the temperature averaged ~0^{o}C for 1 month and the minimum temperature (without snow cover) was -13^{o}C. Considering the entire CAP IV set, the percent survival ranged from 0 – 100% (Fig. 2.5). Merlin and Waxbar were in the 0% group. Luca, Strider, and Eight-twelve had ≥ 95% survival. All PS lines were in the high survival group. Survival in the MAS lines ranged from 0 – 100%. Based on these phenotype results, several points can be made: (i) Merlin and Waxbar are not facultative, as they showed no LTT, (ii) MAS for *VRN-H2* was not sufficient to achieve LTT, and (iii) all PS progeny had LTT equivalent to the parents and Eight-twelve, the LTT check. These phenotypic results can be explained, in part, by the results of GW-AM.

Considering all CAP IV lines, there were significant marker-trait associations on chromosomes 5H and 7H and a strong association on 6H (Fig. 2.6). The 5H and 7H associations are with *VRN-H1* and *VRN-H3*, respectively. On 5H, the most significant markers (3_0590 and 1_1080) are 5 and 1 cM from *VRN-H1*, respectively, according to the consensus map (Close et al., 2009). The most significant marker on 7H was 1_0056, which is ~3cM from *HvFT1* (*VRN-H3*). On 6H, 1_1455 approached the false discovery threshold. There are no genes associated with LTT reported in this region.

***FR-H1* and defining growth habit in Merlin and Waxbar**

Frost Resistance-H1 (*FR-H1*) (most likely a pleiotropic effect of *VRN-H1*) is the most commonly detected determinant of LTT in barley (reviewed by von Zitzewitz et al., 2011). In the CAP IV we did not expect to find a significant effect for *FR-H1* because (i) the winter and facultative parents (Luca, Strider, and Doyce) all have similar levels of LTT and winter alleles at *VRN-H1*, and (ii) our initial genotyping of Waxbar and Merlin led us to believe that they were facultative. We had confirmed that they were *VRN-H2* deletions, using allele-specific primers, and we determined that they had a "winter" allele at *VRN-H1* using three sets of allele-specific primers: BM5.88F/89R, BM5.42F/56R and BM5.42F/86R (Szucs et al., 2006; Szucs et al., 2007). These primers indicated that Waxbar and Merlin were the same as Luca, Strider and Doyce but distinct from Legacy. We therefore selected Waxbar and Merlin, out of all available waxy type barley varieties, due to the presence of this "winter" allele since it meant that *VRN-H1* could be ignored for MAS. However, in view of the phenotype results, Merlin and Waxbar do not meet an essential criterion for facultative

growth habit, which is LTT. There are at least ten different *VRN-H1* alleles described in the literature (Hemming et al., 2009) and these are specifically related to VS, not LTT. Several key issues need to be addressed before rigorously defining facultative growth habit at the phenotype and molecular levels. These include confirmation that *FR-H1* is a pleiotropic effect of *HvBM5A* and, if this is true, determining the functional basis of LTT within *FR-H1*. The intron 1 allele-specific markers in *VRN-H1* do not show significant associations with LTT in this study: the associations are with SNPs in LD with *VRN-H1*.

The two significant SNPs show different patterns of association. For SNP 1_1080, Legacy, Merlin, and Waxbar (all with poor LTT) have the B allele and the average survival of the 29 lines with this allele is 56%. Strider, Luca, Maja and Eight-twelve have the A allele and the 62 lines with this allele have an average survival of 90%. At SNP 3_0590 lines with the same allele (B) as Luca, Strider, Doyce, Legacy, Eight-twelve and Maja had an average of 90% survival whereas those with the Merlin, Waxbar allele had an average survival of 43%, providing further evidence that Merlin and Waxbar are not facultative types. In the LD heat plot (Fig. 2.7) for the ~10cM interval on chromosome 5H where *VRN-H1* is located we found that the two significant SNP markers are in the same LD block as *VRN-H1*. However, the correlation between the two markers is low, suggesting that they may be in LD with different functional polymorphisms in *VRN-H1* (or, if different than *VRN-H1*, the real determinant of LTT). In summary, our findings point to a role for *VRN-H1* and/or linked genes in LTT but they do not provide new insight on the functional

polymorphism(s) leading to enhanced LTT. They underscore the need to use perfect markers based on functional polymorphisms for MAS whenever possible. Finally, the LTT phenotype values of the different allele classes point to Merlin and Waxbar as donors of unfavorable alleles for LTT at *FR-H1*. That MAS lines with Merlin or Waxbar alleles at *FR-H1* would show lower LTT values is not unexpected, given the performance of the parents.

A potential new factor in low temperature tolerance: a significant interaction between *FR-H1* and *PPD-H1*

We found a significant interaction between 3_0590 (in LD with *VRN-H1)* and 3_0871 (a SNP within the photoperiod response gene, *PPD-H1*). There is an important role of PPDS in the mechanism of action of the vernalization genes (Trevaskis et al., 2006). Even though the main effect of *PPD-H1* was not detected in the GW-AM for LTT, we tested the interaction between *PPD-H1* and the significant markers at *VRN-H1* (3_0590 and 1_1080) and *VRN-H3* (1_0056). As shown in Cuesta-Marcos et al. (2010) markers involved in epistatic interactions may not be significant when performing single marker analysis. If lines with the unfavorable A allele (Merlin/Waxbar) at 3_0590 have a B (favorable) allele at *PPD-H1*, the average survival is 61%, as compared to 21% for lines with unfavorable alleles at both loci (AA individuals). The survival percentages for the BB and BA classes were 91% and 93% respectively. These results reveal a possible new role for *PPD-H1* in LTT. To date, allelic variation at this locus has focused on the role of the mutant allele in delaying flowering time (FT) under long-day conditions of spring and summer (Turner

et al., 2005), whereas low temperature injury at the vegetative stage is most likely to occur during the short day conditions of winter.

No effects for *FR-H2*

We found no effect of *FR-H2* on LTT in this germplasm, suggesting that favorable alleles are fixed, even in germplasm with poor LTT. There is precedent in the facultative x spring (Dicktoo × Morex) mapping population where only *FR-H1* is a significant determinant of LTT (Skinner et al., 2006). Von Zitzewitz et al. (2011) found both *FR-H1* and *FR-H2* to be associated with variation in LTT using CAP I and II germplasm. In that analysis, accessions with favorable alleles at *FR-H2* had an average survival of 14%, as compared to 30% for those with *FR-H1* only. Accessions with favorable alleles at both loci had an average survival of 64%.

Another new factor in low temperature tolerance: *VRN-H3*

We found significant associations of markers in, and near, *VRN-H3* with LTT (Fig. 2.6). There are no prior reports implicating *VRN-H3* in LTT. The two significant markers are 3_0893, which represents a SNP in the intron at nucleotide position 264 in *HvFT1* and 1_0056, which is 2.6 cM from 3_0893 and is in high LD with 3_0893. The latter marker was more significant than the former, presaging the challenge that the current data pose in terms of understanding how *VRN-H3* is related to LTT. At SNP 1_0056, Merlin, Waxbar, Strider, Doyce, Legacy, Maja and Eight-twelve have the same allele (B) and this allele is associated with high survival (91%) in 71 CAP IV lines. Luca has the unfavorable (A) allele and this allele is associated with lower survival (50%) in the 18 CAP IV lines in which it is present. Luca itself showed

100% survival and Legacy, while not in the HER trials, is a spring variety with very low survival in prior tests (data not shown). For marker 3_0893, Merlin, Waxbar, Strider, Legacy, Maja and Eight-twelve have the same allele (B), with average survival of 91% in 67 CAP IV lines. Luca and Doyce have the unfavorable (A) allele and this allele is associated with lower survival (53%) in the 17 CAP IV lines. However, Luca and Doyce have excellent LTT. The data from these two markers would suggest that Luca probably contributes an unfavorable allele for LTT even though the variety itself has good LTT. In order to try to understand these results in the context of *HvFT1* gene function, we genotyped the two putative functional polymorphisms in *HvFT1* in the CAP IV panel and consider these findings in relation to the SNP data and the reported function of *VRN-H3*.

Yan et al. (2006) reported that the functional domain in *HvFT1* is in intron I. This domain is represented as SNP 3_0895 and this marker was not significant in the GW-AM. To confirm this lack of significance, we used allele-specific primers (HvFT1.03F and HvFT1.04R and product digestion with *BclI*) (Yan et al., 2006) for this region to genotype the CAP IV and relate it to LTT. Two allele classes were identified: Merlin and Luca have the same allele and the 39 CAP lines with this allele had an average survival of 77%. Waxbar and Strider had the same allele and the 53 CAP IV lines had an average survival of 83%. Therefore, this polymorphism is not predictive, or causal, of the phenotypic variation in LTT observed in the CAP IV. An alternative functional polymorphism is in the promoter at nucleotide position 927 (Cuesta-Marcos et al., 2010). Therefore we re-sequenced the parental panel and

developed the primer set of HvFT1.13F (see Materials and Methods) to genotype the CAP IV. This primer set, due to the differences in restriction sites between parents, generated a total of four amplicons. Accordingly, these data were converted into four biallelic markers and their effects were tested using SAS PROC MIXED with the REML method. The results showed a strong effect of the unique allele found in Luca, with an average survival of 52% (11 individuals) compared with the average survival of the individuals with either of the three other alleles (83%, 87 individuals). All available data from three loci in and near *VRN-H3* point to Luca contributing the unfavorable allele at this locus. Considering the reported function of *HvFT1* in terms of VS and FT under long days (Trevaskis, 2010; Yan et al., 2006), this allele: LTT phenotype pattern challenges interpretation.

The 1_1455 marker: LTT phenotype association on 6H approached the significance threshold for the full CAP IV panel but does not approach significance in either the MAS or PS groups. There are no prior reports of LTT-related genes or QTLs on 6H. There were no significant interactions between markers on chromosomes 5H, 6H and 7H for LTT. Additional experiments will be necessary to determine if the 6H LTT effect is real, and if so, how it impacts LTT.

Genome-wide association mapping of low temperature tolerance in the marker-assisted selection and phenotypic selection subsets

Considering the phenotypic frequency distribution for LTT (Fig. 2.5), all PS lines showed good survival (95-100%) and there was tremendous variation (0-100%) within the MAS germplasm. We therefore conducted GW-AM on the two subsets. As

expected, there were no associations in the PS subset. In the MAS subset there were associations on 5H and 7H that mirrored the full CAP IV set. Therefore, future research into the role of the 7H LTT effect, and the 6H effect if warranted, should focus on the MAS subset.

Vernalization sensitivity and flowering time

Without vernalization, there was a clear distinction between lines that did not flower and those that flowered (Fig. 2.8). Luca and Strider were in the non-flowering (VS) group. Merlin and Waxbar did not require vernalization to flower (FT = 46 and 64 days, respectively). All except two of the PS lines (09OR-65, 09OR-68) were VS. Twelve of the MAS lines were VS and the remainder (50 lines) showed a range of FT, from 31 to 91 days. These results can be explained by the results of GW-AM in the case of *VRN-H2*, but in the case of *VRN-H3* the GW-AM raises more questions than it answers.

Considering all CAP IV lines, there were significant marker: trait associations, on chromosome 7H (Fig. 2.9). The significant markers were in the first intron of *HvFT1*: 3_0894 and 3_0895, at positions 471 and 585, respectively. There was a strong association at 2_0119 (4H), which is 20 cM distal to *VRN-H2*. Given the distance between the two markers it is not likely that this marker is in LD with *VRN-H2*. The annotation of this SNP (encoded catalase-1 protein) does not directly implicate a role in flowering (HarvEST database, available at http://harvest.ucr.edu/HBarley178.exe [verified 7 Nov. 2011]). There are no reports of VS or FT genes in this region of the barley genome. The subsequent discussion is

based on GW-AM within the PS and MAS groups, where a *VRN-H2* associated effect is unique to the PS group and the *VRN-H3* effect is unique to the MAS group.

In the PS group, the GW-AM scan shows a significant association between *VRN-H2* and VS. This can be explained by the deletion of *VRN-H2* in 09OR-65, 09OR-68, Waxbar and Merlin. There was no significant association of *VRN-H3* markers with VS within the PS group. All PS lines are monomorphic for *VRN-H3* markers, except for 09OR-68, which showed a polymorphism at 3_0894. This line is not VS.

The GW-AM scan of the MAS group (Fig. 2.9) showed a significant association at *VRN-H3*, with the most significant marker defined by the primers HvFT1.03F/04R+*BclI* (Yan et al., 2006). As mentioned previously in the section on LTT, these primers amplify the putative functional polymorphism reported by Yan et al. (2006) and target the same polymorphism as 3_0895. Marker 3_0895 showed a significant association in the full CAP IV set and did not approach the false discovery threshold in the MAS population. Merlin and Luca have the same allele at 3_0895 and the 38 lines with this allele had an average days to flowering, without vernalization, of 65 with a range of 39 to 150. Waxbar and Strider had the same allele and the 17 lines with this allele had an average days to flowering, without vernalization, of 97 and a range of 50 to 150. Therefore allele type at this marker does not predict VS and the association is likely due to LD with the functional determinant of VS, being in or near *VRN-H3*.

