Genetic analysis of Barley (*Hordeum vulgare* L.) grain (1,3;1,4)-β-glucan concentration and fine structure.

A Thesis Submitted to the College of
Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Deparment of Plant Sciences
University of Saskatchewan
Saskatoon, Saskatchewan

By
Aron Thomas Cory

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements of a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this university may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised this thesis work or, in their absence, by the Head of the Department or the Dean of the College in which this thesis work was done. It is understood that any copying or publication or use of this thesis or part therof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Request for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department

Plant Sciences

51 Campus Drive,

University of Saskatchewan,

Saskatoon, Saskatchewan, Canada

S7N 5A8

ABSTRACT

(1,3;1,4)-β-glucan accumulated in barley (*Hordeum vulgare* L.) cell walls is an important determinant for grain end-use as food, malt, feed or fuel. As a trait affected by multiple genes and the environment, grain (1,3;1,4)-β-glucan concentration qualifies as a quantitative trait. A major QTL on chromosome 7H contains a cellulose synthase like gene *HvCslF6*, coding for an enzyme associated with (1,3;1,4)-β-glucan biosynthesis. To develop gene based perfect markers, *HvCslF6* was analyzed to determine allelic variation between CDC Bold, a low (1,3;1,4)-β-glucan (~ 3.3 %) cultivar and TR251, a high (1,3;1,4)-β-glucan (~ 5.2 %) genotype. Comparison of the CDC Bold and TR251 nucleotide sequences downstream of the ATG start codon in *HvCslF6* alleles revealed 16 single nucleotide polymorphisms (SNPs) and two indels. The two indels added 16 nucleotides to the first intron of *HvCslF6* of CDC Bold and a single SNP in the third exon changed alanine 590 codon in the CDC Bold sequence to a threonine codon in TR251 allele. Five polymorphic sites were converted into genetic markers and confirmed to select low and high (1,3;1,4)-β-glucan lines in a previously characterized CDC Bold / TR251 doubled haploid genetic mapping population and a novel F_5 recombinant inbred line (RIL) population derived from a Merit / H93174006 (4.8 and 5.3 % (1,3;1,4)-β-glucan) cross. An analysis of parental lines of six populations segregating for (1,3;1,4)-β-glucan concentration validated the association between the TR251 *HvCslF6* haplotype and high (1,3;1,4)-β-glucan concentration in populations showing a (1,3;1,4)-β-glucan quantitative trait locus (QTL) on chromosome 7H.

To further investigate the role of *HvCslF6* alleles, 91 lines of the Merit / H93174006 RIL grown in two environments were phenotyped for (1,3;1,4)-β-glucan grain concentration, cellotriose content (DP3), cellotetraose content (DP4) and cellotriose:cellotetraose (DP3:DP4) ratio. DP3, DP4, (1,3;1,4)-β-glucan and total DP3+DP4 were strongly positively correlated (r>0.9) to each other, suggesting no preference for DP3 or DP4 subunit production in high or low (1,3;1,4)-β-glucan lines. DP3:DP4 ratio showed no strong correlation with any other measured trait. Significant effects arising from genotype and environment were associated with grain (1,3;1,4)-β-glucan concentration, DP3, DP4 and DP3:DP4 ratio. Only DP3:DP4 ratio showed a significant GxE (genotype by environment) interaction. Hereditability of grain (1,3;1,4)-β-glucan concentration was moderate (~ 30 %), DP3 and DP4 had low heritability (> 21 %) and DP3:DP4 ratio had moderate heritability (~ 43 %). Single marker analysis showed an

association between marker CSLF6_4105 and (1,3;1,4)-β-glucan fine structure in Vegreville but not in Castor, supporting significant GxE interaction in (1,3;1,4)-β-glucan fine structure. Association mapping of candidate markers in 119 barley genotypes of diverse origin grown in greenhouses showed that on chromosome 7H, marker CSLF6_4105 was associated only with (1,3;1,4)-β-glucan concentration, while Bmac273e was associated with both (1,3;1,4)-β-glucan concentration and DP3:DP4 ratio. In addition on chromosome 1H, markers Bmac504 and Bmac211 were associated only with DP3:DP4 ratio. This study suggests that DP3:DP4 ratio is strongly affected by genotype and environment.

To identify new markers with (1,3;1,4)-β-glucan concentration, ninety-four two-row spring varieties were genotyped using double digestion Restriction-site Associated DNA (ddRAD) sequencing on an Illumina sequencer. Two bioinformatics pipelines were used to discover and call SNPs for association linkage analysis. SAMtools bioinformatics pipeline identified 9,062 markers and UNEAK identified 3,060 markers, 2,311 of which were identical between both bioinformatics pipelines. Both sets of markers showed excellent coverage of the genome and distinguished the ninety-four varieties into the same subgroups based on geographical region of origin. Association mapping was performed using TASSEL 3.0 and grain (1,3;1,4)-β-glucan concentration was associated with a region on the 5HS telomere by markers generated using both UNEAK and SAMtools. Some putative candidate genes were identified, including a UDP-glucosyltransferase, two phosphorylation signaling proteins and two transcription factors. The markers developed and tested in this study can be used in marker assisted selection to develop barley genotypes with desired (1,3;1,4)-β-glucan concentration.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

.

LIST OF FIGURES

LIST OF ABBREVIATIONS

AACCI	American Association of Cereal Chemists International
AFLP	Amplified Fragment Length Polymorphism
AM	Association Mapping
ANOVA	Analysis of variance
AX	Arabinoxylans
BAC	Bacterial Artificial Chromosome
Btr	Brittle Rachis
BSA	Bovine Serum Albumin
CAPS	Cleaved Amplified Polymorphic Sequences
CesA	Cellulose Synthase Subfamily A
CIM	Composite Interval Mapping
CM	Chloroform Methanol
Csl	Cellulose Synthase-Like
CSR	Class Specific Region
DAP	Days After Pollination
DArT®	Diversity Array Technology
DCB	2,6-Dichlorobenzonitrilen
ddRAD	Double Digestion Restriction-site Associated DNA
DP3:DP4	Cellotriose:Cellotetraose Ratio
DP3	Cellotriose (Degree of Polymerization 3)
DP4	Cellotetraose (Degree of Polymerization 4)
ELISA	Enzyme-Linked Immunosorbent Assay
EMS	Ethyl MethaneSulfonate
Eng1	Endo-1,3-Beta Glucosidase
ERF	Ethylene Response Factor
EST	Expressed Sequence Tag
FAO	Food and Agricultural Organization of the United Nations
FDA	US Food and Drug Administration
FDR	False Discovery Rate
GAX	Glucuronoarabinoxylan
GBS	Genotyping-By-Sequencing
GBSSI	Granule Bound Starch Synthase I
GLM	General Linear Model
GWAS	Genome-Wide Association Studies
GxE	Genotype by Environment
HDAC	Histone Deacetylase Complex
HMW	High Molecular Weight
HPAEC-PAD	High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detection
HVR	Highly Variable Region
Indel	Insertion / Deletion
int-c	Intermedium Gene
int-c.a	Dominant Intermedium Allele

LD	Linkage Disequilibrium
LMW	Low Molecular Weight
LOD	Logarithm of Odds
MAF	Minor Allele Frequency
MCMC	Markov Chain Monte Carlo
mRNA	Messenger RNA
M_W	Molecular Weight
nC	Nanocoulomb
NGS	Next Generation Sequencing
PCA	Principal Component Analysis
PCoA	Principal Coordinate Analysis
QQ	Quintile-Quintile
QTL	Quantitative Trait Locus
RAPD	Random Amplified Polymorphic DNA
RE	Restriction Endonucleases
RIL	Recombinant Inbred Line
RING	Really Interesting New Gene
RLFP	Restriction Fragment Length Polymorphism
SIM	Simple Interval Mapping
SNP	Single Nucleotide Polymorphism
SSCP	Single Strand Conformation Polymorphism
STS	Sequence Tagged Site
TAIR	The Arabidopsis Information Resource
TILLING	Targeting Induced Local Lesions in Genomes
TMD	Transmembrane Domains
UDP	Uridine DiPhosphate
UDP-Glc	Uridine DiPhosphate Glucose
UGE	UDP-glucose-4 epimerase
USDA	United States Department of Agriculture
VCF	Variant Call file
WGS	Whole-Genome Shotgun
YFP	Yellow Flourescent Protein

CHAPTER 1
INTRODUCTION

1.1 Background

Barley (*Hordeum vulgare* L.) is a member of the Triticeae tribe within the Poaceae family and is an important component of global agriculture. After corn, rice and wheat, barley is the fourth most important cereal crop worldwide in terms of production, acreage and consumption (FAO, 2014). Barley is produced in diverse environments from the subarctic to the sub-tropics. Barley has a shorter growing season, can be grown at higher altitudes, matures earlier, requires less heat units and is better adapted to drought than many other cereal crops (Harlan, 1979). Barley grows best on light or sandy loam soils and is more tolerant to salinity and alkalinity than most other cereals. Barley can use soils low in nitrogen content and tolerate soils which are too light or coarse textured for wheat (van Gool and Vernon, 2006). The success of artificial selection on barley is evident by its diverse uses in food, feed and industrial applications. In North America, roughly 65 % of barley production is used for animal feed, 30 % is used in the malting and brewing industry to produce beer and whiskey and two percent is used as food for human consumption (Newman and Newman, 2006).

Barley utilization is influenced by its grain composition. Carbohydrates (starch and non-starch), proteins and lipids are the major storage components in a grain. Vitamins, minerals and minor compounds such as phenolics and pigments are the minor components present in a barley grain (Newman and Newman, 2008). Among the major non-starch carbohydrates accumulated in the barley grains is the dietary fiber, (1,3;1,4)-β-glucan, which is an important factor determining grain use. Traditionally, barley has been selected for lower concentrations of grain (1,3;1,4)-β-glucan to meet the requirements of the malt and feed industries. High concentrations of grain (1,3;1,4)-β-glucan have a direct, deleterious effect during malting (Swanston et al., 1995). Residual (1,3;1,4)-β-glucan leads to highly viscous wort and causes chill haze in bright beer (Bamforth, 1982; Wang et al., 2004). High concentrations of (1,3;1,4)-β-glucan are also problematic when barley is used as a feed for non-ruminants, especially poultry (Gohl et al., 1978).

Non-starch polysaccharides are major components of dietary fibers. Consumption of dietary fiber by humans is associated with reduced incidence of coronary heart disease (Anderson, 1995; Pereira et al., 2004). The Canadian Heart and Stroke Foundation recommends a daily intake of 21 - 38 grams of fiber; specifically, soluble fiber, which may help reduce cholesterol and blood sugar levels (Ames and Rhymer, 2008). The (1,3;1,4)-β-glucan molecule is considered a soluble dietary fiber, as it has unique physical and physiological properties that make it a functional, bio-active ingredient of interest in human food products (Cui and Wood, 2000). Due to its physical propreties, (1,3;1,4)-β-glucan can be used as a thickening agent to modify the look and feel of gravies, salad dressings, and ice cream formulations (Wood, 1986) or as a replacement for fat to develop calorie-reduced foods (Inglett, 1990). In food systems, high viscosity (1,3;1,4)-β-glucan is associated with reduced plasma cholesterol and a better control of postprandial serum glucose levels (Lazaridou and Biliaderis, 2007; Wood, 2007; Klopfenstein and Hoseney, 1987). In 2006, barley was approved to carry the health claim for reducing the risk of cardiovascular disease (FDA news release, 2005). In 2011, human food use for barley increased by 21 % over the four year average (Statistics Canada, 2014) suggesting it as an important growing market for Canadian barley. In summary, the demand for barley in human nutrition in the Western world is increasing due to several health benefits associated with whole grains and dietary fiber. Therefore, high grain (1,3;1,4)-β-glucan concentration is desirable for barley used as food by humans.

(1,3;1,4)-β-glucan is a mixed linked glucan polymer made up by β-D-glucopyranosyl molecules that can form either β-1,3 or β-1,4 linkages. The β-1,4-linked molecules are constituted primarily of non-randomly arranged cellotriose (DP3) and cellotetraose (DP4) units which form a polymer with linkage ratios ranging from 2.0:1 to 4.8:1 (Burton et al., 2011; Collins et al., 2010; Mikkelsen et al., 2013). The ratio of DP3:DP4 in (1,3;1,4)-β-glucan affect the viscosity of (1,3;1,4)-β-glucan, a factor influencing its end use in malting and brewing, feed or food applications (Izydorczyk et al., 2000). The (1,3;1,4)-β-glucan concentration in barley grain can range from very low (≤1 %) to high concentrations (up to 21 %) but normally varies between three to six percent (Munck et al., 2004; Kato et al., 1995; Holtekjølen et al., 2006). Similarly, the concentration and DP3:DP4 ratio of (1,3;1,4)-β-glucan show a larger variation in barley grain than in other commercially produced cereals (Collins et al., 2010; Wood, 2007).