Cuesta-Marcos et al. (2010) reported that a polymorphism at position 927 (SNP927) in the promoter of *VRN-H3* differentiated phenotypes with dominant (T) and recessive (C) alleles in a limited set of genotypes with known *VRN-H3* alleles (e.g. Dairokaku, Tammi, Strider, Luca, Kompolti, Cali-sib, OWB-Dom and Dicktoo). However, when SNP927 was characterized on a broader array of germplasm, the T and C alleles did not show any association with flowering time (Cuesta-Marcos et al., 2010). Casas et al. (2011) assayed the C/T polymorphism in a set of Spanish barley cultivars and found no relationship with VS. Sequence alignment of the MAS parents (Supplemental Fig. 2) revealed that Waxbar has the T allele at SNP927 whereas Luca, Strider and Merlin have the C allele. At position 928, Merlin contains T while all other genotypes have the C allele. Therefore, considering the haplotype defined by SNPs 927 and 928, Waxbar and Merlin can be differentiated from Luca and Strider. This classification corresponds to the observed growth habits of the MAS parents but the association with VS in the MAS population cannot be determined until information is generated from the full population for allele type at SNP928.

Although the 927+928 haplotype corresponds to VS in the parents, neither Merlin nor Waxbar are exceptionally early to flower. In fact, Waxbar is nearly two weeks later to flower than normal spring varieties. The effect of the dominant *VRN-H3* allele is very early flowering (Takahashi and Yasuda, 1971; Yan et al., 2006). Therefore, we hypothesize that Merlin and Waxbar have recessive alleles at *VRN-H3* and that the significant association of *VRN-H3* with VS is due to the effects of these uncharacterized recessive alleles. Experiments are underway to test this hypothesis.

Allelic variation at *VRN-H3* would explain the observed segregation for VS in the MAS population: 12 of the 62 MAS lines are VS. This fits a 3:1 ratio ($p = 0.3$), which could be explained by the epistatic interaction of two loci: *VRN-H1* and *VRN-H3*. *VRN-H2* is not included in this model since we have established that the dominant alleles (VS competent) at this locus are fixed in the MAS population.

This model assumes that Waxbar and Merlin have uncharacterized spring (dominant) alleles at *VRN-H1*, an assumption supported by multi-locus haplotypes for markers near *VRN-H1* (Supplemental Table. 2), where Merlin and Waxbar have the same haplotype as Legacy (which has a spring dominant *Vrn-H1* allele) and differ from all winter types. To validate this two-locus interaction, we tested for significant two-locus interactions between all *VRN-H1* and *VRN-H3* markers that had significant main effects for VS and/or LTT. There was a significant interaction ($p<0.02$) between 3_0590 and 3_0894 (*VRN-H3*). However, 3_0590 is ~5 cM from *VRN-H1* and at this locus Merlin and Waxbar are different from Legacy, which has the same allele as the winter genotypes. A key supporting piece of evidence for *VRN-H1* having a role in this interaction is lacking: a significant main effect association with VS. However, detecting main effects for each of the two QTLs involved in an epistatic interaction may not be possible using GW-AM in a small population, as reported by Cuesta-Marcos et al. (2010).

Perspectives on *VRN-H3*, low temperature tolerance and vernalization sensitivity

Our results reveal an unexpectedly important role for *VRN-H3* (or tightly linked gene/genes) in LTT and VS. *VRN-H3* has not been implicated in LTT and its

role in VS was reported to be limited to very exotic germplasm. In terms of LTT, all available data suggest that Luca confers the unfavorable allele at *VRN-H3*. With the current data, it is not possible to explain this result. Intuitively, one would expect Merlin and Waxbar to be the donors of unfavorable alleles for LTT. Concerning VS, Takahashi and Yasuda (1971) reported that most varieties (VS and non-VS) have the recessive allele at *VRN-H3* and that the dominant allele is limited to exotic germplasm. More recently, the key role of *HvFT1* in flowering time has been elucidated in a series of elegant reports and this research has confirmed that the phenotype conferred by the dominant allele at *VRN-H3* is exceptionally early flowering. None of the parents of the MAS germplasm are exceptionally early flowering, yet one or more carry an allele associated with VS. In the case of VS, as with LTT, it is not possible to explain this result with the current data. Genome-wide association mapping has afforded an exciting new perspective on *VRN-H3* but further research will be necessary to determine the source(s) of the novel alleles. This will be straightforward with biparental populations and the resulting data should allow for building a model explaining the coincidence of VS and LTT QTLs at *VRN-H3*. The presence of non-VS lines with excellent LTT in both the MAS and PT shows that it is not a simple cause: effect relationship in which a lack of VS leads to a precocious vegetative to reproductive transition and a concomitant loss of LTT.

Flowering time

With vernalization, the phenotypic distributions for FT were similar in both the MAS and PS groups, although the latter showed a narrow range of variation (Fig. 2.8).

Marker-assisted selection lines flowered between 37-72 days, whereas PS lines flowered within 52 days. The average FT was not different, with 47 and 48 days in MAS and PS groups, respectively. This phenotypic variation can be explained more satisfactorily by GW-AM than VS.

Considering the full dataset, there was a significant association of 2_0394 (~1 cM from *PPD-H1*) with FT (Fig. 2.10). Two SNPs (3_0871 and 3_0872) in the CCT domain of the *PRR7* gene and an allele-specific assay for *PPD-H1* based on primers on *PRR7* (PRR7.5F/08R + *BstUI*) (Faure et al., 2007) were also significant, but *p* values were slightly higher than for 2_0394. At SNP 2_0394 Luca contributed the early flowering (A) allele in 27 lines with an average FT of 42 days, compared to 51 days in 61 lines with the B allele (Waxbar, Merlin, Strider, Maja, Eight-twelve and Doyce). Considering the functional polymorphism in *PPD-H1*, Luca, Strider, Maja and Doyce are long-day sensitive, with the dominant allele (*Ppd-H1*) whereas Merlin, Waxbar and Legacy are long-day insensitive, with the recessive allele (*ppd-H1*). On average, *Ppd-H1* lines flowered 7 days earlier than *ppd-H1* lines. These results confirm that association mapping based on perfect markers (e.g. 3_0871 in *PRR7*) generates allele type classifications for germplasm that are biologically meaningful whereas germplasm classifications based on allele type at markers in LD with functional polymorphisms can be misleading (e.g. 2_0394). The significant associations of markers in *PPD-H1* and in LD with *PPD-H1* in the full CAP IV set would be expected to be due to segregation in the MAS group, since the only parent in

the PS group with the *ppd-H1* allele is Legacy and no PS line inherited the Legacy allele.

The most significant association in the GW-AM of the MAS group was with 2_0394, as in the full CAP IV set. Classification of alleles based on this marker does not correspond to the known allele states at *PPD-H1,* e.g. Luca (*Ppd-H1*) is correctly differentiated from Waxbar and Merlin (*ppd-H1*), but Strider (*Ppd-H1*) has the same 2_0394 allele as Merlin and Waxbar. On average, lines with the Luca allele at 2_0394 were 6 days earlier than lines with the alternative allele.

In these experiments, plants were grown only under long-day conditions, and as expected there were no significant associations of FT with *PPD-H2*. Allele specific assays for *PPD-H2* (HvFT3.F4/R1, HvFT3.F1/R1 and HvFT3.01F/02R) (Faure et al., 2007) indicated allelic variation is present in both the MAS and PS populations. Further experiments under short-day conditions will be necessary to determine the main effects and possible interactions of *PPD-H2* in this germplasm.

Conclusions

In summary, we demonstrated successful MAS for target alleles at the *WX* and *VRN-H2* loci. The selection for *wxwx* was effective in raising grain β-glucan percentage, as expected. Fixation of dominant alleles at *VRN-H2* did not lead to VS and high LTT as expected. This was due to unexpected variation at *VRN-H1*/*FR-H1* and *VRN-H3*. In the case of *VRN-H1*, our limited characterization of the parents led us to an erroneous conclusion regarding Merlin and Waxbar having favorable alleles at *FR-H1*. Prior to initiating a MAS project, every effort should be made to gain a

complete picture of the structure and function of the target loci. Unfortunately, this may not always be possible. As yet, the basis of differences in LTT ascribed to alternative alleles at *FR-H1* is not understood. In the case of *VRN-H3*, we discovered unexpected allelic variation with effects on both LTT and VS. Novel alleles with dramatic effects may be lurking in parental germplasm and their effects will only be manifested when placed in new genetic backgrounds. Genome-wide association mapping was very useful for validating the effects of perfect markers (e.g. *WX*) and revealing the genome locations of determinants of target traits (e.g. *FR-H1* and *VRN-H3*). In this application, GW-AM is an excellent exploratory tool that generates results more quickly than structured biparental populations. However, unless perfect markers are used with multi-parent pedigrees, it may not be possible to determine the donors of favorable and unfavorable alleles.

REFERENCES

Akerberg A., Liljeberg H., Bjorck I. (1998) Effects of amylose/amylopectin ratio and baking conditions on resistant starch formation and glycaemic indices. Journal of Cereal Science 28:71-80. DOI: 10.1006/jcrs.1997.0173.

Ames N.P., Rhymer C.R. (2008) Issues surrounding health claims for barley. Journal of Nutrition 138:1237S-1243S.

Benjamini Y., Hochberg Y. (1995) Controlling The False Discovery Rate - A Practical and Powerful Approach To Multiple Testing. Journal of the Royal Statistical Society Series B-Methodological 57:289-300.

Box A., Washington J. M., Pallotta M. A., Eglinton J. K. (2003) Development of waxy hulless barley cultivars using marker-assisted selection. Conference proceedings : a joint meeting for the 11th Australian Barley Symposium and the 53rd Australian Cereal Chemistry Conference [electronic resource]: 5p.

Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler E.S. (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633-2635. DOI: 10.1093/bioinformatics/btm308.

Brennan C.S., Cleary L.J. (2005) The potential use of cereal (1 -> 3,1 -> 4)-beta-D-glucans as functional food ingredients. Journal of Cereal Science 42:1-13. DOI: 10.1016/j.jcs.2005.01.002.

Casas A.M., Djemel A., Ciudad F.J., Yahiaoui S., Ponce L.J., Contreras-Moreira B., Pilar Gracia M., Lasa J.M., Igartua E. (2011) HvFT1 (VrnH3) drives latitudinal adaptation in Spanish barleys. Theoretical and Applied Genetics 122:1293-1304. DOI: 10.1007/s00122-011-1531-x.

Close T.J., Bhat P.R., Lonardi S., Wu Y.H., Rostoks N., Ramsay L., Druka A., Stein N., Svensson J.T., Wanamaker S., Bozdag S., Roose M.L., Moscou M.J., Chao S.A.M., Varshney R.K., Szucs P., Sato K., Hayes P.M., Matthews D.E., Kleinhofs A., Muehlbauer G.J., DeYoung J., Marshall D.F., Madishetty K., Fenton R.D., Condamine P., Graner A., Waugh R. (2009) Development and implementation of high-throughput SNP genotyping in barley. Bmc Genomics 10. DOI: 10.1186/1471-2164-10-582.

Cockram J., White J., Leigh F.J., Lea V.J., Chiapparino E., Laurie D.A., Mackay I.J., Powell W., O'Sullivan D.M. (2008) Association mapping of partitioning loci in barley. Bmc Genetics 9. DOI: 1610.1186/1471-2156-9-16.

Crosbie G. B., Chiu P. C., Ross A. S. (2002) Shortened Temperature Program for Application with a Rapid Visco Analyser in Prediction of Noodle Quality in Wheat. Cereal Chemistry Journal 79(4):596-599.

Cuesta-Marcos A., Szucs P., Close T.J., Filichkin T., Muehlbauer G.J., Smith K.P., Hayes P.M. (2010) Genome-wide SNPs and re-sequencing of growth habit and inflorescence genes in barley: implications for association mapping in germplasm arrays varying in size and structure. Bmc Genomics 11. DOI: 70710.1186/1471-2164-11-707.

Dhillon T., Pearce S.P., Stockinger E.J., Distelfeld A., Li C., Knox A.K., Vashegyi I., Vagujfalvi A., Galiba G., Dubcovsky J. (2010) Regulation of Freezing

Tolerance and Flowering in Temperate Cereals: The VRN-1 Connection. Plant Physiology 153:1846-1858. DOI: 10.1104/pp.110.159079.

Domon E., Fuijita M., Ishikawa N. (2002) The insertion/deletion polymorphisms in the waxy gene of barley genetic resources from East Asia. Theoretical and Applied Genetics 104:132-138.

Dubcovsky J., Chen C.L., Yan L.L. (2005) Molecular characterization of the allelic variation at the VRN-H2 vernalization locus in barley. Molecular Breeding 15:395-407. DOI: 10.1007/s11032-005-0084-6.

Dubcovsky J., Loukoianov A., Fu D.L., Valarik M., Sanchez A., Yan L.L. (2006) Effect of photoperiod on the regulation of wheat vernalization genes VRN1 and VRN2. Plant Molecular Biology 60:469-480. DOI: 10.1007/s11103-005-4814-2.

Ehrenbergerova J., Belcredi N.B., Psota V., Hrstkova P., Cerkal R., Newman C.W. (2008) Changes caused by genotype and environmental conditions in beta-glucan content of spring barley for dietetically beneficial human nutrition. Plant Foods for Human Nutrition 63:111-117. DOI: 10.1007/s11130-008-0079-7.

Falush D., Stephens M., Pritchard J.K. (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 164:1567-1587.

Faure S., Higgins J., Turner A., Laurie D.A. (2007) The FLOWERING LOCUS T-like gene family in barley (Hordeum vulgare). Genetics 176:599-609. DOI: 10.1534/genetics.106.069500.

Food and Drug Administrarion: FDA (2006) Food labeling: ‘health claims; dietary fiber from certain foods and coronary heart disease’. In Code of Federal Regulation Title 21 Part 101.

Fowler D.B., Chauvin L.P., Limin A.E., Sarhan F. (1996) The regulatory role of vernalization in the expression of low-temperature-induced genes in wheat and rye. Theoretical and Applied Genetics 93:554-559. DOI: 10.1007/s001220050314.

Francia E., Barabaschi D., Tondelli A., Laido G., Rizza F., Stanca A.M., Busconi M., Fogher C., Stockinger E.J., Pecchioni N. (2007) Fine mapping of a HvCBF gene cluster at the frost resistance locus Fr-H2 in barley. Theoretical and Applied Genetics 115:1083-1091. DOI: 10.1007/s00122-007-0634-x.

Francia E., Rizza F., Cattivelli L., Stanca A.M., Galiba G., Toth B., Hayes P.M., Skinner J.S., Pecchioni N. (2004) Two loci on chromosome 5H determine low-temperature tolerance in a 'Nure' (winter) x 'Tremois' (spring) barley map. Theoretical and Applied Genetics 108:670-680. DOI: 10.1007/s00122-003-1468-9.

Fu D.L., Szucs P., Yan L.L., Helguera M., Skinner J.S., von Zitzewitz J., Hayes P.M., Dubcovsky J. (2005) Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat 273:54. Molecular Genetics and Genomics 274:442-443. DOI: 10.1007/s00438-005-0045-0.

Galiba G., Vagujfalvi A., Li C., Soltesz A., Dubcovsky J. (2009) Regulatory genes involved in the determination of frost tolerance in temperate cereals. Plant Science 176:12-19. DOI: 10.1016/j.plantsci.2008.09.016.

Hamblin M.T., Close T.J., Bhat P.R., Chao S., Kling J.G., Abraham K.J., Blake T., Brooks W.S., Cooper B., Griffey C.A., Hayes P.M., Hole D.J., Horsley R.D., Obert D.E., Smith K.P., Ullrich S.E., Muehlbauer G.J., Jannink J.-L. (2010) Population Structure and Linkage Disequilibrium in US Barley Germplasm: Implications for Association Mapping. Crop Science 50:556-566. DOI: 10.2135/cropsci2009.04.0198.

Hayes P.M., Liu B.H., Knapp S.J., Chen F., Jones B., Blake T., Franckowiak J., Rasmusson D., Sorrells M., Ullrich S.E., Wesenberg D., Kleinhofs A. (1993) Quantitative Trait Locus Effects and Environmental Interaction in a Sample of North-American Barley Germplasm. Theoretical and Applied Genetics 87:392-401.

Hemming M.N., Fieg S., Peacock W.J., Dennis E.S., Trevaskis B. (2009) Regions associated with repression of the barley (Hordeum vulgare) VERNALIZATION1 gene are not required for cold induction. Molecular Genetics and Genomics 282:107-117. DOI: 10.1007/s00438-009-0449-3.

Hemming M.N., Peacock W.J., Dennis E.S., Trevaskis B. (2008) Low-temperature and daylength cues are integrated to regulate FLOWERING LOCUS T in barley. Plant Physiology 147:355-366. DOI: 10.1104/pp.108.116418.

Hu G., Burton C. (2008) Modification of standard enzymatic protocol to a cost-efficient format for mixed-linkage (1 -> 3,1 -> 4)-beta-D-glucan measurement. Cereal Chemistry 85:648-653. DOI: 10.1094/cchem-85-5-0648.

Hung P.V., Maeda T., Morita N. (2006) Waxy and high-amylose wheat starches and flours - characteristics, functionality and application. Trends in Food Science & Technology 17:448-456. DOI: 10.1016/j.tifs.2005.12.006.

Kikuchi R., Kawahigashi H., Ando T., Tonooka T., Handa H. (2009) Molecular and Functional Characterization of PEBP Genes in Barley Reveal the Diversification of Their Roles in Flowering. Plant Physiology 149:1341-1353. DOI: 10.1104/pp.108.132134.

Kim W., Johnson J.W., Graybosch R.A., Gaines C.S. (2003) Physicochemical properties and end-use quality of wheat starch as a function of waxy protein alleles. Journal of Cereal Science 37:195-204. DOI: 10.1006/jcrs.2002.0494.

Kosova K., Prasil I.T., Vitamvas P. (2008) The relationship between vernalization-and photoperiodically-regulated genes and the development of frost tolerance in wheat and barley. Biologia Plantarum 52:601-615. DOI: 10.1007/s10535-008-0120-6.

Luna A., Nicodemus K.K. (2007) snp.plotter: an R-based SNP/haplotype association and linkage disequilibrium plotting package. Bioinformatics 23:774-776. DOI: 10.1093/bioinformatics/btl657.

Mahfoozi S., Limin A.E., Hayes P.M., Hucl P., Fowler D.B. (2000) Influence of photoperiod response on the expression of cold hardiness in wheat and barley. Canadian Journal of Plant Science 80:721-724.

McCleary B.V., Codd R. (1991) MEASUREMENT OF (1- 3),(1- 4)-Beta-D-Glucan in Barley and Oat - A Streamlined Enzymatic Procedure. Journal of the Science of Food and Agriculture 55:303-312. DOI: 10.1002/jsfa.2740550215.

Patron N.J., Smith A.M., Fahy B.F., Hylton C.M., Naldrett M.J., Rossnagel B.G., Denyer K. (2002) The altered pattern of amylose accumulation in the endosperm of low-amylose barley cultivars is attributable to a single mutant allele of granule-bound starch synthase I with a deletion in the 5 '-non-coding region. Plant Physiology 130:190-198. DOI: 10.1104/pp.005454.

Pritchard J.K., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Roy J.K., Smith K.P., Muehlbauer G.J., Chao S., Close T.J., Steffenson B.J. (2010) Association mapping of spot blotch resistance in wild barley. Molecular Breeding 26:243-256. DOI: 10.1007/s11032-010-9402-8.

Shin J-H., Blay S., McNeney B. and Graham J. (2006). LDheatmap: An R Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms. J Stat Soft, 16 Code Snippet 3

Skinner J., Szucs P., von Zitzewitz J., Marquez-Cedillo L., Filichkin T., Stockinger E.J., Thomashow M.F., Chen T.H.H., Hayes P.M. (2006) Mapping of barley homologs to genes that regulate low temperature tolerance in Arabidopsis. Theoretical and Applied Genetics 112:832-842. DOI: 10.1007/s00122-005-0185-y.

Skinner J.S., von Zitzewitz J., Szucs P., Marquez-Cedillo L., Filichkin T., Amundsen K., Stockinger E.J., Thomashow M.F., Chen T.H.H., Hayes P.M. (2005) Structural, functional, and phylogenetic characterization of a large CBF gene family in barley. Plant Molecular Biology 59:533-551. DOI: 10.1007/s11103-005-2498-2.

Stockinger E.J., Cheng, H. and Skinner, J.S. (2006) Structural organization of barley CBF genes coincident with QTLs for cold hardiness. In Chen T.H.H., Uemura M. and Fujikawa S., (ed.) Cold Hardiness in Plants. Molecular Genetics, Cell Biology and Physiology. Oxford, UK. 53–63.

Szucs P., Karsai I., von Zitzewitz J., Meszaros K., Cooper L.L.D., Gu Y.Q., Chen T.H.H., Hayes P.M., Skinner J.S. (2006) Positional relationships between photoperiod response QTL and photoreceptor and vernalization genes in barley. Theoretical and Applied Genetics 112:1277-1285. DOI: 10.1007/s00122-006-0229-y.

Szucs P., Skinner J.S., Karsai I., Cuesta-Marcos A., Haggard K.G., Corey A.E., Chen T.H.H., Hayes P.M. (2007) Validation of the VRN-H2/VRN-H1 epistatic model in barley reveals that intron length variation in VRN-H1 may account for a continuum of vernalization sensitivity. MGG Molecular Genetics and Genomics 277:249-261. DOI: :10.1007/s00438-006-0915-8.

Takahashi R., and Yasuda S. (1971) Genetics of earliness and growth habit in barley. In R.A. Nilan (ed.) Barley genetics II. Proc.Intl. Barley Genet. Symp., 2nd, Pullman, WA. 6–11 July 1969. 388–408.

Tester R. F. and W. R. Morrison. (1990) Swelling and Gelatinization of Cereal Starches. I. Effects of Amylopectin, Amylose, and Lipids. Cereal Chem 67:551-557.

Trevaskis B. (2010) The central role of the VERNALIZATION1 gene in the vernalization response of cereals. Functional Plant Biology 37:479-487. DOI: 10.1071/fp10056.

Trevaskis B., Hemming M.N., Dennis E.S., Peacock W.J. (2007) The molecular basis of vernalization-induced flowering in cereals. Trends in Plant Science 12:352-357. DOI: 10.1016/j.tplants.2007.06.010.

Trevaskis B., Hemming M.N., Peacock W.J., Dennis E.S. (2006) HvVRN2 responds to daylength, whereas HvVRN1 is regulated by vernalization and developmental status. Plant Physiology 140:1397-1405. DOI: 10.1104/pp.105.073486.

Turner A., Beales J., Faure S., Dunford R.P., Laurie D.A. (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. Science 310:1031-1034. DOI: 10.1126/science.1117619.

von Zitzewitz J., Condon F., Corey A., Cuesta-Marcos A., Filichkin T., Haggard K., Fisk SP., Smith K.P., Muehlbauer G.J., Karsai I., Hayes P.M. (2011) The genetics of winterhardiness in barley: perspectives from genome-wide association mapping. The Plant Genome 4:76–91 DOI: 10.3835/plantgenome2010.12.0030.

Wang G., Schmalenbach I., von Korff M., Leon J., Kilian B., Rode J., Pillen K. (2010) Association of barley photoperiod and vernalization genes with QTLs for flowering time and agronomic traits in a BC(2)DH population and a set of wild barley introgression lines. Theoretical and Applied Genetics 120:1559-1574. DOI: 10.1007/s00122-010-1276-y.

Wang H., Smith K.P., Combs E., Blake T., Horsley R.D., and Muehlbauer G.J. (2011) Effect of population size and unbalanced data sets on QTL detection using genome-wide association mapping in barley breeding germplasm. Theoretical and Applied Genetics. DOI: 10.1007/s00122-011-1691-8.

Waugh R., Jannink J.-L., Muehlbauer G.J., Ramsay L. (2009) The emergence of whole genome association scans in barley. Current Opinion in Plant Biology 12:218-222. DOI: 10.1016/j.pbi.2008.12.007.

Xue Q., Wang L., Newman R.K., Newman C.W., Graham H. (1997) Influence of the hulless, waxy starch and short-awn genes on the composition of barleys. Journal of Cereal Science 26:251-257.

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

Yan L.L., Loukoianov A., Blechl A., Tranquilli G., Ramakrishna W., SanMiguel P., Bennetzen J.L., Echenique V., Dubcovsky J. (2004) The wheat VRN2 gene is

a flowering repressor down-regulated by vernalization. Science 303:1640-1644. DOI: 10.1126/science.1094305.

Zhang G.P., Wang J.M., Chen J.X. (2002) Analysis of beta-glucan content in barley cultivars from different locations of China. Food Chemistry 79:251-254. DOI: 10.1016/s0308-8146(02)00127-9.