This suggests that ample opportunity exists to genetically alter (1,3;1,4)-β-glucan concentration and composition in barley grain to meet the requirements of specific end-users.

Grain (1,3;1,4)-β-glucan concentration in barley is influenced by both genetic (Powell et al. 1985), and environmental factors (Morgan and Riggs, 1981; Perez-Vendrell et al., 1996). The trait is inherited in a quantitative manner and can be associated with Quantitative Trait Loci (QTLs) located on all the seven barley chromosomes (Han et al., 1995; Baum et al., 2003; Islamovic et al., 2013; Molina-Cano et al., 2007). However, at present, no barley chromosomes have been associated with (1,3;1,4)-β-glucan structure elements such as DP3:DP4 ratio or concentrations of cellotrioses and cellotetroses.

The study of quantitative traits in crops has rapidly evolved during the last three decades, due to progress made in DNA markers technology and their ability to detect genotypic differences. Most of the QTL mapping studies use segregating progeny derived from a cross between parents with contrasting phenotypes. Various technologies are available to detect QTLs on the genome and association mapping via linkage disequilibrium (LD) has shown the potential to resolve QTLs to individual (candidate) genes (Oraguzie et al., 2007). Recently developed high-throughput DNA chip technology, diversity array technology (DArT®) and genotype by sequencing (GBS) methods have increased the availability of barley markers for marker trait association (http://wheat.pw.usda.gov). In this study, two types of mapping populations were used to develop and verify functional markers for concentration of grain (1,3;1,4)-β-glucan and its structure.

1.2 Objectives

1) Develop and validate a perfect marker within *HvCslF6* for grain (1,3;1,4)-β-glucan;
2) Determine the heritability of (1,3;1,4)-β-glucan fine structure;
3) Use association mapping to identify novel markers for (1,3;1,4)-β-glucan concentration in barley.

CHAPTER 2

REVIEW OF LITERATURE

2.1 Grasses

Grasses are members of the Poaceae (Gramineae) family of higher plants and include agronomically important cereal species such as corn, rice, wheat and barley (Gaut, 2002) that supply around 50 % of calories in human diet. In addition, domestic animals are raised on diets wholly or partially consisting of grasses and / or cereal grains. All grasses likely evolved from a common ancestor, making the *Poaceae* family monophyletic (Devos, 2005). This family includes approximately 10,000 species classified into 700 genera. Four main subfamilies make up 90 % of all grasses. The Chloridoideae, which include finger millet and teff, is recognized by their C4 anatomy and the structure of their microhairs. The Bambusoideae comprise of the woody and herbaceous bamboos, characterized by asymmetrically lobed mesophyll cells. The Panicoideae have paired flowers with the upper one generally being hermaphrodite and the lower one reduced or staminate. Maize, sorghum, common millet and fox-millet are members of this subfamily. Pooideae subfamily was originally shown by Russian cytogeneticist Avdulov to contain grasses with a base number of seven chromosomes (x=7), but recent genetic studies have included several other grasses (Kellogg, 2001). Wheat, oats, rye and barley along with a number of lawn and pasture grasses are members of Pooideae. The *Triticeae* tribe of Pooideae consists of 350 to 500 species (Bothmer et al., 1995), among which several important cereal and forage crops such as wheat (*Triticum* spp.), barley (*Hordeum vulgare L.*), rye (*Secale cereale L.*) and crested wheatgrass (*Agropyron cristatum*) are included. However, the taxonomic delimitation of the tribe is not fully resolved (Barkworth, 1992).

Within *Hordeum* four basic haplomes occur: the H haplome in *H. vulgare* and *H. bulbolsum*, the Xa haplome in *H. marinum*, the Xu haplome in the *H. murinum* complex, and the I haplome in all other *Hordeum* species. Beyond division by haplome, success rates of various *Hordeum* species crosses delineate three gene pools. The primary gene pool of *H. vulgare* subsp. *vulgare* consists of all forms of cultivated barley and the wild progenitor (subsp. *spontaneum*). This pool has no sterility barriers and can exchange genes easily (von Bothmer and Komatsuda, 2011). Modern barley varieties (*Hordeum vulgare* subsp. *vulgare)* are direct descendents of

Hordeum vulgare ssp. s*pontaneum* C. Koch. Cultivated barley and *H. spontaneum* have similar plant morphology but the cultivared form has a shatter resistant rachis, broader leaves, larger grains, and shorter stems and awns (von Bothmer and Komatsuda, 2011). Modern barley (*Hordeum vulgare* L.) is the result of centuries of artificial selection (Newman and Newman, 2006).

Barley was once thought to be composed of a non-homogeneous group of morphologically similar but unrelated entities with distinct shattering and spike types due to two complementary brittle rachis (*Btr*) genes (Zohary et al., 2012). The shattering characteristic of *H. spontaneum* is caused by the brittle rachis; spikes disarticulate above each rachis node to form wedge-shaped spikelets, which helps seed dispersal in the wild (von Bothmer and Komatsuda, 2011). A second form of brittle rachis phenotype is the breaking of the rachis leading to complete loss of the spike (Kandemir et al., 2000). Disarticulation scars in wild barley are smooth, whereas in cultivated barley, threshing produces rough scars on grains broken from rachis segments (Tanno and Willcox, 2006). Mutations in two tightly linked genes, *Btr1* and *Btr2*, on the 3H chromosome resulted in shatter resistant barley (Azhaguvel et al., 2006). All cultivated forms of barley carry the recessive alleles for one or both *Btr* genes.

2.2 Origin of barley

Archaeological evidence has shown that barley (*Hordeum vulgare* L.) was one of the earliest cultivated cereal grains, and directions on how to successfully grow barley have been found inscribed on clay tablets from ancient Sumer in Lower Mesopotamia dating from 1700 B.C. (Harlan, 1979). The origin of cultivated barley remains controversial with two separate centers of origin proposed (Bothmer and Komatsuda, 2012). Initially, the Fertile Crescent in the region occupied presently by portions of Israel, Iran, Iraq, Turkey and Syria (Harlan, 1979) was proposed as the sole center of origin, but recent genomic studies have also suggested regions in the Far East including Northern India and Tibet as a separate area of origin (Igartua et al., 2013). The fact that barley was domesticated not once but twice is a testament to its importance in modern agriculture. Energy and nutrients derived from barley directly or indirectly (as livestock feed) make up a major portion of daily human diet.

2.3 Production and distribution of barley

Globally, barley is the 12th most important crop in terms of production and the fourth ranked cereal after maize, rice and wheat in terms of consumption, harvested area and production (FAO, 2014). In 2013, approximately 145 million metric tonnes of barley were produced worldwide. In the same year, 10.2 million metric tonnes were produced in Canada, making it the fourth largest commercial barley producer in the world after Russia, Germany and France (FAO, 2014). One reason for the success of barley as a crop worldwide is its ability to grow in diverse environments (McIntosh et al., 1993).

Cultivated barley has the one of the widest ecological ranges among the cereals. As a cool climate crop, barley grows best in temperatures between 15 - 30 °C and can tolerate colder or warmer temperatures in low humidity environments. Although barley prefers to grow in cool climates, it is not particularly winter hardy. Therefore, barley is grown as a summer crop in temperate areas and as a winter crop in tropical climates. Barley has a shorter growing season than wheat and oats, matures earlier, has evolved to use less heat units than other crops and can be grown at higher altitudes than oats and wheat, giving it a wider ecological range than any other cereal. Thus, barley can be grown in areas of high altitude such as the steppes of Tibet or above the Arctic Circle in Alaska, Finland and Norway. Among the small grain cereals, barley has the lowest transpiration rate and requires the least amount of water. Barley is better adapted to drought, through high water use efficiency, than many other cereal crops. Barley can tolerate annual rainfalls ranging from 190 - 1760 mm but it is susceptible to water-logging and does best in areas with annual precipitation of 500 - 1000 mm (van Gool and Vernon, 2006). It also responds well in irrigated areas (Harlan, 1979). Barley is not well suited to tropical warm, humid environments (Nevo, 1992). Excessive heat after anthesis can reduce grain weight and negatively affect malting characteristics (van Gool and Vernon, 2006). Barley does best on light or sandy loam soils and can tolerate soils which are too light or coarse textured for wheat. Barley has a higher tolerance to salinity and alkalinity than most other cereals. Despite this, barley is susceptible to soil acidity and aluminum toxicity and usually thrives on soils with a pH of 7 - 8. Malting grade barleys are grown on soils with low nitrogen to obtain optimal protein content (van Gool and Vernon, 2006).

2.4 Barley types and classifications

2.4.1 Two-row and six-row types

Barley is classified by the number of kernel rows in a spike. Two forms are most commonly cultivated; two-row and six-row barley. Spikelets are arranged in triplets which alternate along the rachis. In *H. spontaneum* and two-row barley, only the central spikelet is fertile and the lateral spikelets are sterile. All three spikelets are fertile in six-row barley. The number of fertile spikelets on the ear is controlled by at least five genetic factors: *vrs1, vrs2, vrs3, vrs4* and *int-c* (Koppolu et al., 2013); the predominant factor is in most cases the *vrs1* locus on the 2H chromosome (Powell et al., 1990). The six-rowed spike gene (*vrs1*) is genetically recessive and originates from a mutation in a homeobox gene (*HvHox1*) containing a leucine zipper motif (Komatsuda et al., 2007). The cultivated forms of barley (*Hordeum vulgare L.*) may also include a group of barley lines derived from crosses between two- and six-row barley that were previously denoted as *H. intermedium* (Jui et al., 1997). For full development of the lateral spikelets in six-row barley, the additional action of the intermedium gene (*int-c*) is needed. The gene underlying the *int-c* locus is an ortholog (*HvTB1*) of the maize domestication gene TEOSINTE BRANCHED1 (Ramsay et al., 2011). A combination of recessive alleles in the *Vrs1* allele and dominant *int-c* allele are needed to complete the shift from two-row to six-row cultivars (Gymer, 1978). Fertility and an increase in the size of lateral spikelets are considerably enhanced by *int-c.a* (the dominant *int-c* allele) in combination with *Vrs1vrs1* heterozygotes (Lundqvist and Lundqvist, 1987). Haplotype analysis of the *vrs1* region indicates that six-row varieties have been independently selected from domesticated two-row varieties at least three times throughout history (Komatsuda et al., 2007).

Two-row barley grains are more circular, have higher test and grain weight, and are higher in starch and protein than six-row kernels (Tanno and Takeda, 2004; Marquez-Cedillo et al., 2001). Two-row genotypes are more lodging resistant and grains have lower amounts of (1,3;1,4)-β-glucan compared to six-row varieties (Berry et al., 2006). The higher number of kernels per spike in the six-row cultivars makes them attractive for production as yield is 20 - 27 % higher than in two-row varieties. The increased number of kernels per spike in six-row varieties is also associated with a reduced amount of tillering; thus spikes per plant compensate for the reduced number of seeds per spike in two-row compared to six-row genotypes (Lundqvist et al., 1997; Kirby and Riggs, 1978).

2.4.2 Hulled and hulless grain

A second classification of barley is based on the adherence of the hull to the kernel. Wild barley and most cultivated forms are hulled (covered caryopsis) with the palea and lemma firmly adhered to the pericarp epidermis at maturity. Hulless varieties have hulls which are easily removed during threshing. The hulless or "naked" trait is controlled by a single locus *nud* (for *nudum*) on chromosome arm 7HL. Positional cloning and analysis of the *nud* locus has shown that the hulless lines have a 17 kb deletion which removes a gene encoding an ethylene response factor (ERF) transcription factor (Taketa et al., 2004; 2008). This ERF gene is express predominantly on the ventral side of the testa in hulled varieties. An absence of staining by Sudan black B in hulless barley has lead to the postulation that the identified ERF family transcription factor also control lipid synthesis at the testa (Kakeda et al., 2011).