Figure 2.1. Scheme for germplasm development of fall-sown barley with waxy and normal starch using marker-assisted selection (MAS) and phenotypic selection (PS). Parental lines were Luca(L), Merlin(M), Strider(S), Waxbar(W), Maja(Ma), Legacy(Le) and Doyce(D)**.**

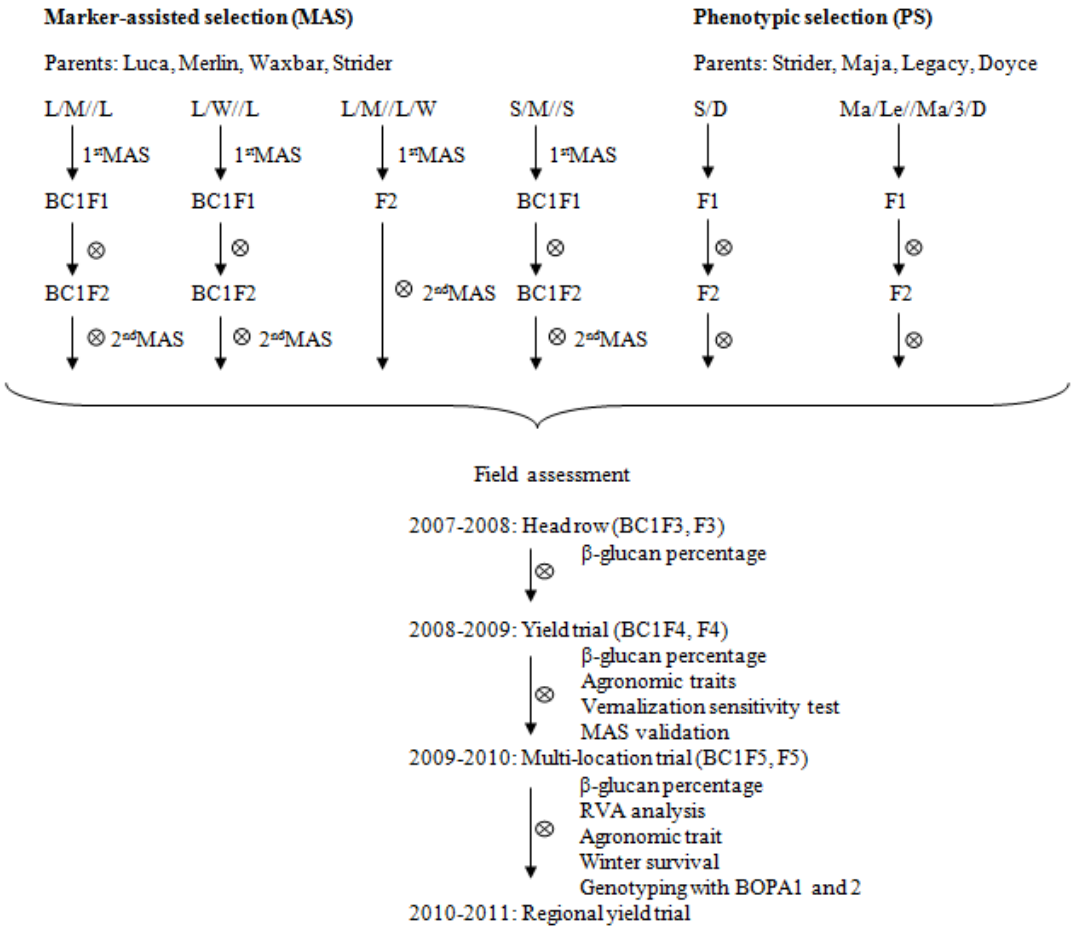


Figure 2.2. Peak time vs peak viscosity in barley CAP IV samples grown at Corvallis, Oregon. Triangle indicates the phenotypic selection (PS) group with non-waxy starch. Circle indicates marker-assisted selection (MAS) group with waxy starch. 09OR-18 and 09OR-21 are intermediate lines. 09OR-16 is a *WxWx* MAS line.

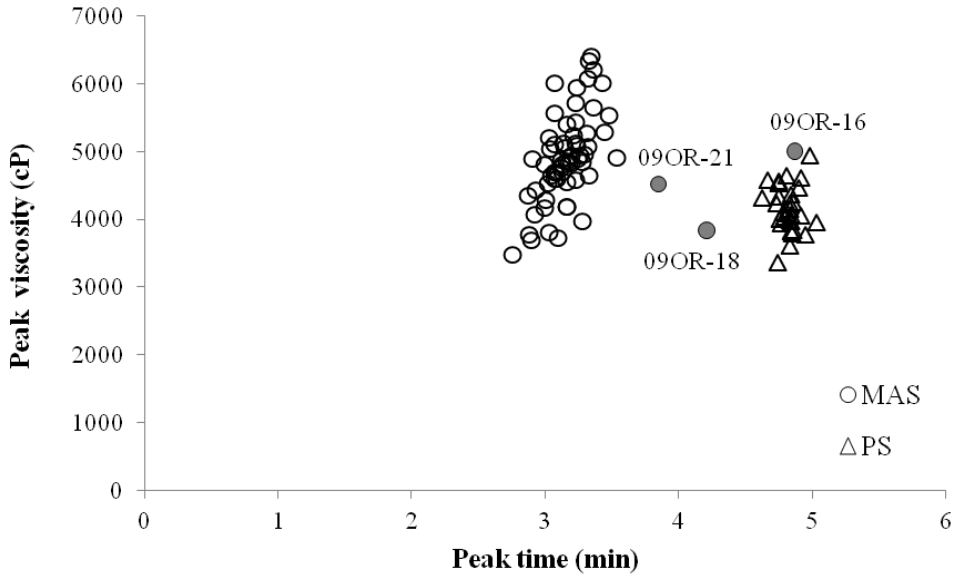


Figure 2.3. Phenotypic frequency distribution for grain β-glucan percentage at A) Hermiston (HER), and B) Corvallis (CVO), Oregon. Four parents were used as checks. Merlin and Waxbar are waxy varieties. Luca and Strider have normal starch. No data are shown for Waxbar because it did not survive the winter.

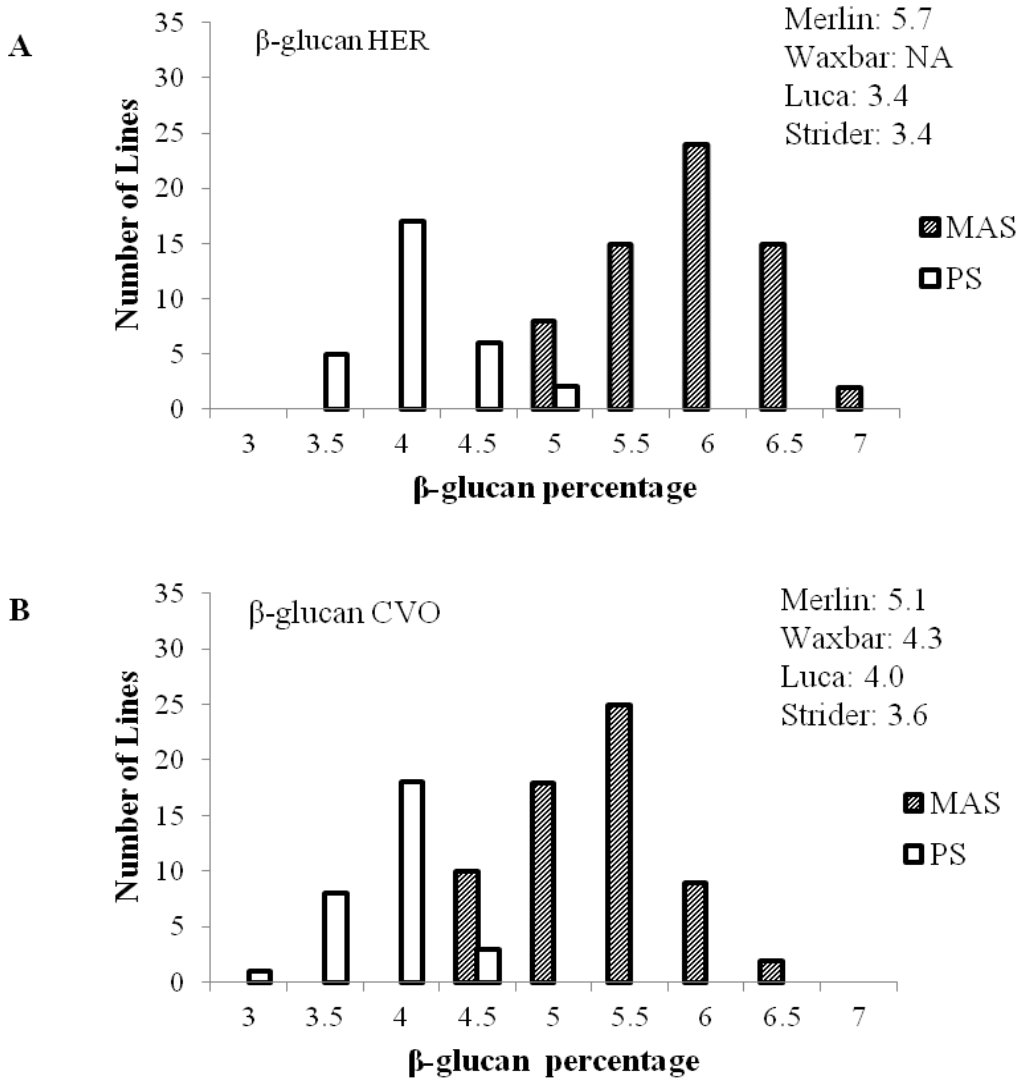


Figure 2.4. Grain β-glucan percentage at Hermiston (HER), Oregon genome-wide association scans using three datasets. A) full CAP IV dataset (*n* = 99) B) marker-assisted selection (MAS) dataset (*n* = 69) C) phenotypic selection (PS) dataset (*n* = 65). The significance threshold was determined as described in the text.

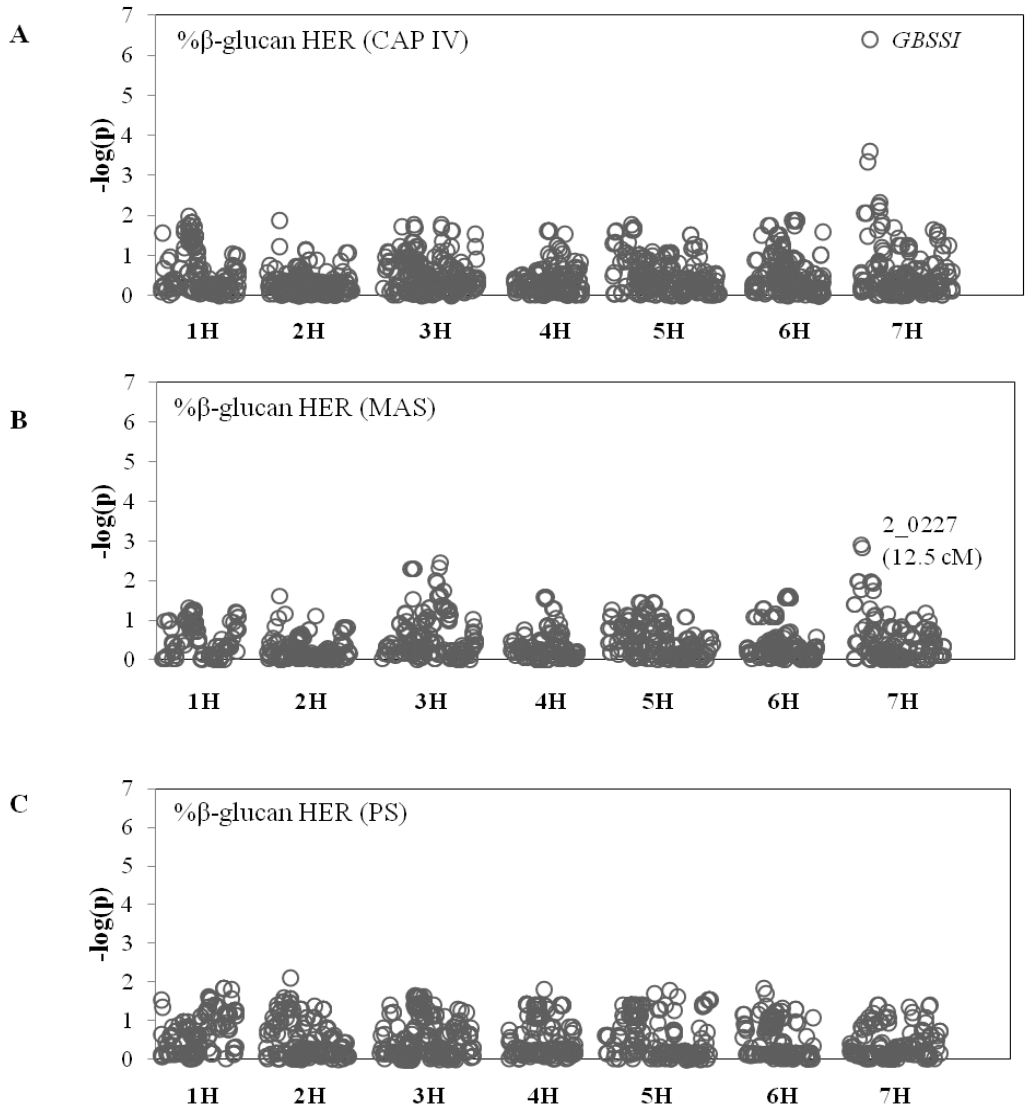


Figure 2.5. Phenotypic frequency distribution for low temperature tolerance (LTT) evaluated at Hermiston (HER), Oregon. The data represents the mean of two replications with standard error shown of four parents and checks.

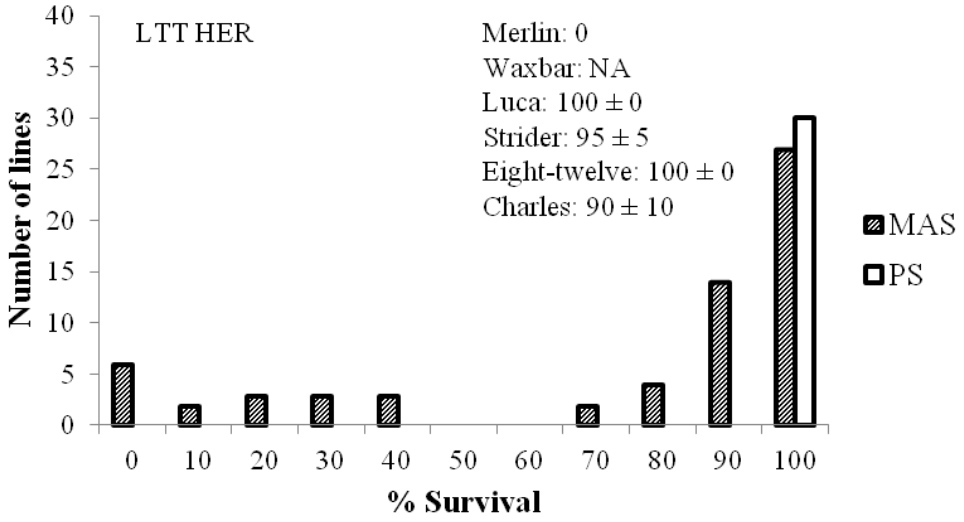


Figure 2.6. Low temperature tolerance (LTT) at Hermiston (HER) genome-wide association scans using three datasets. A) full CAP IV dataset (*n* = 99) B) marker-assisted selection (MAS) dataset (*n* = 69) C) phenotypic selection (PS) dataset (*n* = 65). The significance threshold was determined as described in the text.

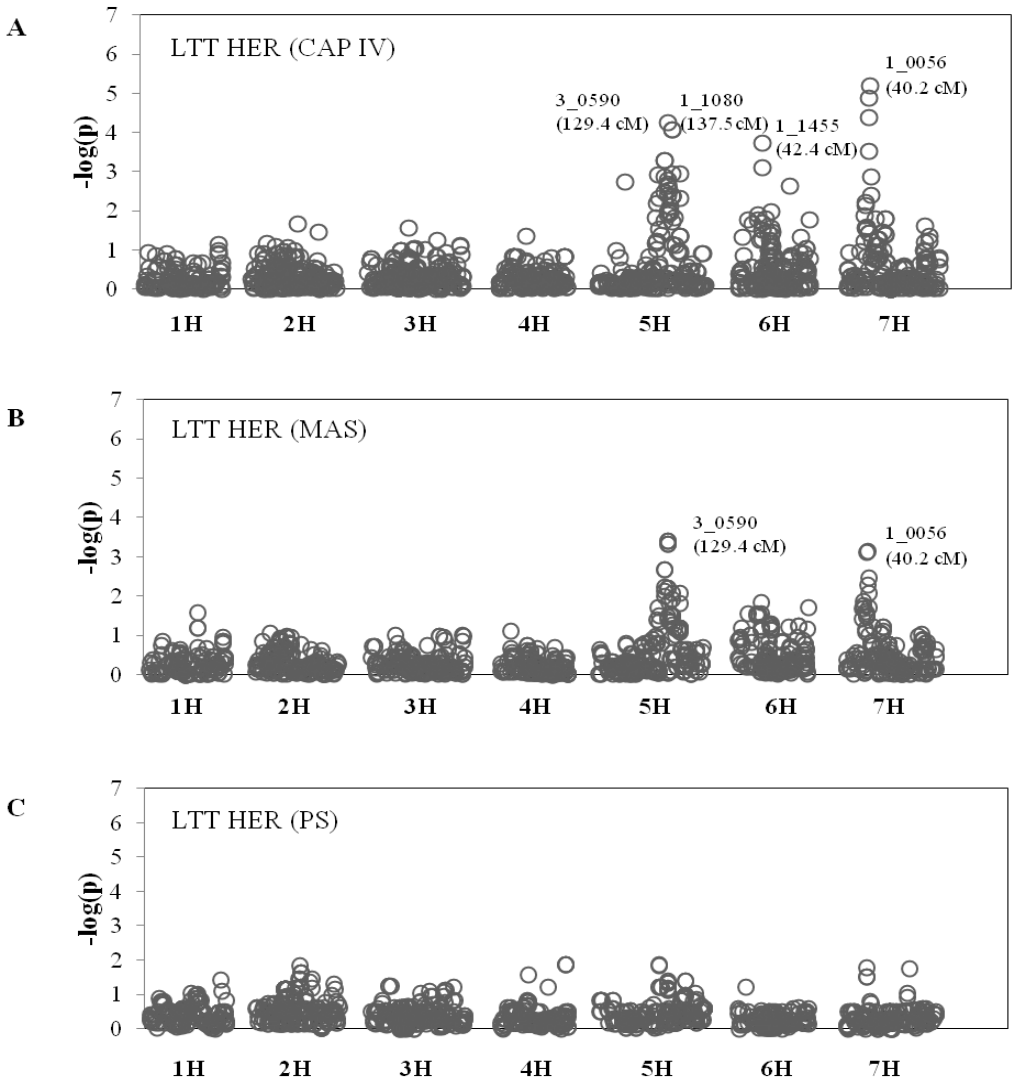


Figure 2.7. Linkage disequilibrium heat plot for 11 cM region on chromosome 5H containing *HvBM5A* (*VRN-H1*). The consensus map of *HvBM5A* is indicated. SNP 1_1080 and 3_0590 showed significant association with low temperature tolerance.

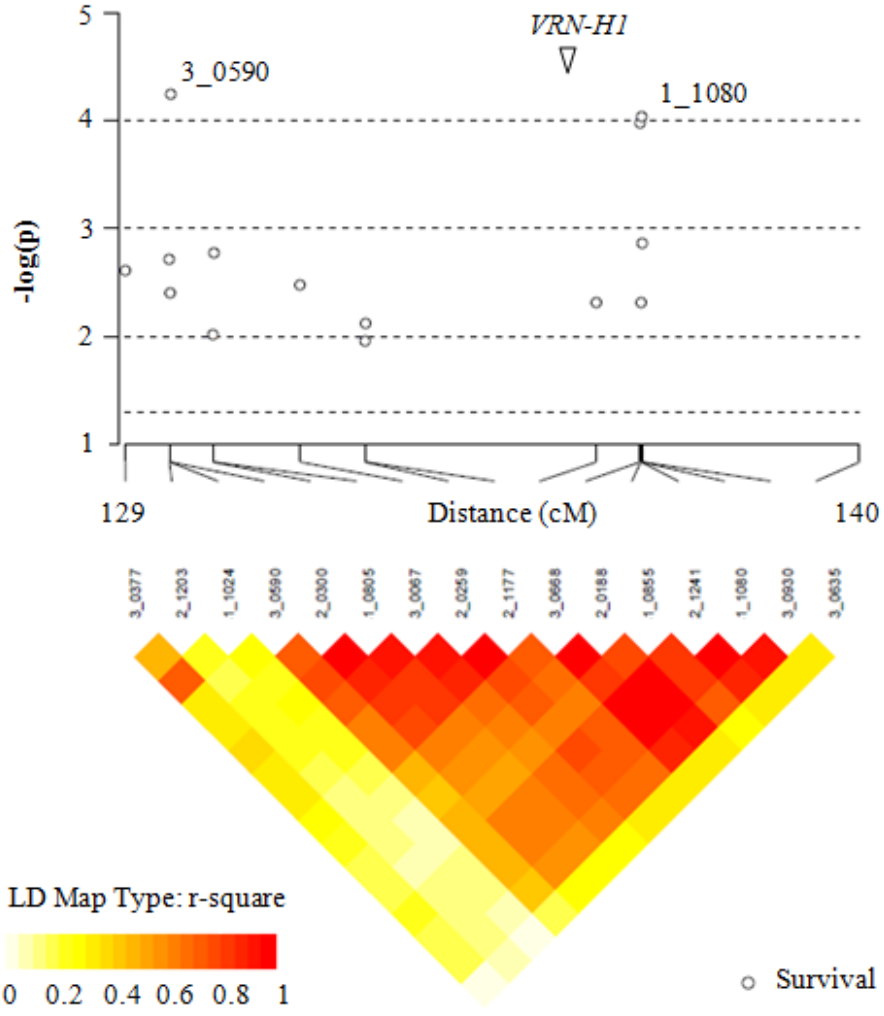
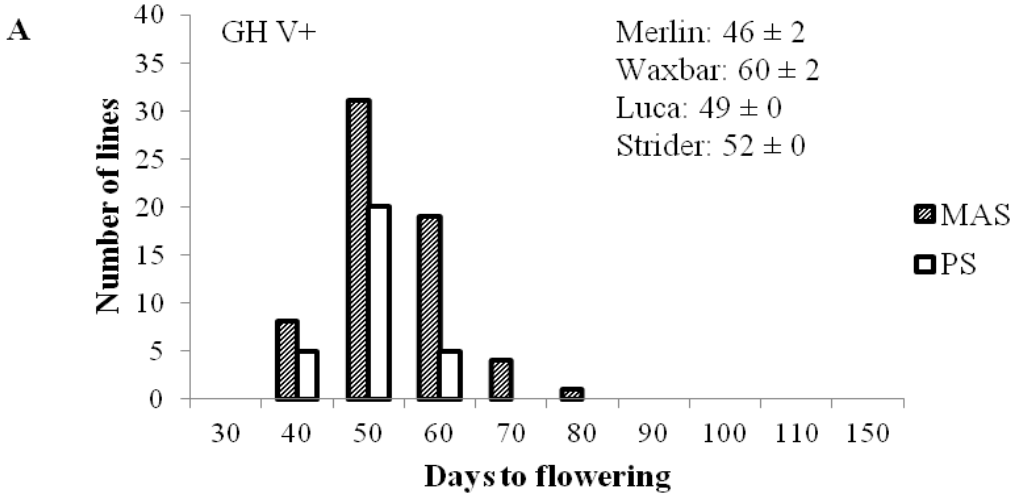


Figure 2.8. Phenotypic frequency distributions for vernalization sensitivity experiments including the marker-assisted selection (MAS) and phenotypic selection (PS) groups within the Oregon CAP IV association mapping panel. Days to flowering were evaluated under greenhouse (GH) conditions (A) with and (B) without vernalization (V+ and V-). The data represents the mean of two replicates. Mean and standard error are shown for the four parents. No standard errors are shown for Luca and Strider because they did not flower without vernalization and were assigned the arbitrary value of 150.



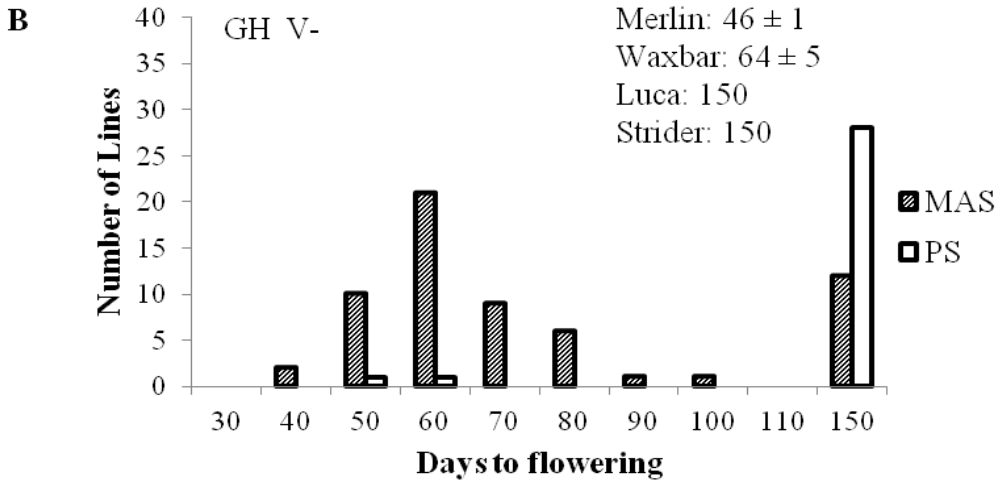


Figure 2.9. Days to flowering genome-wide association scans of unvernalization treatment (V-) under greenhouse (GH) conditions using three datasets. A) full CAP IV dataset (*n* = 99) B) marker-assisted selection (MAS) dataset (*n* = 69) C) phenotypic selection (PS) dataset (*n* = 65). The significance threshold was determined as described in the text.