2.4.3 Market classes

In Canada, barley is divided into three classes based on end-use: malting, general purpose and food. The two principal markets for Canadian barley are malt and animal feed industries. The faster fermentation rate of barley compared to other cereals makes it an attractive grain for maltsters. Only varieties within the malting barley variety designation list are eligible for the malting grades. Malting barley may be covered (hulled) or hulless varieties and are selected for uniform germination, production of an extract that is low in protein and low in soluble fiber. Two-row barley usually has a lower protein content and higher fermentable sugar content than six-row barley and is most commonly used in English style ale and traditional German beer. Six-row barley is commonly used in American lager style beers, along with adjuncts such as corn and rice. If barley cannot be sold at a premium for malting and brewing, it becomes general purpose grade routinely destined for livestock feed. Only about 20 percent of malting barley production is actually selected for malting each year. The other 80 percent is used general purpose grades. General purpose grades can include hulled and hulless barley not selected for malting, food or registered feed varieties. Registered malting varieties can be relegated to general purpose, but registered feed barley is not considered suitable for malting and brewing and must be used for livestock feed. Hulless barley is used primarily for non-ruminant animal feed, but is increasingly marketed for human consumption (http://www.agr.gc.ca). Food barley can be any barley variety (hulless or hulled) that has been selected for a food market. In the last several

years there has been a growing interest from food processors to incorporate barley in food products due to the fiber and nutritional composition of the grain.

Globally, approximately two percent of barley produced is used for food (Baik and Ullrich, 2008). Uniform sized, thin-hulled, bright yellow-white, plump, medium hard and clean barley is usually selected for food use. Only a few hulless barley genotypes meet these requirements with minimal processing. Barley is nutritionally dense with high carbohydrate concentrations, high dietary fiber, moderate amounts of protein and is a relatively good source of selenium, phosphorous, copper and manganese (Ames et al., 2006). Consumption of barley grain products has a positive effect on human health as it reduces transit time of fecal matter by increasing bulk and lowering the frequency of hemorrhoids and colon cancer (Tsai et al., 2004). Meals containing (1,3;1,4)-β-glucan also slow down absorption of carbohydrates in the gastro intestinal tract and stimulate reverse cholesterol transport (Bourdon et al., 1999). Furthermore, it has been shown that fermentation of insoluble fiber from barley sources in the large intestine produces short-chain fatty acids such as butyric acid which maintain colon health (Behall et al., 2004). Propionic acid which inhibit HMG-CoA reductase and thereby lowers cholesterol biosynthesis in the liver (Erkkila et al., 2005) and acetic acid which provide fuel for liver and muscle cells (Liu, 2004) are also produced upon barley fiber consumption. As consumption of diets rich in fiber is associated with a reduction in the risk of developing coronary heart disease (Anderson, 1995; Pereira et al., 2004), the US Food and Drug Administration (FDA) has allowed whole grain barley and barley-containing products to carry a claim that they reduce the risk of coronary heart disease (FDA News Release, 2005). This health declaration is largely attributed to the (1,3;1,4)-β-glucan content in barley grain.

2.5 Major structural components of barley grain

The barley kernel can be subdivided into three main components: 1) bran, 2) endosperm and 3) embryo. Bran consists of the hull and aleurone layer. On average, 14 % of the kernel weight can be attributed to bran. According to the USDA Nutrient Database, bran is low in protein (< 3 %) and trace minerals (3 - 5 %) but contains relatively high amounts of B vitamins (3 - 6 %) and fiber (4.5 - 15 %). The removal of the hulls from grains causes a relative increase of nutrients in the caryopsis due to the reduction in fiber content in the remaining fraction (Table 2.1) (McGuire and Hockett, 1981; MacGregor and Fincher, 1993, Bhatty and Rossnagel, 1988).

The hull is the adhered lemma and palea, which comprises from 9 - 13 % of the grain by weight and consists mainly of cellulose, lignin, arabinoxylans, and silica (Henry, 1988). The aleurone layer is the outer two or three cell layers of the grain underneath the hull. These cells are thick-walled and contain lipids, proteins and minerals. The aleurone layer contains unique structures called 'aleurone grains'. These are storage bodies which contain two types of inclusion bodies called globoids and crystalloids containing phytin and 8S globulin storage proteins, respectively (Gubatz and Shewry, 2011). The cell wall of the aleurone layer is composed mainly of arabinoxylans (~ 67 %) and (1,3;1,4)-β-glucan (~ 25 %) (Bacic and Stone, 1981).

The endosperm is by far the largest component of a barley kernel constituting approximately 83 % of the grain by weight. The endosperm begins at the sub-aleurone layer, which comprises the first two to three cell layers beneath the aleurone layer. The subaleurone layer cells can be distinguished from the endosperm cells as they are rich in proteins but poor in starch granules. Below this layer, the cells have higher numbers of starch granules and less protein (Gubatz and Shewry, 2011). The endosperm is composed of carbohydrates (70 - 77 %) and proteins (12 - 16 %) with trace amounts of minerals and vitamins (USDA Nutrient database). Starch and the cell wall polysaccharides, such as (1,3;1,4)-β-glucan and arabinoxylans are the major carbohydrates of the endosperm. The endosperm contains very little raffinose, sucrose or monosaccharide sugars. Barley differs from other cereals in that the (1,3;1,4)-β-glucan content of the endosperm is similar to that of the grain as a whole (Henry, 1988).

The embryo (or germ) is located on the dorsal side of the caryopsis near the end attached to the rachis. The embryo contains the tissues necessary for new plant growth. The main storage proteins in the embryo are the 8S globulin storage proteins (Heck et al., 1993). While starch is present in the embryo during development, it is almost completely absent in the mature embryo (Duffus and Cochrane, 1993). Raffinose represents up to nine percent of the dry weight of a barley embryo while sucrose represents 12 to 15 % on a dry weight basis (Henry, 1988). Oil can account for as much as 20 % of the embryo dry weight. Nearly 90 % of the oils found in the germ are in the form of triglycerides.

Table 2.1. Composition of hulless and hulled barley grain.

Compound	Hulless		Hulled	
	Mean*	Range*	Mean*	Range*
Starch	63.4	60.1 - 75.2	58.2	57.0 - 65.4
Protein[1]	14.1	12.1 - 16.6	13.7	8.2 - 15.4
Alanine	0.47		0.44	
Arginine	0.64		0.60	
Cytosine	0.31		0.28	
Glycine	0.44		0.42	
Glutamic acid	3.27		2.98	
Histidine	0.28		0.26	
Isoleucine	0.46		0.43	
Lysine	0.41		0.41	
Methionine	0.28		0.21	
Phenylalanine	0.73		0.68	
Proline	1.43		1.32	
Serine	0.57		0.54	
Threonine	0.45		0.42	
Tryptophan	0.23		0.22	
Tyrosine	0.42		0.37	
Valine	0.63		0.59	
Sugars	2.9	0.7 - 4.2	3.0	0.5 - 3.3
Lipids	3.1	2.7 - 3.9	2.2	1.9 - 2.4
Fiber	13.8	12.6 - 15.6	20.2	18.8 - 22.6
Ash	2.8	2.3 - 3.5	2.7	2.3 - 3.0

[1]-Kjeldahl method (Nx6.25)

*-(g / Kg dry weight)

Source: Adapted from Åman and Newman, 1986; Newman and Newman 2005

2.6 Nutrient composition of barley grain

2.6.1 Lipids

Barley grain contains 1 - 3 % lipids depending on the variety (Jacobsen et al., 2005) but levels as high as 5.3 % have been reported for some genotypes (Bhatty, 1982). Barley contains five major fatty acids, including palmitic (C16:0), stearic (C18:0), oleic (C18:1), linoleic (C18:2), and linolenic (C18:3) acids (Moreau et al., 2007). Linoleic acid is the most abundant fatty acid in barley, but the concentrations of the other five major fatty acids varies with genotype. In most barley varieties, linoleic acid is followed by palmitic, oleic, linolenic, and stearic acids in decreasing amounts. Several other fatty acids are also present, including myristic (C14:0), palmitoleic (C16:1), arachidic (C20:0), eicosanoic (C20:1), behenic (C22:0), erucic (C22:1), lignoceric (C24:0), and nervonic (C24:1) acids. Neither reduced phytate concentration nor environment influences lipid content or composition. However, genotype can significantly influence fatty acid profiles and total lipid content in barley grains (Bravi et al., 2012).

Several factors influence fat composition of barley kernels. Hulless varieties generally have about one percent higher lipid concentrations than hulled varieties (Åman et al., 1985). Starch composition of the endosperm can affect the lipid concentrations, as lipids in the endosperm are present as lysophospholipids in complex with amylose. Therefore, lipids concentration increases with higher concentration of amylose in grain starch (Morrison, 1993).

2.6.2 Proteins

Barley kernel protein concentration is an important quality consideration for malting and brewing, food and feed industries. Barley grain protein usually ranges from 6 - 20 % on a dry weight basis (Jadhav et al., 1998; Holtekjølen et al., 2008) and is similar to wheat (10 - 15 %), but higher than maize (7 - 13 %) (Simmonds, 1978). However, genotype and environment affect protein accumulation in barley grain (Aniskov et al., 2008). Hulless barley has on average 1.1 % higher protein concentration than hulled barley (Ullrich, 2002). For malting barley, an optimal range of 9 - 12 % protein is desired (LaFrance and Watts, 1986).

The main storage proteins of the *Tritceae* tribe are prolamins, of which hordein storage proteins in barley grain make up 36 - 49 % of the total protein content. In the grain, hordein and protein contents are generally strongly influenced by soil nitrogen availability (Kirkman et al., 1982) as hordeins are relatively high in glutamine and asparagine in addition to proline (Rastogi

and Oaks, 1986). The storage proteins in barley are widely characterized as low molecular weight (LMW) and high molecular weight (HMW) hordeins.

LMW proteins are a family of trypsin and α-amylase inhibitors that range in weight from 16.5 to 22 kDa (Gubatz and Shewry, 2011). LMW hordeins have been referred to as avenin-like proteins or A-hordeins based on sequence similarities (Kan et al., 2006) but was originally denoted as chloroform methanol (CM) proteins by Aragonillo et al. (1981). Genes encoding the CM proteins are located on 2H, 3H, 4H, 6H and 7H chromosomes. CM proteins consist of seven major subunits BTAI-CMa, BTAI-CMb, BTAI-CMc, BTAI-CMd, BTAI-CMe, BMAI-1 and BDAI-1 which can occur in monomers, dimers or tetramers. Monomeric forms of BMAI-1 and BT1-CMe are inhibitors of α-amylase and trypsin while the dimer and tetrameric forms inhibit only α-amylase. None of the CM proteins inhibit endogenous barley amylases but are only active against amylases from different organisms (Carbonero and Garcia-Olmedo, 1999).

HMW hordeins range in weight from 35-100 kDa and can be further subdivided into B, C, D, γ and γ3 hordeins (Gubatz and Shewry, 2011; Anderson et al., 2013). B-hordeins can account for 70 - 80 % of the total hordeins followed by C hordeins at 10 - 20 % (Kaczmarczyk et al., 2012). D, γ and γ3 hordeins make up a minor fraction of the storage proteins. D-hordeins are found in the central cells of the starchy endosperm while most of the other storage proteins are located in the sub-aleurone layer (Gubatz and Shewry, 2011). Loci for the B, C, D and γ hordeins are the *hor-2* locus on 5HS, *hor-1* locus on 5HS, *hor-3* locus on 5HL and *hor-5* locus on 1HS (Shewry, 1993). Hordeins are known to be a causative agent in celiac disease (Sollid, 2000), an autoimmune enteropathy leading to damage of the gastrointestinal mucosa (Rallabhandi, 2012).

Barley, similar to other cereals, is low in essential amino acids including lysine, methionine, tryptophan and threonine for animal and human nutrition (Newman and Newman, 1992). Low lysine levels in cereals are a major concern when balancing nutrition for animal feeds (Foster and Prentice, 1987). Protein content can influence mineral content, such as calcium, phosphorus iron and copper.

2.6.3 Minerals

The barley kernel can vary in mineral (ash) content from 2 - 3 % depending on the genotype (Liu et al., 1975; Marconi et al., 2000). Whole grain barley contains several important macronutrients such as P, K, Mg, and Ca and micronutrients such as Fe, Zn, Mn, and Cu (Table

2.2). The majority of the minerals in barley are found in the outer layers of the kernel, aleurone and embryo (Liu et al., 2007; Stewart et al., 1988; Ockenden et al., 2004). About 15 - 20 % of the Fe is located in the pericarp of the mature barley grain, whereas the endosperm (including the aleurone) contains about 70 % and the embryo only 7 - 8 % (Duffus and Rosie, 1976).

Most of the phosphorus in barley is present as phytate. The phytate molecules serve as mineral reserves incorporated into aleurone grains within the aluerone layer. Phytate is a mixed salt of phytic acid (*myo* -inositol - 1,2,3,4,5,6 - hexa *kis* phosphate) which has become a major focus in barley breeding programs. The chelating properties of phytic acid negatively impacts mineral bioavailability, as bound cations being less available for absorption by monogastric animals. The phosphate groups of phytic acid form negatively charged sites, which can form salts with mineral cations, such as K^+, Mg^{2+}, Ca^{2+}, Mn^{2+} , Zn^{2+}, Ba^{2+}, or Fe^{2+} (Lott et al., 2000). When the grain is used as a major component in diets for human nutrition or animal feed, the high intake of phytic acid can result in deficiencies of iron, zinc, magnesium, and / or calcium.

2.6.4 Vitamins and minor compounds

Cereals are well known sources of vitamins in the human diet. With the exception of vitamin A, D, K, B_{12} and C, barley is a good source of vitamins and choline. Of all the cereals, barley contains the highest amounts of fat soluble vitamin E (tocols) with all eight isomers; four tocopherols: α-T, β-T, γ-T, δ-T and four tocotrienols: α-T3, β-T3, γ-T3, δ-T3 (Morrison, 1978). The majority of tocopherols are found in the embryo, whereas the tocotrienols are evenly dispersed through the grain (Peterson, 1994). Vitamin B_1 (thiamine) is found primarily in the aleurone and scutellum, B_2 (riboflavin) is found in the aleurone and endosperm and B_3 (niacin) is found mostly in the aleurone layer (Newman and Newman, 2008). Barely contains the highest concentration of niacin (B_3) of all the major cereal crops. Vitamin B_9 (folate) is higher than that of both oats and wheat (Anderson et al., 2008).

A range of phyotchemicals including phenolic acids, flavanoids, and alkyl resorcinols are found in barley. The two major phenolic acids found in barley are pro-anthocyanidins and anthocyanidins. The pro-anthocyanidins can cause haze formation in beer (Jende-Strid 1993) and possess high antioxidant activity and potential health benefits (Beecher, 2004). Anthocyanidins occur in the pericarp and aleurone of pigmented grains and can influence barley color making it

Table 2.2. Mineral composition of hulled barley (dry matter basis).

Minearal	Mean (g / 100g)	Range (g / 100g)
Calcium	0.05	0.03 - 0.06
Phosphorus	0.35	0.26 - 0.44
Potassium	0.47	0.36 - 0.58
Magnesium	0.14	0.10 - 0.18
Sodium	0.05	0.01 - 0.08
Chlorine	0.14	0.11 - 0.18
Sulfur	0.20	0.16 - 0.24
Silicon	0.33	0.15 - 0.42
Copper	6.25×10^{-4}	$2.0 - 9.0 \times 10^{-4}$
Iron	45.7×10^{-4}	$36.0 - 85.0 \times 10^{-4}$
Manganese	19.4×10^{-4}	$17.0 - 20.0 \times 10^{-4}$
Zinc	34.4×10^{-4}	$19.0 - 35.0 \times 10^{-4}$
Selenium	0.4×10^{-4}	$0.2 - 0.5 \times 10^{-4}$
Cobalt	0.7×10^{-5}	$0.05 - 0.10 \times 10^{-4}$

Source: Adapted from Newman and Newman, 2008.

range from yellow to purple, violet, blue or black (Baik and Ullrich, 2008). The seed coat contains small amounts of alkylresorcinols. Alkylresorcinols have been associated with a number of biological effects ranging from reducing weight gain in livestock, when consumed in large quantities, to antioxidant and anticancer action in human, when consumed in small quantities (Ross et al., 2003).

2.6.5 Carbohydrates

Barley kernels consist of roughly 78 - 83 % carbohydrates on a dry weight basis. Only 2 - 3.5 % of the carbohydrates are simple sugars or oligosaccharides (Figure 2.1). The majority of the carbohydrates are stored in the form of starch granules in the endosperm. The remaining polysaccharides are largely found in the cell wall in the form of arabinoxylans, cellulose and (1,3;1,4)-β-glucans (Figure 2.2) (MacGregor and Fincher, 1993).

Monosaccharides such as glucose, fructose, fucose, arabinose, xylose, ribose, deoxyribose, galactose and mannose are produced in the grain and incorporated into oligo and polysaccharides, glycosides, glycolipids or glycoproteins (Holtekjølen et al., 2008). Free glucose makes up 0.1 - 1.4 % of the dry weight of barley and fructose ranges from trace amounts to 0.5 % (Åman et al., 1985). Most other sugars are rapidly converted or incorporated into larger carbohydrates.

The most abundant disaccharides in barley are sucrose and maltose. Sucrose concentrations range from 0.74 - 3.9 % in the mature caryopsis on a dry-weight basis (MacGregor and Fincher, 1993; Åman et al., 1985) but can accumulate to as high as seven pecrent in waxy cultivars (Batra et al., 1982). The majority of free sucrose is found in the embryo (80 %). Sucrose is an important precursor for starch and non-starch polysaccharides such as callose, (1,3;1,4)-β-glucan and cellulose. Maltose accumulates in the endosperm as a result of starch amylolytic activity and constitutes 0.1 - 0.2 % of grain dry weight (Sopanen and Lauriere, 1989). The maltose concentration can be higher in waxy genotypes (Nielson et al., 2009) or in plants exposed to high temperatures during grain filling (Hogy et al., 2013).

Oligosaccharides are polymers of 3 - 20 monosaccharides units (Chibbar et al., 2004). Raffinose represents 0.16 - 0.80 % dry weight in the kernel (MacGregor and Fincher, 1993; Andersen et al., 2005) and is predominantly (80 %) found in the embryo (Andersen et al., 2005)

Monosaccharides

α-D-Glucose CID: 79025	β-D-Glucose CID: 64689	β-D-Fructose CID: 439709	α-D-Xylose CID: 6027	α-D-Galactose CID: 439357
α-D-Fucose CDI: 444200	α-L-Arabinose CID: 469731	α-D-Mannose CID: 185698	α-D-Ribose CID: 5779	D-(-)-Deoxyribose CID: 9828112

Disaccharides

Sucrose CID: 5988	Maltose CID: 6255

Oligosaccharides

Raffinose CID: 10542	Fructan (Levan) CID: 440946
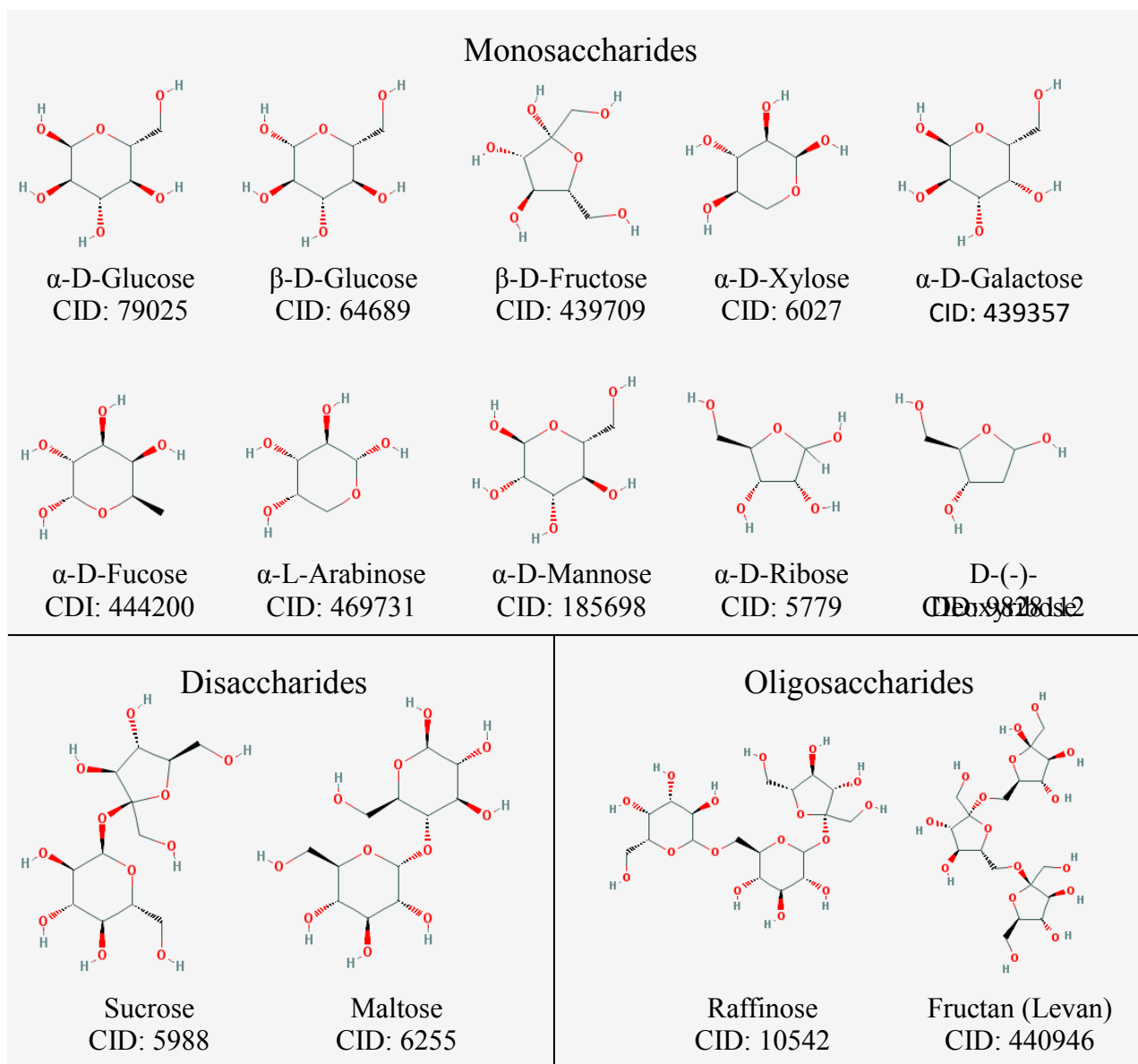
Figure 2.1. Chemical structure of common mono, di and oligosaccharides found in barley. Structure database source: http://www.ncbi.nlm.nih.gov/pccompound.

17

Cell wall polysaccharides

Starch polysaccharides

Arabinoxylan
(standard)

Amylose

Arabinogalactan
(standard)

Main chain

Amylopectin

(1,3;1,4)-β-glucan

Cellulose

β-(1,3)-xylan
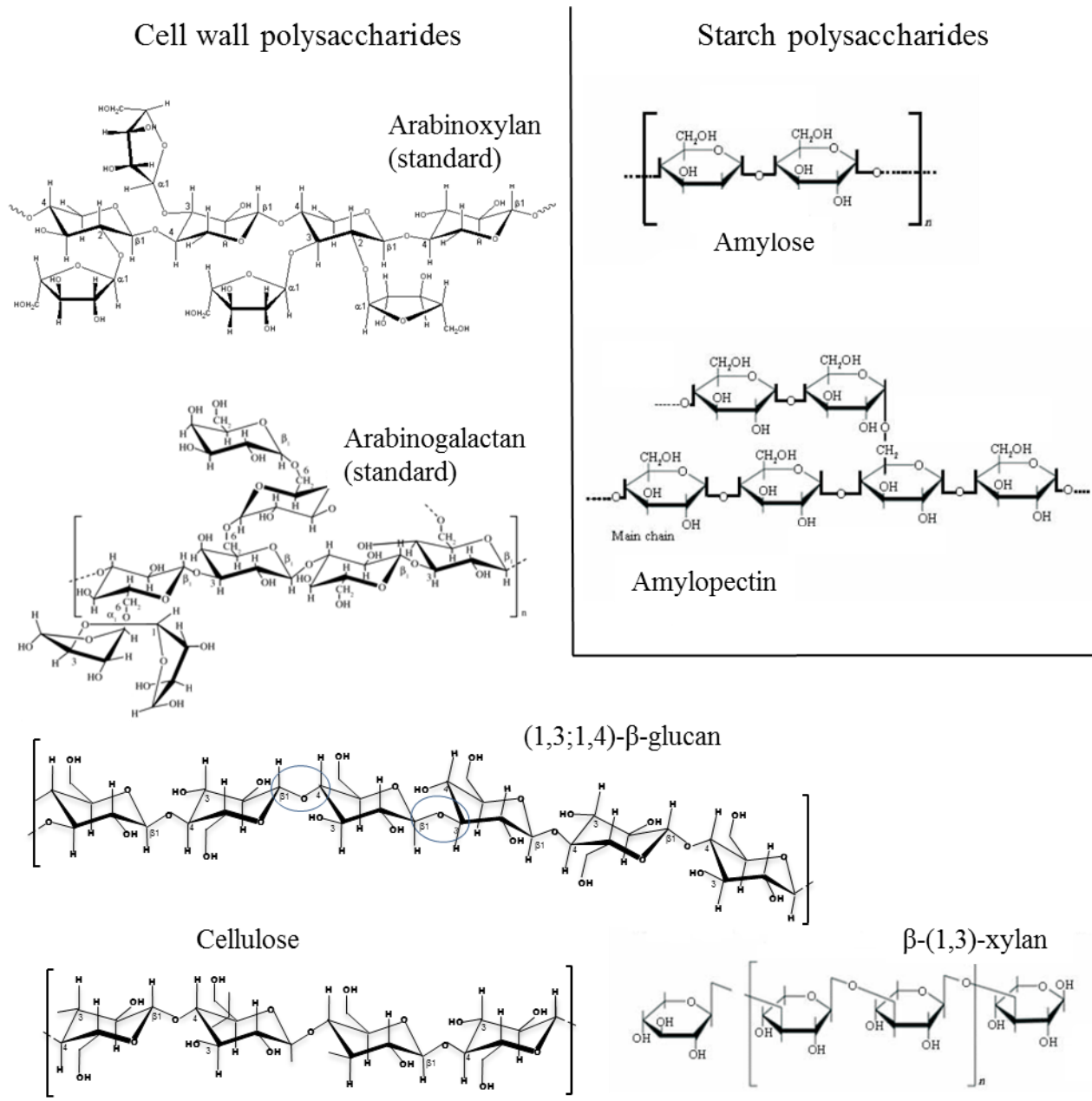
Figure 2.2. Chemical structure of select cell wall and starch polysaccharides.