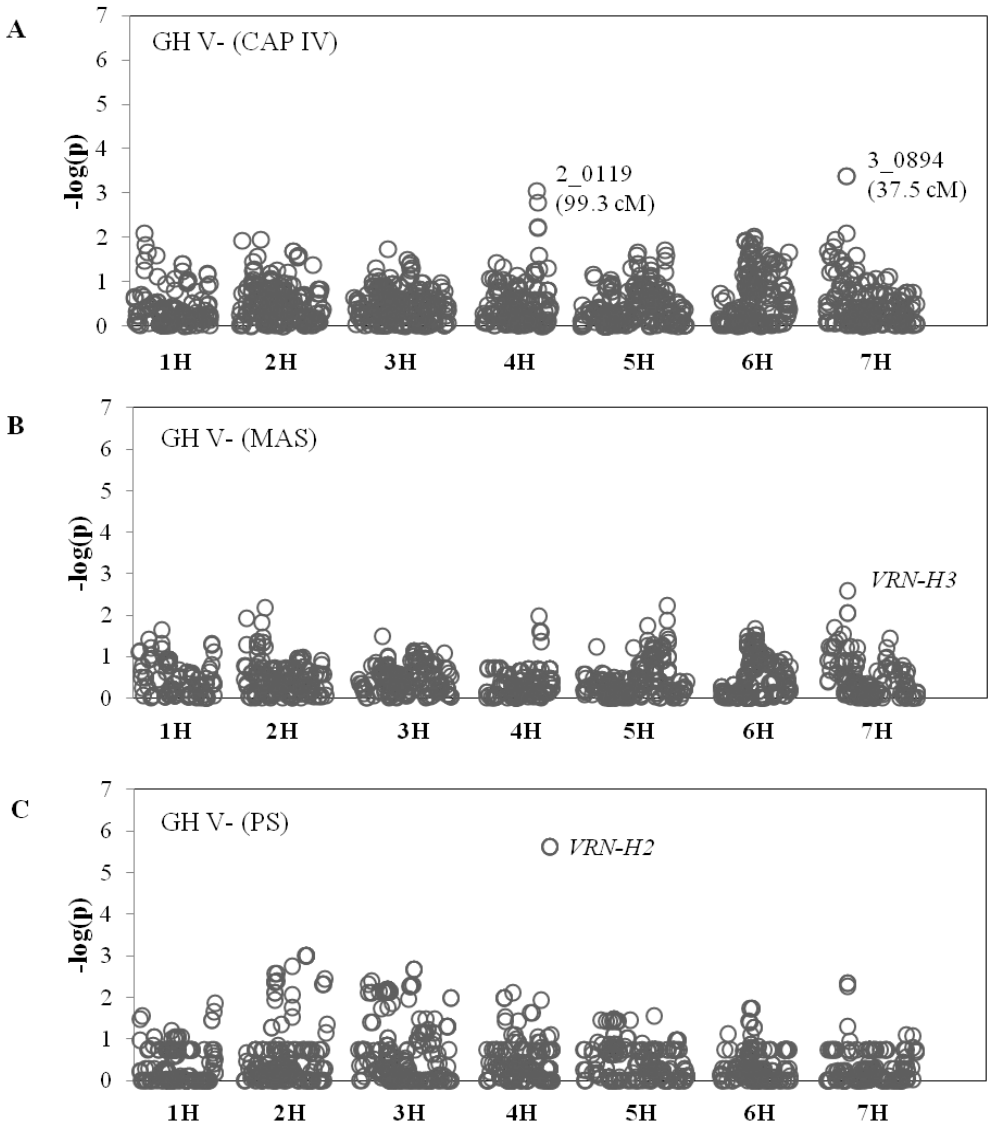
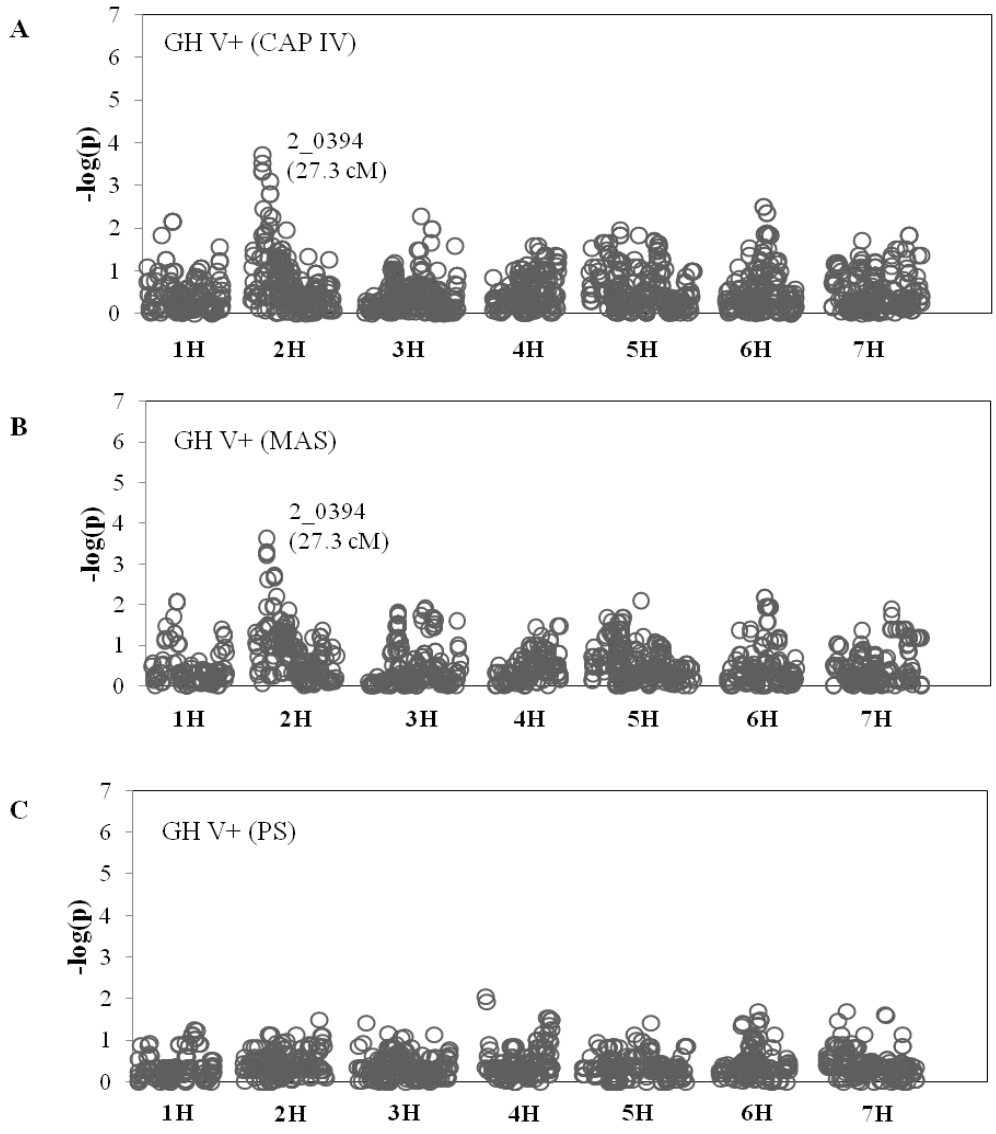


Figure 2.10. Days to flowering genome-wide association scans of vernalization treatment (V+) under greenhouse (GH) conditions using three datasets. A) full CAP IV dataset (*n* = 99) B) marker-assisted selection (MAS) dataset (*n* = 69) C) phenotypic selection (PS) dataset (*n* = 65). The significance threshold was determined as described in the text.



GENERAL CONCLUSIONS

With the advent of high-throughput NGS technology, it is now straightforward and cost-effective to generate vast amounts of high-quality DNA sequence data. The Illumina GoldenGate assay was adopted by the Barley CAP project and its European and South American counterparts (AGOUEB and FONTAGRO). Barley researchers have the necessary data (ESTs) already in place to utilize this technology. Researchers on all crops can benefit from sequenced RAD markers, a new SNP-based marker that allows for simultaneous high-density SNP discovery and genotyping. We demonstrated the effectiveness of sequenced RAD markers for linkage map construction and QTL mapping by using the OWB population. An EST-based SNP map, with marker information generated by the Illumina GoldenGate assay was available in the OWB population. Therefore, we could compare RAD vs. Illumina-based maps, and combine the RAD and Illumina-based data. The RAD markers provided full genome coverage. Locus ordering and map distance estimates are in agreement with the Illumina-based map. The RAD dataset merged seamlessly with the Illumina dataset and was useful for QTL mapping. We conclude that the RAD markers will be useful for genetic analyses and breeding applications in many species.

The addition of RAD data to the OWB mapping population will make it an even more robust international resource for genetics research and instruction. The OWB map is populated with all types of marker data, making it useful for map integration. In terms of instruction, the population can be used to demonstrate the use

of different marker types (e.g. dominant and codominant) and to reinforce the fundamental concept that marker assay platforms vary but the basis of allelic variation remains the same.

A high quality linkage map is necessary for generating the best estimates of QTL location, effect, and interaction. Once QTLs have been detected, candidate genes can be identified, cloned, and characterized. Linked, or ideally perfect markers can be identified for MAS. Marker-assisted selection can accelerate the breeding process and make selection more effective in the early generations of a breeding program. Perfect markers based on functional polymorphisms will be most useful for MAS, particularly in cases where there are numerous opportunities for recombination during generation advance. We used perfect markers for target alleles at the waxy (*WX*) and *VRN-H2* genes to rapidly develop winter habit food barley. Selection for *wxwx* genotype was intended to raise grain β-glucan percentage and selection for *Vrn-H2Vrn-H2* was intended to cause VS and therefore maximum LTT. We were successful in using MAS to fix target alleles at both loci but we achieved only one of the target phenotypes (higher β-glucan percentage). We did not achieve VS due to two factors (i) incomplete information about *VRN-H1* at the outset of the project and (ii) unexpected allelic variation at *VRN-H3* with a large effect on VS and LTT.

The discovery of allelic variation at *VRN-H1* and *VRN-H2* was achieved through GW-AM. The GW-AM process involved integration of field-based phenotypic data and full-genome Illumina SNP data, followed by the appropriate analyses to account for population structure, and to determine the significance of

marker: trait associations. In this analysis, the genome coordinates of significant markers are established by reference to a consensus linkage map. The OWB population is a key contributor of map information to the consensus map.

Genome-wide association mapping validated the effectiveness of MAS for *wxwx* and MAS provided one of two germplasm resources (the other was a *WxWx* germplasm array) for validating the effectiveness of GW-AM. In this albeit circular fashion, the usefulness of perfect markers was demonstrated. Genome-wide association mapping proved to be a powerful tool for detecting the general genome location of a gene (or genes) determining target phenotypes. We were able to determine, through GW-AM, that Merlin and Waxbar do not have recessive alleles at *VRN-H1*, as we had erroneously concluded based on limited information at the outset of the project. In this case, further MAS efforts in this germplasm could be based on selection against Merlin and Waxbar alleles. However, this endeavor would benefit from information on the functional polymorphism, not markers in LD with the functional polymorphism. Genome-wide association mapping revealed an unexpected effect of *VRN-H3* on LTT and VS. This information will be a key in further enhancing our newly-created winter food barley germplasm. However, additional research will be necessary to establish the functional basis of the association. Currently, we have markers in LD with the determinant(s) of VS and LTT but we do not know what alleles to target in MAS. Genome-wide association mapping can be a more efficient QTL detection tool than biparental QTL mapping. The former can be applied directly to any array of breeding germplasm segregating for the trait of

interest. The latter requires construction of a structured population followed by genotyping, phenotyping, and analysis. However, the extra work has a benefit: QTL effects and locations are determined simultaneously. In the case of GW-AM, follow-up biparental mapping may be necessary to determine allele effect.

In summary, this research has contributed to the development of linkage mapping resources in barley, via RAD mapping in the OWB population. It has confirmed the value of reference populations, such as the OWBs in developing the consensus linkage maps necessary for determining the genome coordinates of markers that are significant in GW-AM. We showed that genetic resources, such as the OWB, can be used for biparental mapping of QTL determining agronomic fitness traits. We created agronomically relevant food barley germplasm with higher levels of grain β-glucan. Finally, we discovered an important role of *VRN-H3* in VS and LTT which may allow plant breeders to improve their barley germplasm and/or develop varieties with high LTT.

BIBLIOGRAPHY

Agrama H.A., Eizenga G.C., Yan W. (2007) Association mapping of yield and its components in rice cultivars. Molecular Breeding 19:341-356. DOI: 10.1007/s11032-006-9066-6.

Akhunov E., Nicolet C., Dvorak J. (2009) Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. Theoretical and Applied Genetics 119:507-517. DOI: 10.1007/s00122-009-1059-5.

Alonso-Blanco C., Aarts M.G.M., Bentsink L., Keurentjes J.J.B., Reymond M., Vreugdenhil D., Koornneef M. (2009) What Has Natural Variation Taught Us about Plant Development, Physiology, and Adaptation? Plant Cell 21:1877-1896. DOI: 10.1105/tpc.109.068114.

Baird N.A., Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z.A., Selker E.U., Cresko W.A., Johnson E.A. (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. Plos One 3. DOI: 10.1371/journal.pone.0003376.

Barchi L., Lanteri S., Portis E., Acquadro A., Vale G., Toppino L., Rotino G.L. (2011) Identification of SNP and SSR markers in eggplant using RAD tag sequencing. Bmc Genomics 12. DOI: 30410.1186/1471-2164-12-304.

Brennan C.S., Cleary L.J. (2005) The potential use of cereal (1 -> 3,1 -> 4)-beta-D-glucans as functional food ingredients. Journal of Cereal Science 42:1-13. DOI: 10.1016/j.jcs.2005.01.002.

Burton R.A., Collins H.M., Kibble N.A.J., Smith J.A., Shirley N.J., Jobling S.A., Henderson M., Singh R.R., Pettolino F., Wilson S.M., Bird A.R., Topping D.L., Bacic A., Fincher G.B. (2011) Over-expression of specific HvCslF cellulose synthase-like genes in transgenic barley increases the levels of cell wall (1,3;1,4)-beta-D-glucans and alters their fine structure. Plant Biotechnology Journal 9:117-135. DOI: 10.1111/j.1467-7652.2010.00532.x.

Burton R.A., Wilson S.M., Hrmova M., Harvey A.J., Shirley N.J., Stone B.A., Newbigin E.J., Bacic A., Fincher G.B. (2006) Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)-beta-D-glucans. Science 311:1940-1942. DOI: 10.1126/science.1122975.