Table 2.3. Select vitamins and minor compounds in barley.

Compound	Mean (µg / g)	Range (µg / g)
Total tocol content	55.0	46.2 - 68.8
Alkylresorcinols	55.0	32 – 103
Phenolic acids	463	254 – 675
Thiamin	5.2	4.0 - 6.5
Riboflavin	1.8	1.2 - 2.9
Niacin	63.3	46.0 - 80.0
Pantothenic acid	5.1	2.8 - 8.0
Biotin (B7)	0.14	0.13 - 0.15
Sterols	1049	899 – 1153
Choline	1290	920 – 2200

Source: Adapted from Newman and Newman, 2006; Andersson et al., 2008)

where it plays a role during desiccation and germination (Sreenivasulu et al., 2008). Early determinations of fructan concentrations in barley kernels reported 0.019 - 0.97 % per dry weight basis (Henry, 1988), but higher concentrations (1.5 - 2.0 %) were obtained in a recent study (Hogy et al., 2013). Fructans are oligosaccharides of fructosyl residues which are comprised of up to 10 fructosyl units (MacLeod, 1953) and are thought to enhance drought tolerance in barley (Janthakahalli, 2004).

2.6.6 Starch

Monosaccharide polymers of greater than 20 units are called polysaccharides and are synthesised for both storage and structural purposes in plants (Chibbar et al., 2004). Starch, being the major carbohydrate reserve, is the most extensively studied and economically important component of the barley kernel. Grain starch concentrations can be up to 75 % in some hulless varieties but normally ranges from 45 - 65 % in barley grain. Protein, non-starch polysaccharides and lipid concentrations tend to be low in varieties producing high starch concentrations in the endosperm (Newman and McGuire, 1985).

Amylose and amylopectin are two distinct glucan polymers of which starch is comprised. Amylose is a high molecular weight (up to 10^6 kDa), linear polymer of α-(1-4) linked D-glucose units with minimal (> 0.5 %) α-(1,6) linkages forming branch points (BeMiller and Whistler, 2009; Hung et al., 2008; Takeda et al., 1990). In contrast, amylopectin is highly branched with approximately five percent of all linkages made up of α-(1,6) linkages forming branch points. Each branch consists of 20 - 30 glucose units making the amylose molecule nearly spherical when looked at as a whole (Hizukuri, 1985). Barley cultivars most often contain an amylopectin to amylose ratio of 3:1 but certain genotypes diverge greatly from this ratio. The term 'waxy barley' is used for genotypes with a high amylopectin concentration (95 - 100 %) (Bhatty and Rossnagel, 1997), whereas barley genotypes with high amylose concentrations (40 - 70 %) are described as 'amylotype' (Delcour and Hoseney, 2010).

2.6.7 Dietary fiber

Dietary fibre consists of a mixture of components with a varying degree of solubility. Due to differences in definitions used by countries and research groups, a committee was appointed by the American Association of Cereal Chemists International (AACCI) to define

fiber. In 2001, the committee defined dietary fibre as "the edible parts of plants or analogous carbohydrates that are resistant to digestion and absorption in the human small intestine with complete or partial fermentation in the large intestine" (DeVries, 2001). The committee also concluded that dietary fibre include polysaccharides, oligosaccharides, lignin and associated substances. The most common dietary fibre polysaccharides in cereals are cellulose, arabinoxylan, and (1,3;1,4)-β-glucan.

2.7 Cell walls

The plant cell wall is a complex, composite material made of polysaccharides networks and proteins. It is an important structural element providing mechanical support, defining plant and tissue morphology. The cell wall is responsible for responding to injury or threat, allowing for transport of nutrients and relaying information throughout the plant (Pilling and Hofte, 2003). Two major classes of cell walls are present in plants. Dicots and the non-commelinoid monocots predominantly have type I cell walls, that contain equal amounts of cellulose and cross-linking xyloglucan with minor amounts of arabinoxylans, glucomannans and galacto-glucomannans embedded in a pectin matrix of homogalacturonans and rhamnogalacturonon I and II (Yong et al., 2005). The cell walls of grasses and commelinoid monocots have type II cell walls that contain cellulose microfibrils similar to type I cell walls, but (1,3;1,4)-β-glucan and glucuronoarabinoxylans (GAX), and not xyloglucan, are the major tethering molecules. The mature cell walls are pectin poor and possess extensive interconnecting networks of phenylpropanoids (Yong et al., 2005).

Cell wall polysaccharides represent roughly ten percent of the total carbohydrate content of a barley kernel. The primary polysaccharide in cell walls of a barley kernel is (1,3;1,4)-β-glucan. Within the endosperm cell wall, (1,3;1,4)-β-glucan accounts for 75 % of the polysaccharide content, 20 % is contributed by arabinoxylans and 3 - 4 % cellulose (Fincher and Stone, 1986).

2.7.1 Cellulose

Cellulose is the major structural polysaccharide of plants. It is the major component in husk and outer layers of cereal grains but makes up only 3 - 4 % of the total carbohydrates in the grain (Henry, 1988). It is an unbranched linear molecule composed of (1,4) linked β-D-glucose

residues, which easily associates with itself creating a durable crystal structure. This association and its high order make cellulose insoluble and therefore resistant to pathogen and enzymatic attacks.

2.7.2 Hemicellulose

The term "hemicellulose" is a general term applied to the polysaccharide components of plant cell walls other than cellulose, or to plant cell walls polysaccharides extractable by dilute alkaline solutions. The chemical structures of hemicelluloses consists of long chains of a variety of pentoses, hexoses, and their corresponding uronic acids. Pentosan is a general term used for hemicellulose and refers to polysaccharides yielding pentoses on hydrolysis. The pentosan content of barley grain ranges from 4 - 7 % and is generally derived from cell wall fractions. While the outer protective layers of the grain have high concentration of pentosans (Hashimoto et al., 1987), the endosperm contains only 22 % of total barley pentosans (Henry, 1987). The concentration of barley pentosans vary with the genotype, but are also highly influenced by environmental factors (Henry, 1986).

2.7.3 Arabinoxylans and arabinogalactans

Cereal arabinoxylans (AX) are a heterogeneous group with varying substitution patterns and degree of polymerisation (Vinkx and Delcour, 1996). A (1,4)-β-D-xylopyranosyl backbone gives this group its name but different substituents can be present depending on the plant tissue. The major molecules linked to the AX backbone are α-L-arabinofuranose residues attached by (1,3) and / or (1,2) glycosidic linkages. A second category of minor substituents bound to the xylan backbone are uronic acids attached to the C2 atom of the xylose residue (Fincher, 1975). The side chain substitutions may produce several combinations of sidegroups (Bengtsson et al., 1992; Hoffmann et al., 1991; Gruppen et al., 1993) thereby determining the solubility of arabinoxylan (Amado and Neukom, 1985). Hulless barley genotypes have a significantly higher substitution of arabinose in the AX than hulled grain (Holtekjølen, et al., 2008).

Arabinogalactans consist of a highly branched structure in which galactopyranosyl units are bound through (1,3) and (1,6) glycosidic linkages and are predominantly found in the endosperm. Only single arabinose units are β-linked to the galactose chain (Neukom and Markwalder, 1975; Amado and Neukom, 1985). The arabinogalactan polymer is covalently

22

bound to a hydroxyproline-rich protein forming an arabinogalactan glycoprotein (Fincher and Stone, 1974).

2.8 (1,3;1,4)-β-glucan

(1,3;1,4)-β-glucan is a linear polymer made up of glucose molecules that form either β-(1,3) or β-(1,4) linkages. The β-(1,4)-linked molecules are constituted primarily (~ 90 %) of non-randomly arranged subunits of cellotriose [Degree of Polymerization 3 (DP3)] and cellotetraose (DP4). Inclusion of higher MW subunits (DP5+) can account for approximately nine percent of (1,3;1,4)-β-glucan content (Lazaridou et al., 2004). The (1,3)-linkages cause irregularities in the structure of the molecule, which make the β-glucans partly soluble by preventing close packing of the chains (Jadhav et al., 1998; Jiang and Vasanthan, 2000). Reported apparent molecular weights (M_w) for barley (1,3;1,4)-β-glucan have a wide range from 1.5×10^5 to 2.5×10^6 (Irakli et al., 2004; Lazaridou et al., 2003; 2004).

(1,3;1,4)-β-glucan is rarely found outside of the grass family in the plant kingdom (Sørensen et al., 2008). Within the Poales four distinct types of cell wall confirmations based on (1,3;1,4)-β-glucan deposition have been recognized (Trethewey et al., 2005). *Bromeliaceae*, *Typhaceae* and *Sparganiaceae* families do not accumulate (1,3;1,4)-β-glucan in their cell walls. The cyperoid clade comprised of *Cyperaceae*, *Juncaeae* and *Thurniaceae* have trace amounts of (1,3;1,4)-β-glucan in cells with non-lignified cell walls. The *Restionaceae* subclade has low concentrations (0.1 % of cell wall content) in non-lignified cell walls. *Poaceae* subclade, which includes *L. multiflorum, Zea mays, Avena sativa* and *Hordeum vulgare* have the highest concentrations of (1,3;1,4)-β-glucan in their non-lignified cell walls. (1,3;1,4)-β-glucan is synthesised abundantly and transiently accumulated in primary cell walls of expanding vegetative organs, such as growing leaves, but it is rarely found in mature cells outside of the grain (Gibeaut and Carpita 1991; Fincher and Stone, 1986). The mature cell walls of the aleurone and starchy endosperm in cereals like barley are unusual in that fully expanded cells contain significant quantities of (1,3;1,4)-β-glucan and low levels of cellulose (Bacic and Stone, 1981).

In barley, (1,3;1,4)-β-glucans are deposited early in endosperm development, near the end of cellularization around four days after pollination (DAP). (1,3;1,4)-β-glucan appears along already formed and developing anticlinal and periclinal cell walls in early endosperm development (Wilson et al., 2006). Deposition occurs after the initial installation of callose and

cellulose which define the cell wall, but before the other major cell wall components such as hetero-(1,4)-β-D-mannan and arabino-(1,4)-β-D-xylan. (1,3;1,4)-β-glucan accumulates throughout the endosperm and becomes uniformly distributed by 10 DAP (Wilson et al., 2012). At 12 DAP, the endosperm cells closest to the aleurone layer show a marked decrease or complete absence of (1,3;1,4)-β-glucan. By 16 DAP, (1,3;1,4)-β-glucan production increases until all endosperm cells, including the sub-aleurone cell layers, contain some (1,3;1,4)-β-glucan. By 18 DAP, the aleurone cell layers are markedly lower in (1,3;1,4)-β-glucan compared to the starchy endosperm, and the starchy endosperm cells contain a uniform amount of (1,3;1,4)-β-glucan (Wilson et al., 2012). The aleurone layer can have cell walls with approximately 26 % w / w (1,3;1,4)-β-glucan and 71 % arabinoxylan, while (1,3;1,4)-β-glucan may account for as much as 75 % of total polysaccharides in endosperm cell walls (Fincher and Stone, 1986). The cell walls of starchy endosperm do not exhibit secondary cell wall thickening, but can accumulate enough (1,3;1,4)-β-glucan to account for up to 18 % of the total glucose content in the grain (Burton and Fincher, 2009). (1,3;1,4)-β-glucan concentrations in barley can range from very low (≤ 1 %) to up to 21 %, but normally vary between three to six percent (Munck et al., 2004; Kato et al., 1995; Holtekjølen et al., 2006).

2.8.1 (1,3;1,4)-β-glucan effects on malt and feed

Grain and / or wort (1,3;1,4)-β-glucan concentration is often included in the malting quality evaluation of barley. Wort is the liquid extracted from the mashing process during the brewing of beer or whiskey. High concentrations of (1,3;1,4)-β-glucan in the endosperm cell walls have a direct, deleterious effect on malting (Swanston et al., 1995). High accumulation of (1,3;1,4)-β-glucan in the cell wall may result in insufficient degradation of cell walls, hampering the diffusion of enzymes and kernel reserves needed for germination, therefore decreasing malt extract. Residual (1,3;1,4)-β-glucan will also cause highly viscous wort and chill haze in bright beer (Bamforth, 1982; Wang et al., 2004). Positive correlations have been found between (1,3;1,4)-β-glucan and milling energy requirements, increased pearling time or acid detergent fraction (Tohno-Oka et al., 2004; Laido et al., 2009). (1,3;1,4)-β-glucan is also problematic when barley is used as a feed for non-ruminants, especially poultry (Gohl et al., 1978). The same deleterious effects encountered in malting are thought to reduce nutrient absorption in pigs and poultry.

2.8.2 (1,3;1,4)-β-glucan and human nutrition

High (1,3;1,4)-β-glucan content is desirable for barley used as functional, bio-active ingredient in human nutrition (Wood and Beer, 1998). The Canadian Heart and Stroke Foundation recommend a daily intake of 21 - 38 g of fiber per day. Studies have shown as little as 3 g per day of soluble fibers are sufficient to reduce cholesterol and blood glucose levels (Ames and Rhymer, 2008). These effects can be acheived with barley (1,3;1,4)-β-glucan incorporated into foods to reduce the glycemic response; thus making (1,3;1,4)-β-glucan an attractive natural food additive (Cavallero et al., 2002). Consumption of dietary fiber, such as (1,3;1,4)-β-glucan, is also associated with a reduction in the risk of developing coronary heart disease (Anderson, 1995; Pereira et al., 2004). As mentioned earlier, the US Food and Drug Administration (FDA) have allowed whole grain barley and barley-containing products to carry a claim that they reduce the risk of coronary heart disease (FDA News Release, 2005).

The structural features of (1,3;1,4)-β-glucan are important determinants of their physical properties and functionality. An increased viscosity of cereal (1,3;1,4)-β-glucan is associated with reduced plasma cholesterol and a better control of postprandial serum glucose levels in humans (Bhatty, 1999; Klopfenstein and Hoseney, 1987; Lazaridou and Biliaderis, 2007; Wood, 2007). In addition to physiological effects associated with solution viscosity enhancement, (1,3;1,4)-β-glucan have been shown to gel at different rates under selected conditions (Cui and Wood, 2000; Lazaridou et al., 2003). For this reason (1,3;1,4)-β-glucan can be utilized as thickening agent to modify the texture and appearance of formulated foods (Wood, 1986) or to imitate properties of fat to develop calorie reduced food products (Inglett, 1990). Gelling properties of purified (1,3;1,4)-β-glucan extracted from lichenase, wheat, barley and oats are positively correlated with both high molecular weight and high DP3 content (Lazaridou et al., 2004). The physical features influencing commercially important gelling properties include ratios of β-(1,3), β-(1,4) linkages, presence and the amount of long cellulose-like fragments and ratio of cellotriosyl / cellotetraosyl units (Izydorczyk and Biliaderis, 2007). Barley has the highest and most varied DP3:DP4 ratios among cereal crops with values ranging from 2.8:1 to 3.3:1. Rye DP3:DP4 ratios are close to barley (3.0:1 to 3.2:1) and oats have the lowest and least divergent DP3:DP4 ratios of 2.1:1 to 2.3:1 (Wood et al., 1994). In extreme cases, ratios ranging from 2.1:1 to 4.8:1 have been reported for some barley genotypes (Burton et al., 2011; Collins et al., 2010).

2.8.3 Starch and (1,3;1,4)-β-glucan

Starch characteristics have a large influence on (1,3;1,4)-β-glucan content in barley. High concentrations of (1,3;1,4)-β-glucan are associated with low-amylose barley genotypes (*waxy*) (Ullrich et al., 1986) showing low or no expression of Granule Bound Starch Synthase I (GBSSI) gene [E.C.2.4.1.11]. GBSSI is encoded from the *waxy* locus located on the short arm of chromosome 7H and catalyzes the synthesis of amylose (Kramer and Blander, 1961; Tabata, 1961; Kleinhofs, 1997; Rohde et al., 1988). Genotypes with null alleles at the *waxy* locus produce starch with an amylopectin content of 90 - 98 % and an amylose content of only 2 - 10 % (Nagashima and Ishikawa, 1995; Washington et al., 2000). Waxy barley starch grains have been shown to be consistently smaller and contained slightly less starch than non-waxy grain (Tester and Morrison, 1992). High grain (1,3;1,4)-β-glucan concentrations are accompanied by high content of free sugars and phytoglycogens in many mutant genotypes associated with the *waxy* locus in barley (Fujita et al., 1999; Newman and Newman, 1992). Like *waxy* barley, the high amylose barley genotypes also show (1,3;1,4)-β-glucan concentrations above six percent (Izydorczyk et al., 2000)

Mutants with the *lys5f* and *lys5g* allele and a pseudo waxy line had reduced starch phenotype (Greber et al., 2000). Genotypes with *lys5f* allele had grain (1,3;1,4)-β-glucan concentration as high as 19.8 % compared to 13.3 % for *lys5g* and 6.5 % for the wild type. However, when the total carbohydrate content was compared as a sum of starch and (1,3;1,4)-β-glucan, the wild type and mutant *(lys5f* and *lys5g*) lines had comparable content (Munck et al., 2004). Similar carbohydrate compensation effect was also observed in a barley true waxy genotype 841878 similar to the high lysine barley lines with reduced starch content (Munck et al., 2004). These results led Munck et al. (2004) to postulate that genes regulating (1,3;1,4)-β-glucan synthesis are closely coupled to and compete with those genes that regulate starch synthesis in barley developing endosperm.

A model has been proposed by Islamovic et al. (2013) in which (1,3;1,4)-β-glucan and amylose levels are regulated by multiple genes interconnected by glucose availability. This model is based on genetic marker data and proposes that *Ugp2*, *CesA2* and *CslF6* interact to control (1,3;1,4)-β-glucan synthesis. The starch biosynthetic machinery, in particular the amylose / amylopectin production, has a large influence on the production of (1,3;1,4)-β-glucan through competition for glucose precursors. Inhibition in GBSSI or Glu-6-P-isomerase activity would

block amylose synthesis, while of inhibition of *CslF6* would result in low (1,3;1,4)-β-glucan. Blockage of either pathway would result in substantial increases in the alternative polysaccharide (Figure 2.3).

2.9 (1,3;1,4)-β-glucan synthase

(1,3;1,4)-β-glucan synthase is one of the few biosynthetic enzymes that can produce *in vitro* glucan polymers identical to those produced *in vivo* (Buckeridge et al., 2004). These experiments require Golgi membranes suggesting (1,3;1,4)-β-glucan synthase is a membrane-associated enzyme. Barley endosperm (1,3;1,4)-β-glucan synthase is most active at pH 9.0 and 25 °C and requires more than 2 mM Mg^{2+} for maximum activity. Enzyme activity of barley endosperm (1,3;1,4)-β-glucan synthase increases with higher UDP-Glc concentrations but decreases at concentrations above 3 mM UDP-Glc. The specific activity of (1,3;1,4)-β-glucan synthase in barley endosperm is 200 – 400 pmol / min / mg protein (Tsuchiya et al., 2005). This specific activity (1,3;1,4)-β-glucan synthase in the endosperm is much lower than the 4 - 5 nmol / min / mg protein observed for barley seedlings at 1 mM UDP-Glc (Becker et al., 1995), suggesting that different genetic factors contribute to (1,3;1,4)-β-glucan synthase in these respective tissue types. To complicate matters, (1,3;1,4)-β-glucan synthase can be converted to callose [(1,3)-β-glucan] synthase (EC 2.4.1.34) *in vitro* on disruption of an intact plasma membrane or Golgi membranes, but this artificial conversion is not thought to be reflective of processes *in vivo* (Buckeridge et al., 2004).

The activity of (1,3;1,4)-β-glucan synthase *in vitro* is not always correlated with (1,3;1,4)-β-glucan concentrations in the endosperm. Tsuchiya et al. (2005) studied (1,3;1,4)-β-glucan synthase activity and (1,3;1,4)-β-glucan concentrations in caryopses harvested at 11–22 days after flowering to determine the relationship between enzyme activity and the amount of (1,3;1,4)-β-glucan deposition. It was found that enzyme activity varied between genotypes but increased enzyme activity did not always correspond to increased (1,3;1,4)-β-glucan concentration in the mature seed.
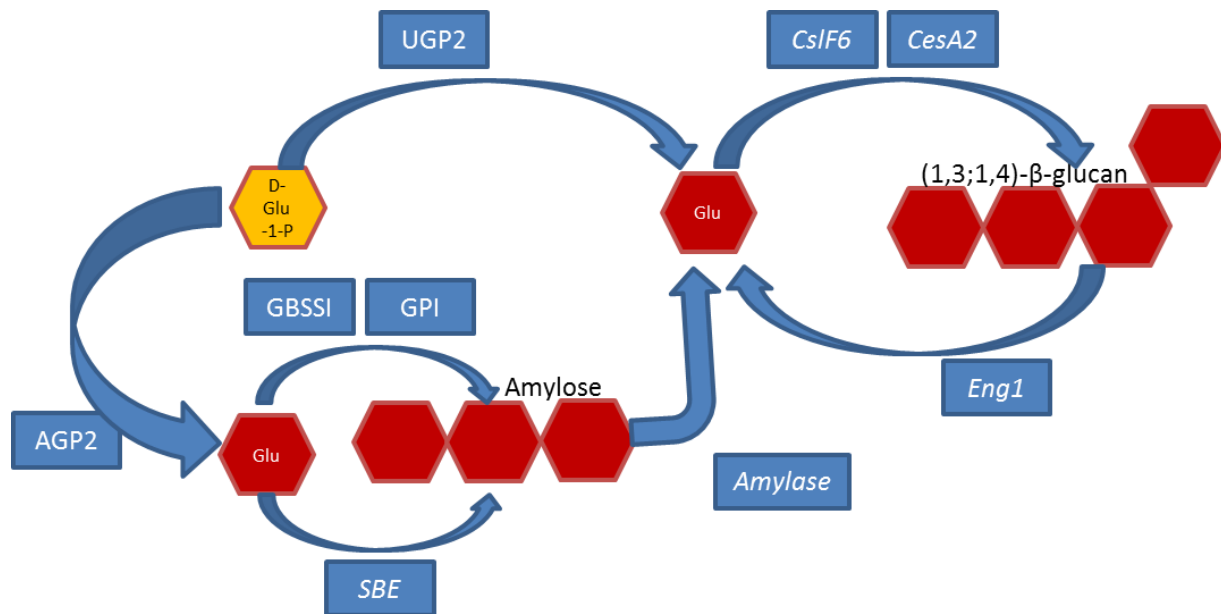
Figure 2.3. Islamovic model of (1,3;1,4)-β-glucan synthesis. Model for the role of carbon partitioning of glucose precursors between starch and (1,3;1,4)-β-glucan synthesis (adapted from Islamovic et al. 2013). (1,3;1,4)-β-glucan content in the grain as determined by the competitive action of amylose and (1,3;1,4)-β-glucan synthesis and degradation. *CslF6* and *CesA2* work in an undefined manner to create (1,3;1,4)-β-glucan. AGP2: glucose-6-phosphate adenyltransferase 2; Eng1: endo-1,3-β-glucosidase 1; GBSSI: granule bound starch synthase I; GPI: glucose-6-phosphate isomerase; SBE: Starch branching enzyme; UGP2: UTP-1-phosphate uridyltransferase.