Chen A., Reinheimer J., Brule-Babel A., Baumann U., Pallotta M., Fincher G.B., Collins N.C. (2009) Genes and traits associated with chromosome 2H and 5H regions controlling sensitivity of reproductive tissues to frost in barley. Theoretical and Applied Genetics 118:1465-1476. DOI: 10.1007/s00122-009-0995-4.

Chutimanitsakun Y., Nipper R.W., Cuesta-Marcos A., Cistue L., Corey A., Filichkina T., Johnson E.A., Hayes P.M. (2011) Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. Bmc Genomics 12. DOI: 10.1186/1471-2164-12-4.

Cistue L., Cuesta-Marcos A., Chao S., Echavarri B., Chutimanitsakun Y., Corey A., Filichkina T., Garcia-Marino N., Romagosa I., Hayes P.M. (2011) Comparative mapping of the Oregon Wolfe Barley using doubled haploid lines derived from female and male gametes. Theoretical and Applied Genetics 122:1399-1410. DOI: 10.1007/s00122-011-1540-9.

Close T.J., Bhat P.R., Lonardi S., Wu Y.H., Rostoks N., Ramsay L., Druka A., Stein N., Svensson J.T., Wanamaker S., Bozdag S., Roose M.L., Moscou M.J., Chao S.A.M., Varshney R.K., Szucs P., Sato K., Hayes P.M., Matthews D.E., Kleinhofs A., Muehlbauer G.J., DeYoung J., Marshall D.F., Madishetty K., Fenton R.D., Condamine P., Graner A., Waugh R. (2009) Development and implementation of high-throughput SNP genotyping in barley. Bmc Genomics 10. DOI: 10.1186/1471-2164-10-582.

Collard B.C.Y., Jahufer M.Z.Z., Brouwer J.B., Pang E.C.K. (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142:169-196. DOI: 10.1007/s10681-005-1681-5.

Costa J.M., Corey A., Hayes P.M., Jobet C., Kleinhofs A., Kopisch-Obusch A., Kramer S.F., Kudrna D., Li M., Riera-Lizarazu O., Sato K., Szucs P., Toojinda T., Vales M.I., Wolfe R.I. (2001) Molecular mapping of the Oregon Wolfe Barleys: a phenotypically polymorphic doubled-haploid population. Theoretical and Applied Genetics 103:415-424.

Cuesta-Marcos A., Szucs P., Close T.J., Filichkin T., Muehlbauer G.J., Smith K.P., Hayes P.M. (2010) Genome-wide SNPs and re-sequencing of growth habit and inflorescence genes in barley: implications for association mapping in germplasm arrays varying in size and structure. Bmc Genomics 11. DOI: 70710.1186/1471-2164-11-707.

Darvasi A., Weinreb A., Minke V., Weller J.I., Soller M. (1993) Detecting Marker-QTL Linkage and Estimating QTL gene Effect and Map Location Using a Saturated Genetic-Map. Genetics 134:943-951.

Davey J.W., Hohenlohe P.A., Etter P.D., Boone J.Q., Catchen J.M., Blaxter M.L. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics 12:499-510. DOI: 10.1038/nrg3012.

Dhillon T., Pearce S.P., Stockinger E.J., Distelfeld A., Li C., Knox A.K., Vashegyi I., Vagujfalvi A., Galiba G., Dubcovsky J. (2010) Regulation of Freezing Tolerance and Flowering in Temperate Cereals: The VRN-1 Connection. Plant Physiology 153:1846-1858. DOI: 10.1104/pp.110.159079.

Doblin M.S., Pettolino F.A., Wilson S.M., Campbell R., Burton R.A., Fincher G.B., Newbigin E., Bacic A. (2009) A barley cellulose synthase-like CSLH gene mediates (1,3;1,4)-beta-D-glucan synthesis in transgenic Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 106:5996-6001. DOI: 10.1073/pnas.0902019106.

Domon E., Fuijita M., Ishikawa N. (2002) The insertion/deletion polymorphisms in the waxy gene of barley genetic resources from East Asia. Theoretical and Applied Genetics 104:132-138.

Emebiri L.C. (2009) EST-SSR markers derived from an elite barley cultivar (Hordeum vulgare L. 'Morex'): polymorphism and genetic marker potential. Genome 52:665-676. DOI: 10.1139/g09-040.

European Food Safety Authority: (EFSA 2009) Scientific Opinion on the substantiation of health claims related to beta-glucans and maintenance of normal blood cholesterol concentrations (ID 754, 755, 757, 801, 1465, 2934) and maintenance or achievement of a normal body weight (ID 820, 823) pursuant to Article 13(1) of Regulation (EC) No 1924/20061. In Summary: EFSA Journal 7(9):1254.

Flint-Garcia S.A., Thornsberry J.M., Buckler E.S. (2003) Structure of linkage disequilibrium in plants. Annual Review of Plant Biology 54:357-374. DOI: 10.1146/annurev.arplant.54.031902.134907.

Fowler D.B., Chauvin L.P., Limin A.E., Sarhan F. (1996) The regulatory role of vernalization in the expression of low-temperature-induced genes in wheat and rye. Theoretical and Applied Genetics 93:554-559. DOI: 10.1007/s001220050314.

Food and Drug Administrarion: FDA (2006) Food labeling: 'health claims; dietary fiber from certain foods and coronary heart disease'. In Code of Federal Regulation Title 21 Part 101.

Francia E., Barabaschi D., Tondelli A., Laido G., Rizza F., Stanca A.M., Busconi M., Fogher C., Stockinger E.J., Pecchioni N. (2007) Fine mapping of a HvCBF gene cluster at the frost resistance locus Fr-H2 in barley. Theoretical and Applied Genetics 115:1083-1091. DOI: 10.1007/s00122-007-0634-x.

Francia E., Rizza F., Cattivelli L., Stanca A.M., Galiba G., Toth B., Hayes P.M., Skinner J.S., Pecchioni N. (2004) Two loci on chromosome 5H determine low-temperature tolerance in a 'Nure' (winter) x 'Tremois' (spring) barley map. Theoretical and Applied Genetics 108:670-680. DOI: 10.1007/s00122-003-1468-9.

Ganal M.W., Altmann T., Roeder M.S. (2009) SNP identification in crop plants. Current Opinion in Plant Biology 12:211-217. DOI: 10.1016/j.pbi.2008.12.009.

Gupta P.K., Roy J.K., Prasad M. (2001) Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Current Science 80:524-535.

Gupta P.K., Rustgi S., Kulwal P.L. (2005) Linkage disequilibrium and association studies in higher plants: Present status and future prospects. Plant Molecular Biology 57:461-485. DOI: 10.1007/s11103-005-0257-z.

Gupta P.K., Rustgi S., Mir R.R. (2008) Array-based high-throughput DNA markers for crop improvement. Heredity 101:5-18. DOI: 10.1038/hdy.2008.35.

Han F., Ullrich S.E., Chirat S., Menteur S., Jestin L., Sarrafi A., Hayes P.M., Jones B.L., Blake T.K., Wesenberg D.M., Kleinhofs A., Kilian A. (1995) Mapping if

Beta-glucan Content and Beta-glucanase Activity Loci in Barley-Grain and Malt. Theoretical and Applied Genetics 91:921-927.

Hayes P.M., Liu B.H., Knapp S.J., Chen F., Jones B., Blake T., Franckowiak J., Rasmusson D., Sorrells M., Ullrich S.E., Wesenberg D., Kleinhofs A. (1993) Quantitative Trait Locus Effects and Environmental Interaction in a Sample of North-American Barley Germplasm. Theoretical and Applied Genetics 87:392-401.

Hori K., Kobayashi T., Shimizu A., Sato K., Takeda K., Kawasaki S. (2003) Efficient construction of high-density linkage map and its application to QTL analysis in barley. Theoretical and Applied Genetics 107:806-813. DOI: 10.1007/s00122-003-1342-9.

Hu G., Burton C. (2008) Modification of standard enzymatic protocol to a cost-efficient format for mixed-linkage (1 -> 3,1 -> 4)-beta-D-glucan measurement. Cereal Chemistry 85:648-653. DOI: 10.1094/cchem-85-5-0648.

Li J.Z., Baga M., Rossnagel B.G., Legge W.G., Chibbar R.N. (2008) Identification of quantitative trait loci for beta-glucan concentration in barley grain. Journal of Cereal Science 48:647-655. DOI: 10.1016/j.jcs.2008.02.004.

Limin A., Corey A., Hayes P., Fowler D.B. (2007) Low-temperature acclimation of barley cultivars used as parents in mapping populations: response to photoperiod, vernalization and phenological development. Planta 226:139-146. DOI: 10.1007/s00425-006-0475-x.

Matsuki J., Yasui T., Sasaki T., Fujita M., Kitarnura Y. (2008) Effects of the barley amo1 and wax genes on starch structure and physicochemical properties. Starch-Starke 60:279-285. DOI: 10.1002/star.200800006.

Miller M.R., Atwood T.S., Eames B.F., Eberhart J.K., Yan Y.L., Postlethwait J.H., Johnson E.A. (2007b) RAD marker microarrays enable rapid mapping of zebrafish mutations. Genome Biology 8. DOI: 10.1186/gb-2007-8-6-r105.

Miller M.R., Dunham J.P., Amores A., Cresko W.A., Johnson E.A. (2007a) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Research 17:240-248. DOI: 10.1101/gr.5681207.

Molina-Cano J.L., Moralejo M., Elia M., Munoz P., Russell J.R., Perez-Vendrell A.M., Ciudad F., Swanston J.S. (2007) QTL analysis of a cross between European and North American malting barleys reveals a putative candidate gene for beta-glucan content on chromosome 1H. Molecular Breeding 19:275-284. DOI: 10.1007/s11032-006-9075-5.

Myles S., Peiffer J., Brown P.J., Ersoz E.S., Zhang Z., Costich D.E., Buckler E.S. (2009) Association Mapping: Critical Considerations Shift from Genotyping to Experimental Design. Plant Cell 21:2194-2202. DOI: 10.1105/tpc.109.068437.

Nelson J.C., Wang S., Wu Y., Li X., Antony G., White F.F., Yu J. (2011) Single-nucleotide polymorphism discovery by high-throughput sequencing in sorghum. Bmc Genomics 12. DOI: 35210.1186/1471-2164-12-352.

Nevo E. (1992) Origin, evolution, population genetics and resources for breeding of wild barley *Hordeum spontaneum* in the fertile crescent. In Shewry P.R. (ed)

Barley: genetics, biochemistry, molecular biology and biotechnology. CAB International Oxon, UK. 19-43.

Pan A., Hayes P.M., Chen F., Chen T.H.H., Blake T., Wright S., Karsai I., Bedo Z. (1994) Genetic-analysis of The Components of Winterhardiness in Barley (HORDEUM-VULGARE L). Theoretical and Applied Genetics 89:900-910.

Patron N.J., Smith A.M., Fahy B.F., Hylton C.M., Naldrett M.J., Rossnagel B.G., Denyer K. (2002) The altered pattern of amylose accumulation in the endosperm of low-amylose barley cultivars is attributable to a single mutant allele of granule-bound starch synthase I with a deletion in the 5 '-non-coding region. Plant Physiology 130:190-198. DOI: 10.1104/pp.005454.

Pfender W.F., Saha M.C., Johnson E.A., Slabaugh M.B. (2011) Mapping with RAD (restriction-site associated DNA) markers to rapidly identify QTL for stem rust resistance in Lolium perenne. Theoretical and Applied Genetics 122:1467-1480. DOI: 10.1007/s00122-011-1546-3.

Pins J.J., Kaur H. (2006) A review of the effects of barley beta-glucan on cardiovascular and diabetic risk. Cereal Foods World 51:8-11. DOI: 10.1094/cfw-51-0008.

Rafalski A. (2002) Applications of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology 5:94-100. DOI: 10.1016/s1369-5266(02)00240-6.

Reinheimer J.L., Barr A.R., Eglinton J.K. (2004) QTL mapping of chromosomal regions conferring reproductive frost tolerance in barley (Hordeum vulgare L.). Theoretical and Applied Genetics 109:1267-1274. DOI: 10.1007/s00122-004-1736-3.

Romagosa I., Han F., Ullrich S.E., Hayes P.M., Wesenberg D.M. (1999) Verification of yield QTL through realized molecular marker-assisted selection responses in a barley cross. Molecular Breeding 5:143-152.

Rowe H.C., Renaut S., Guggisberg A. (2011) RAD in the realm of next-generation sequencing technologies. Molecular Ecology 20:3499-3502. DOI: 10.1111/j.1365-294X.2011.05197.x.

Roy J.K., Smith K.P., Muehlbauer G.J., Chao S., Close T.J., Steffenson B.J. (2010) Association mapping of spot blotch resistance in wild barley. Molecular Breeding 26:243-256. DOI: 10.1007/s11032-010-9402-8.

Sato K., Nankaku N., Takeda K. (2009) A high-density transcript linkage map of barley derived from a single population. Heredity 103:110-117. DOI: 10.1038/hdy.2009.57.

Semagn K., Bjornstad A., Xu Y. (2010) The genetic dissection of quantitative traits in crops. Electronic Journal of Biotechnology 13. DOI: 1410.2225/vol13-issue5-fulltext-14.