2.9.1 Control of (1,3;1,4)-β-glucan synthesis

Environment has an effect on (1,3;1,4)-β-glucan concentration and fine structure. High soil nitrogen and grain protein are positively correlated to (1,3;1,4)-β-glucan content (Güler, 2003; Hang et al., 2007). Moisture and (1,3;1,4)-β-glucan grain concentrations have been shown to be negatively correlated (Morgan and Riggs 1981; Molina-Cano et al., 1995) with drought stressed plants have higher concentrations than well watered plants (Perez-Vendrell et al., 1996). No evidence exists in barley to suggest environment effects fine structure, but in oats, wet environments were associated with higher frequency of DP3 fragments (Doehlert and Simsek, 2012).

An analysis of the (1,3;1,4)-β-glucan fine structure of Beach, CDC Dancer and HiFi oat varieties grown in six environments showed that genotype and environment are significant in determining the fine structure of the polymer (Doehlert and Simsek, 2012). The high (1,3;1,4)-β-glucan line HiFi had lower DP3 frequency than the other two genotypes leading the authors to speculate that increased action of (1,3;1,4)-β-glucan synthase may cause an increased competition for substrate causing a reduced frequency of DP3 subunits. Furthermore, wet environments were associated with higher frequency of DP3 subunits. It was suggested that superior growing conditions might provide more UDP-Glc (uridine diphosphoglucose) substrate in the cellular environment resulting in more efficient production of DP3 subunits. In an extreme example of environmental control, (1,3;1,4)-β-glucan synthesis is shown to be affected by oxygen availability. In submerged rice seedlings both the (1,3;1,4)-β-glucan synthase activity and the expression of *CslF6* was reduced compared to dryland controls. The reduction in (1,3;1,4)-β-glucan synthase activity was partially attributed to oxygen depletion as partial recovery of *CslF6* expression and (1,3;1,4)-β-glucan synthase activity could be achieved by providing bubble aeration (Kimpara et al., 2008).

Carbon partitioning may have an effect on (1,3;1,4)-β-glucan concentrations and fine structure. Ullrich et al. (1986) showed that waxy varieties were not only higher in (1,3;1,4)-β-glucan content but the viscosity of alkali and acid extracted (1,3;1,4)-β-glucan was also higher compared to non-waxy barley. Wood et al. (2003) compared covered and naked, long and short awn, and waxy and non-waxy barley genotypes. Analysis of the products released from (1,3;1,4)-β-glucan by lichenase digestion showed no association with awn length or difference in hulled vs. hulless phenotype. In this same study, comparison between waxy and non-waxy barley

cultivars show that the waxy phenotype has a significant effect on the (1,3;1,4)-β-glucan fine structure by increasing the relative amount of DP3 subunits (Wood et al., 2003). This is consistent with the findings of Mikkelson et al. (2013) who investigated DP3:DP4 in *lys5f* mutants of barley. Differences between the *lys5f* and its wild type Bomi in DP3:DP4 ratios were found to be significant. The low starch, high (1,3;1,4)-β-glucan *lys5f* mutant is enriched in DP3 subunits, giving the barley a ratio profile ranging from 4.6:1 to 4.8:1, whereas Bomi ranged from 3.3:1 to 3.6:1 (Mikkelson et al., 2013).

An early study looking at expression of *CslF6* suggested that (1,3;1,4)-β-glucan synthesis might be controlled by *CslF6* gene expression. High (1,3;1,4)-β-glucan variety Himalaya show increased transcription of *CslF6* compared to an elite malting variety Sloop, with low (1,3;1,4)-β-glucan concentration (Burton et al., 2008). More recently studies suggest *CslF6* expression may play a smaller role than originally suggested. As mentioned earlier, *lys5f* and *lys5g* have an increased grain (1,3;1,4)-β-glucan content compared to the wild type parent. *Lys5g* and *lys5f* accumulate high concentrations of (1,3;1,4)-β-glucan during development. Despite this, expression levels of *CslF6* and *CslH* were higher in wild type compared to the mutant varieties throughout development. These results suggest a sensing and signalling system acting at the cell wall to control expression of *CslF6* (Christensen and Scheller, 2012). In the most extreme example in this study was that of the lys3a mutant. The high lysine mutant *lys3a* is associated with hypermethylation and down regulation of several genes during endosperm development, including *CslF6*. Throughout endosperm development, *CslF6* expression was 1000 - fold higher in the wild type compared to the *lys3a* mutant. Despite this near silencing of the *CslF6* gene (1,3;1,4)-β-glucan concentration was comparable to Bomi until 32 - 50 days after flowering. At maturation the *lys3a* mutant contained four percent (1,3;1,4)-β-glucan concentration whereas the parent Bomi contained seven percent (1,3;1,4)-β-glucan concentration (Christensen and Scheller, 2012).

2.9.2 Models of (1,3;1,4)-β-glucan synthesis

Non-cellulosic polysaccharides of the plant cell walls are believed to be synthesized in the Golgi apparatus. Nascent hemicellulose is transported via Golgi-derived vesicles to the plasma membrane where it is deposited into the periplasmic space and eventually incorporated into the cell wall. *In vitro* studies suggest that the (1,3;1,4)-β-glucan synthetic machinery is

located within Golgi membranes (Gibeaut and Carpita, 1994). Urbanowicz et al. (2004) postulated that (1,3;1,4)-β-glucan synthase in maize (*Zea mays*) seedlings is localized in the Golgi apparatus and consists of two or three separate glycosyltransferases. *In vitro* synthesis of (1,3;1,4)-β-glucan from maize coleoptiles requires Golgi vesicle fractions, UDP-Glc, and either Mn^{2+} or Mg^{2+} as cofactors (Meikle et al., 1991; Gibeaut and Carpita 1993; Becker et al., 1995). The combined glucan synthase complex possesses an active site on the cytosolic face of the Golgi membranes and extrudes the growing glucan chains into the lumen of the Golgi using a supply of uridine diphosphoglucose (UDP-Glc) mediated by a Golgi-localized sucrose synthase. Light proteolysis using proteinase K and 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) treatment at 0.2 % (w / v) of cell extract reduces the (1,3;1,4)-β-glucan synthase activity, (1,3;1,4)-β-glucan molecular mass and also decreases the amount of DP3 produced indicating the catalytic site is exposed to the cytosol and embedded in a lipid membrane (Urbanowicz et al., 2004). Analysis of products has led to a model suggesting multiple sites on the (1,3;1,4)-β-glucan synthase complex synthesizing three β-(1,4) linkages at a time, followed by a β-(1,3) linkage under optimal conditions (Buckeridge et al., 1999). If UDP-Glc concentrations are below a certain threshold and unable to fill the three spaces available, the complex will lengthen to the β-(1,4) series and skip the β-(1,3) linkage creating an irregular pattern of β-(1,3) and β-(1,4) linkages (Figure 2.4). The proposed mechanisms for (1,3;1,4)-β-glucan synthesis inherently suggests environment would play a large role in the fine structure of (1,3;1,4)-β-glucan (Buckeridge et al., 1999). Recently this model has been called into question due to the lack of evidence of accumulation of (1,3;1,4)-β-glucan in Golgi vesicles of developing plants (Wilson et al., 2012).

To account for the lack immunochemical detection of (1,3;1,4)-β-glucan, a two phase assembly system has been proposed (Burton et al., 2010). The model proposes that individual cellodextrin units are synthesized separately by the cellulose synthase like family of proteins, either a *CslF* or *CslH* (Figure 2.5). The cellodextrins produced are anchored to recyclable lipids or proteins within the Golgi vesicle and transported to the plasma membrane. At the plasma membrane an as yet unidentified enzyme, such as a callose synthase, that randomly links the (1,4)-β-oligosaccharides with (1,3)-β-linkages creating the full length (1,3;1,4)-β-glucan
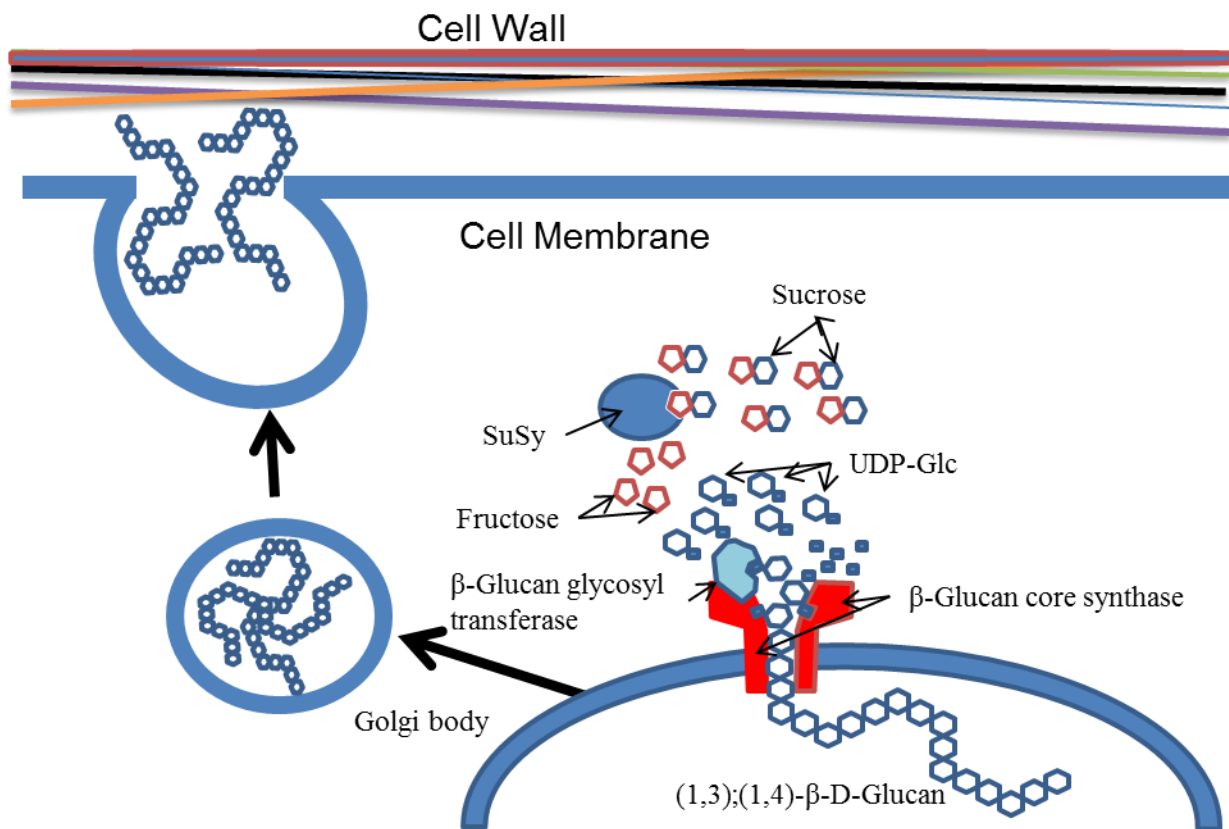
Figure 2.4. Urbanowicz model of (1,3;1,4)-β-glucan synthesis. (1,3;1,4)-β-glucan is created by the action of a complex containing two to three separate enzymes situated in the Golgi membrane. The growing (1,3;1,4)-β-glucan chain is extruded into a Golgi vesicle and is then transferred to the cell membrane where the full length molecule is then released to incorporate into the cell wall (Buckeridge et al., 1999; Urbanowicz et al., 2004).
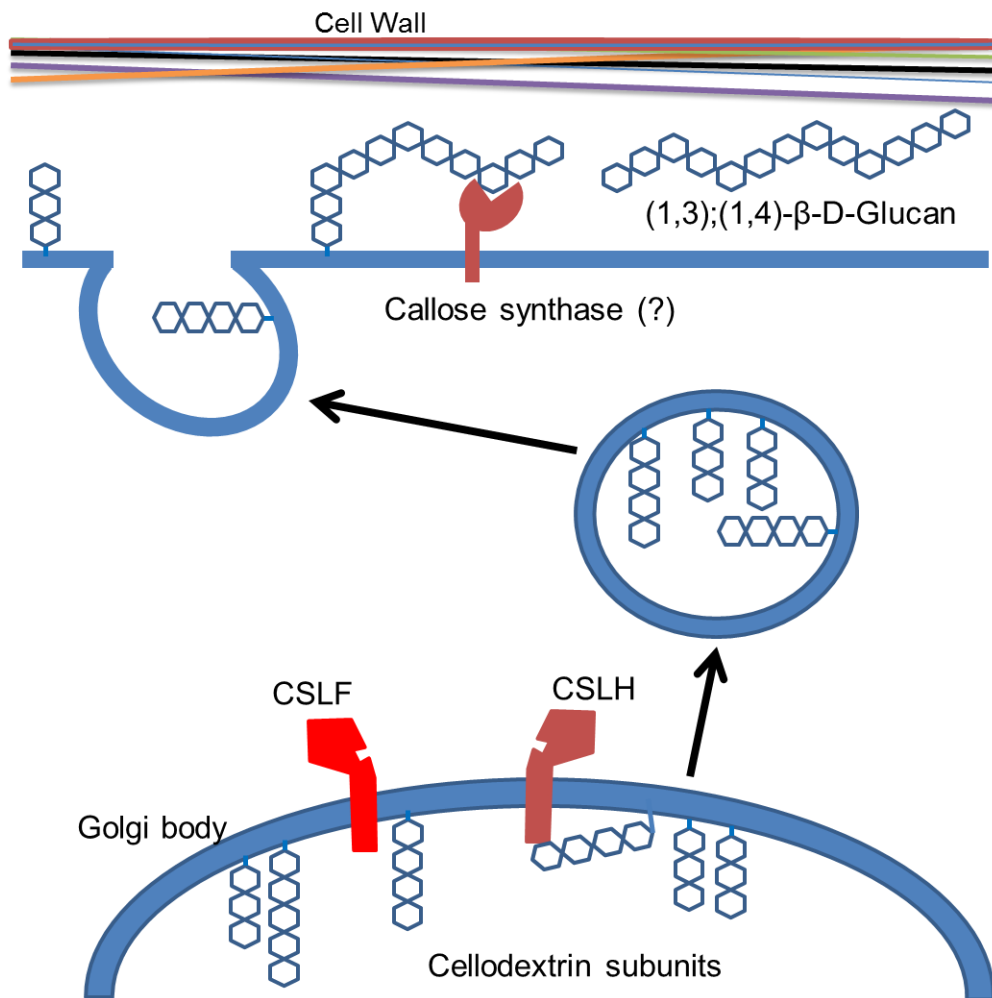
Figure 2.5. Two phase model of (1,3;1,4)-β-glucan synthesis (Burton et al., 2010). Cellodextrin molecules are synthesised separately by cellulose synthase like molecules in the Golgi membrane. Cellodextrin subunits are attached to phospholipids on the Golgi vesicle and transported to the cell membrane where yet unidentified enzyme(s) join the subunits to form (1,3;1,4)-β-glucan before being released to become incorporated into the cell wall.