Skinner J., Szucs P., von Zitzewitz J., Marquez-Cedillo L., Filichkin T., Stockinger E.J., Thomashow M.F., Chen T.H.H., Hayes P.M. (2006) Mapping of barley homologs to genes that regulate low temperature tolerance in Arabidopsis. Theoretical and Applied Genetics 112:832-842. DOI: 10.1007/s00122-005-0185-y.

Skinner J.S., von Zitzewitz J., Szucs P., Marquez-Cedillo L., Filichkin T., Amundsen K., Stockinger E.J., Thomashow M.F., Chen T.H.H., Hayes P.M. (2005) Structural, functional, and phylogenetic characterization of a large CBF gene family in barley. Plant Molecular Biology 59:533-551. DOI: 10.1007/s11103-005-2498-2.

Szucs P., Karsai I., von Zitzewitz J., Meszaros K., Cooper L.L.D., Gu Y.Q., Chen T.H.H., Hayes P.M., Skinner J.S. (2006) Positional relationships between photoperiod response QTL and photoreceptor and vernalization genes in barley. Theoretical and Applied Genetics 112:1277-1285. DOI: 10.1007/s00122-006-0229-y.

Szucs P., Skinner J.S., Karsai I., Cuesta-Marcos A., Haggard K.G., Corey A.E., Chen T.H.H., Hayes P.M. (2007) Validation of the VRN-H2/VRN-H1 epistatic model in barley reveals that intron length variation in VRN-H1 may account for a continuum of vernalization sensitivity. Molecular Genetics and Genomics 277:249-261. DOI: 10.1007/s00438-006-0195-8.

Szucs P., Blake V.C., Bhat P.R., Chao S., Close T.J., Cuesta-Marcos A., Muehlbauer G.J., Ramsay L.V., Waugh R., Hayes P.M. An integrated resource for barley linkage map and malting quality QTL alignment. The Plant Genome 2009, 2:134-140. DOI: 10.3835/plantgenome2008.01.0005.

Takahashi R., and Yasuda S. (1971) Genetics of earliness and growth habit in barley. In R.A. Nilan (ed.) Barley genetics II. Proc.Intl. Barley Genet. Symp., 2nd, Pullman, WA. 6–11 July 1969. 388–408.

Taketa S., Yuo T., Tonooka T., Tsumuraya Y., Inagaki Y., Haruyama N., Larroque O., and Jobling S.A. (2011) Functional characterization of barley betaglucanless mutants demonstrates a unique role for CslF6 in (1,3;1,4)-β-D-glucan biosynthesis. J. Exp. Bot. DOI: 10.1093/jxb/err285.

Tanksley S.D., Nelson J.C. (1996) Advanced backcross QTL analysis: A method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theoretical and Applied Genetics 92:191-203. DOI: 10.1007/s001220050114.

Trevaskis B. (2010) The central role of the VERNALIZATION1 gene in the vernalization response of cereals. Functional Plant Biology 37:479-487. DOI: 10.1071/fp10056.

Turner A., Beales J., Faure S., Dunford R.P., Laurie D.A. (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. Science 310:1031-1034. DOI: 10.1126/science.1117619.

Vales M.I., Schon C.C., Capettini F., Chen X.M., Corey A.E., Mather D.E., Mundt C.C., Richardson K.L., Sandoval-Islas J.S., Utz H.F., Hayes P.M. (2005) Effect of population size on the estimation of QTL: a test using resistance to barley stripe rust. Theoretical and Applied Genetics 111:1260-1270. DOI: 10.1007/s00122-005-0043-y.

van Os H., Andrzejewski S., Bakker E., Barrena I., Bryan G.J., Caromel B., Ghareeb B., Isidore E., de Jong W., van Koert P., Lefebvre V., Milbourne D., Ritter E., van der Voort J., Rousselle-Bourgeois F., van Vliet J., Waugh R., Visser

R.G.F., Bakker J., van Eck H.J. (2006) Construction of a 10,000-marker ultradense genetic recombination map of potato: Providing a framework for accelerated gene isolation and a genomewide physical map. Genetics 173:1075-1087. DOI: 10.1534/genetics.106.055871.

Wang G., Schmalenbach I., von Korff M., Leon J., Kilian B., Rode J., Pillen K. (2010) Association of barley photoperiod and vernalization genes with QTLs for flowering time and agronomic traits in a BC(2)DH population and a set of wild barley introgression lines. Theoretical and Applied Genetics 120:1559-1574. DOI: 10.1007/s00122-010-1276-y.

Wang H., Smith K.P., Combs E., Blake T., Horsley R.D., and Muehlbauer G.J. (2011) Effect of population size and unbalanced data sets on QTL detection using genome-wide association mapping in barley breeding germplasm. Theoretical and Applied Genetics. DOI: 10.1007/s00122-011-1691-8.

Xue Q., Newman R.K., Newman C.W., McGuire C.F. (1991) Waxy Gene Effects on Beta-glucan, Dietary Fiber Content and Viscosity of Barleys. Cereal Research Communications 19:399-404.

Xue Q., Wang L., Newman R.K., Newman C.W., Graham H. (1997) Influence of the hulless, waxy starch and short-awn genes on the composition of barleys. Journal of Cereal Science 26:251-257.

Yan J., Shah T., Warburton M.L., Buckler E.S., McMullen M.D., Crouch J. (2009) Genetic Characterization and Linkage Disequilibrium Estimation of a Global Maize Collection Using SNP Markers. Plos One 4. DOI: e845110.1371/journal.pone.0008451.

Yan J., Yang X., Shah T., Sanchez-Villeda H., Li J., Warburton M., Zhou Y., Crouch J.H., Xu Y. (2010) High-throughput SNP genotyping with the GoldenGate assay in maize. Molecular Breeding 25:441-451. DOI: 10.1007/s11032-009-9343-2.

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

Yu H., Xie W., Wang J., Xing Y., Xu C., Li X., Xiao J., Zhang Q. (2011) Gains in QTL Detection Using an Ultra-High Density SNP Map Based on Population Sequencing Relative to Traditional RFLP/SSR Markers. Plos One 6. DOI: e1759510.1371/journal.pone.0017595.

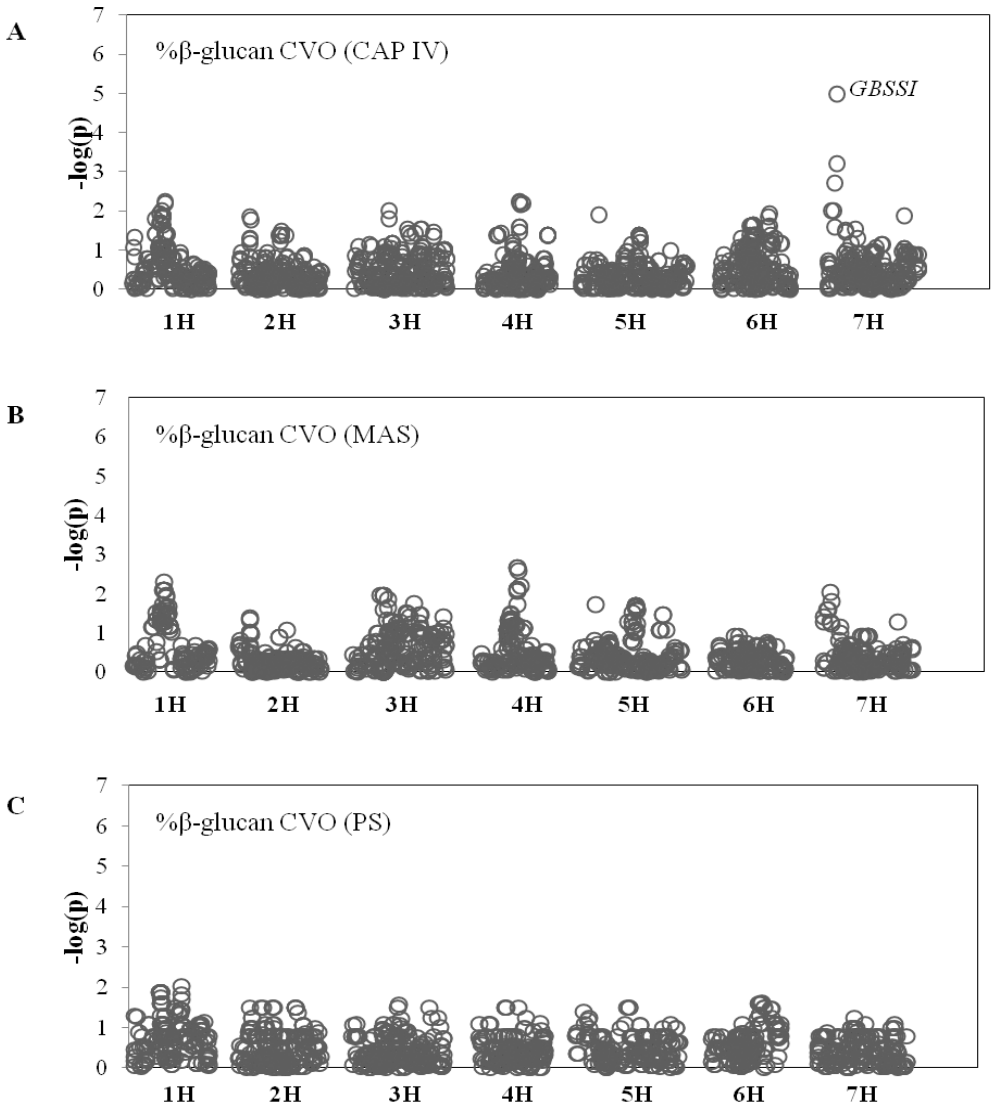
Yu J.M., Buckler E.S. (2006) Genetic association mapping and genome organization of maize. Current Opinion in Biotechnology 17:155-160. DOI: 10.1016/j.copbio.2006.02.003.

von Zitzewitz J., Condon F., Corey A., Cuesta-Marcos A., Filichkina T., Haggard K., Fisk SP., Smith K.P., Muehlbauer G.J., Karsai I., Hayes P.M. (2011) The genetics of winterhardiness in barley: perspectives from genome-wide association mapping. The Plant Genome 4:76–91 DOI: 10.3835/plantgenome2010.12.0030.

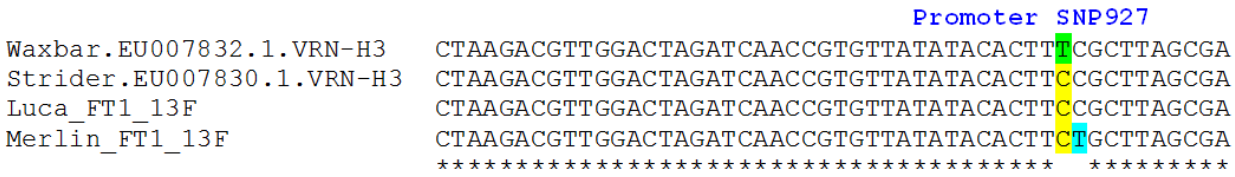
Zhao K., Aranzana M.J., Kim S., Lister C., Shindo C., Tang C., Toomajian C., Zheng H., Dean C., Marjoram P., Nordborg M. (2007) An Arabidopsis example of association mapping in structured samples. Plos Genetics 3. DOI: e410.1371/journal.pgen.0030004.

APPENDIX

Supplement Figure 1. Grain β-glucan percentage at Corvallis (CVO), Oregon genome-wide association scans using three datasets. A) full CAP IV dataset (*n* = 99) B) marker-assisted selection (MAS) dataset (*n* = 69) C) phenotypic selection (PS) dataset (*n* = 65). The significance threshold was determined as described in the text.



Supplement Figure 2. Sequence alignment among marker-assisted selection parents.



Supplement table 1. Average monthly temperature and total monthly rainfall at two locations in Oregon during the growing period in 2010.

	Monthly Temperature (^{o}C)		**Monthly Precipitation (mm)**	
	Hermiston	**Corvallis**	**Hermiston**	**Corvallis**
September	17.9	16.6	1.8	27.9
October	9.9	10.9	34.0	72.6
November	4.9	7.2	9.4	193.3
December	-2.9	2.3	26.4	157.0
January	3.7	7.5	39.9	134.6
February	5.3	7.9	15.0	103.9
March	7.8	7.9	10.9	141.2
April	11.1	9	18.8	101.9
May	13.7	11.2	31.8	78.0
June	18.5	14.7	33.5	66.0
July	23.3	18.7	4.1	2.0
August	21.8	18.2	7.9	12.7

Supplemental Table 2. Multi-locus haplotypes of seven parents on chromosome 5H in the vicinity of *VRN-H1*. 3_0883 is a SNP in the intron I of *HvBM5*.

Pilot OPA Name	Position (cM)	Spring/Facultative parents			Winter Parents			
		Merlin	Waxbar	Legacy	Strider	Luca	Maja	Doyce
3_0590	129.4	A	A	B	B	B	B	B
1_0805	130.1	A	A	B	B	B	B	B
2_0259	132.6	B	B	A	A	A	A	A
3_0883	135.7	A	A	-	A	A	A	-
3_0869	136.4	B	B	B	A	A	A	A
1_0095	137.2	A	A	A	B	B	B	B
1_1080	137.2	B	B	B	A	A	A	A