molecule. Upon completion the polysaccharide is released into the periplasmic space and incorporates into the cell wall.

2.10 Candidate genes for (1,3;1,4)-β-glucan synthesis

2.10.1 Monocot-unique 23 kDa protein (P23k)

The *P23k* gene transcripts increased with increasing photosynthetic activity, thus suggesting that it plays a role in sugar metabolism (Oikawa et al., 2007). The P23k protein is abundant in vascular bundles and sclerenchyma where secondary wall formation is active in barley leaves (Oikawa et al., 2007) and may be involved in cell wall synthesis. Virus induced gene silencing of *P23k* gene leads to abnormal leaf development, decrease in calcofluor staining of cell wall polysaccharides and down-regulation *CslF6* gene involved in (1,3;1,4)-β-glucan synthesis (Oikawa et al., 2009). *P23k* is not expressed in developing barley grains but a closely related protein Jasmonate induced protein-23 (*JIP-23*) is localized to tissues involved in sugar transport in the endosperm (Oikawa et al., 2007).

2.10.2 CesA / Csl superfamily

The type I polysaccharide synthases, a group of enzymes thought to be responsible for plant cell wall biosynthesis, are encoded by members of a large multigene family, the cellulose synthase and cellulose synthase-like superfamily (CesA / Csl). This family is divided into several subgroups including the Cellulose synthase (CesA) subfamily and Cellulose synthase-like (Csl) subfamilies A to J (Richmond and Somerville, 2000; Farrokhi et al., 2006). The protein structure is well conserved within the CesA / Csl superfamily. The CesA family of proteins differs from the Csl family due to the inclusion of an N-terminal cysteine-rich region that form two zinc-binding RING-finger domains (Kurek et al., 2002). The latter are predicted to add stability to the CesA heterodimer subunits within the cellulose synthase rosette. All members of the superfamily contain eight transmembrane domains (TMD), with clusters of two N-terminal and six C-terminal predicted TMD dividing the protein into three major cytoplasmic domains. The N-terminal domain of both the CesA and Csl families include a highly variable region (HVR) which is separated from the large central domain by TMD 1 and 2, while TMD 3 to 8 in turn separate the central domain from the short C-terminal sequence. The central domain contains the catalytically active D,D,D,QxxRW signature characteristic of most glycosyltransferases

(Coutinho et al., 2003; Vergara and Carpita, 2001). The central domain is highly conserved with the exception of about 70 residues, named the class-specific region (CSR), which differs between paralogs but is well conserved between orthologs (Vergara and Carpita, 2001).

2.10.3 Cellulose synthase like (*Csl*) family

The Csl family can be subdivided into nine sub-families CslA to CslJ (CslI is omitted to avoid confusion). This family of genes is thought to encode proteins involved in the synthesis of various non-cellulosic β-linked cell wall polymer backbones. *CslF*, *CslH* and *CslJ* have been shown to be excusive to the Poaceae family (Farrokhi et al., 2006).

2.10.3.1 CslF subfamily

The *CslF* group of genes are exclusive to the commelinoid monocotyledon group of land plants (Burton et al., 2006). The first *CslF* genes in barley were identified through positional cloning of the 2H QTL found in the Steptoe / Morex DH population (Burton et al., 2006). Genetic mapping using the Steptoe / Morex doubled haploid population, revealed that a locus on chromosome 2H in the interval of ABG019 - ABC162 near the centromere had the largest effect on barley grain (1,3;1,4)-β-glucan concentration (Han et al., 1995). A syntenous region in rice identified a cluster of rice *CslF* genes at the barley chromosome 2H locus (Burton et al., 2006). Heterologous expression of rice *CslF* genes in Arabidopsis synthesized (1,3;1,4)-β-glucan in these plants which normally do not produce this polysaccharide. Four *CslF* genes from barley were identified in this region. Additional, *CslF* genes were identified on the barley chromosomes 1H, 5H, and 7H (Burton et al., 2008). *In silico* analysis of the Morex genome identified an additional three *CslF* genes (Schreiber et al., 2014); *CslF11* on the long arm of 7H, *CslF12* in the cluster of *CslF* genes on 2H and *CslF13* on the long arm of 2H. *CslF13* is predicted to be a pseudogene due to a stop codon before the final three transmembrane domains on the C terminal. Gene expression studies have shown that of all the *CslF* genes, *CslF6* and *CslF9* are the most abundantly expressed in the barley developing endosperm (Burton et al., 2008).

Transgenic experiments have indicated that the degree of solubility of (1,3;1,4)-β-glucan may be influenced by genetic factors in cereal grains. Over expression of *CslF4* in barley (cv Golden Promise) shifted the DP3:DP4 ratio higher (from 2.8:1 to 3.1:1), suggesting CSLF4 is preferentially producing cellotriose subunits. Conversely, overexpression of *CslF6* in the same

genetic background shifted the DP3:DP4 ratio to 2.1:1, indicating CSLF6 is preferentially produces cellotetraose subunits (Burton et al., 2011). However, RNAi mediated inhibition of *TaCslF6* in wheat (*Triticum aestivum*) did not affect the DP3:DP4 ratio but reduced both the molecular weight and total concentration of (1,3;1,4)-β-glucan (Nemeth et al., 2010).

2.10.3.2 CslF6

Transgenic studies which either overexpress or knock down individual *HvCslF* genes have strengthened the argument that individual genes from this family influence the concentration and fine structure of (1,3;1,4)-β-glucan (Burton et al., 2011; Nemeth et al., 2010). Overexpression of *CslF4* and *CslF6* increases (1,3;1,4)-β-glucan deposition but also reduces the ratio of DP3:DP4 from 2.8:1 to 2.1:1 in trangenic plants. Screening of a barley Targeting Induced Local Lesions in Genomes (TILLING) population identified a (1,3;1,4)-β-glucan-less phenotype (Tonooka et al., 2009). Molecular characterization of the (1,3;1,4)-β-glucan-less phenotype revealed three C253Y, G638D and G660D changes in the CSLF6 protein (Taketa et al., 2012). These mutations are close to the conserved aspartic acid residues, and the authors speculate that these amino acid shifts disrupt the nucleotide sugar binding domain. Screening of an ethyl methanesulfonate (EMS) mutagenized Harrington barley population identified a mutant line m351 with reduced (1,3;1,4)-β-glucan accumulation in the grains. Upon backcrossing the mutant line to the Harrington cultivar, genotypes with only 1.4 % grain (1,3,1,4)-β-glucan as compared to 5.2 % grain (1,3;1,4)-β-glucan in the parent Harrington were obtained (Hu et al., 2014). Genetic mapping of the phenotype and sequencing of the *CslF6* gene revealed a single point mutation causing a substitution of alanine to threonine at position 849. This mutation is positioned within the sixth transmembrane domain and was suggested to reduce the protein stability (Hu et al., 2014).

2.10.3.3 CslH

Cellulose synthase like H was identified in barley by expressed sequence tag (EST) database mining and subsequently confirmed through sequencing of bacterial artificial chromosome (BAC) clones from a Morex barley BAC library (Doblin et al., 2009). *CslH* maps to the short arm of 2H in close proximity to a cluster of *CslF* genes. The predicted amino acid sequence contains the conserved D, D, D, QFKRW motifs within the cytoplasmic domain and

six transmembrane domains. *CslH* expression can be detected at low levels in barley leaf and within the developing grain. Transgenic Arabidopsis plants producing a histidine-tagged CSLH show that HvCSLH is present in the Golgi but no presence at the plasma membrane could be demonstrated. Expression of *HvCslH* in Arabidopsis also produces low concentrations of (1,3;1,4)-β-glucan which accumulate in the cell wall.

2.11 Genetic markers

(1,3;1,4)-β-glucan concentration is a quantitative trait, suggested to be influenced by the additive effects of between three to five genetic factors (Powell et al., 1985), in addition to environmental factors (Morgan and Riggs, 1981; Perez-Vendrell et al., 1996). The study of quantitative traits in crops has rapidly evolved during the last two decades, due to the propensity of molecular markers, which have become useful tools to accelerate crop improvement. Ideally, a marker used for genetic mapping corresponds to a specific locus on the genome. Markers can be morphological, biochemical or molecular. Reliable morphological markers are usually visible (awns, row type, absence of hull, etc.) but restricted in number. Biochemical markers (isozymes and proteins) were among the first to be used in genetic diversity studies and to some extent for grain quality improvement (Buckler and Thornsberry, 2002). However, biochemical markers have several limitations such as being prone to post-translational modifications, are limited in number, labour intensive and therefore, rarely used at present for crop improvement. Instead, molecular markers (DNA-based markers) have become the pre-dominant marker system for genetic analysis and crop improvement. Molecular markers have the advantage over biochemical markers in that they are stable and not influenced by environment or life stage of the plant. A summary of advantages and disadvantages of a select group of marker systems is summarised in Table 2.4.

2.11.1 Simple sequence repeats (SSRs)

Microsatellites or simple sequence repeats (SSRs) are a class of repetitive sequences found in all eukaryotic organisms (Litt and Luty, 1989; Duran et al., 2009). SSRs are short, tandem repeats of monomers between 1 - 6 bp long that are thought to have arisen through DNA polymerase slippage during DNA replication (Levinson and Gutman, 1987). SSRs can occur as perfect repeats without interruption, imperfect repeats which include interruptions by non-repeat

Table 2.4. Advantages and disadvantages of different marker systems.

Marker system	Isozyme	RLFP	RAPD	AFLP	SSR	DArT	CAPS	STS	SSCP	SNP array	GBS
Quantity of DNA	NA	High	Low	Low	Low	Low	Low	Low	Low	Low	Low
Hybridization / PCR	NA	Hybrid	PCR	PCR	PCR	Hybrid / PCR	PCR	PCR	PCR	Hybrid	PCR
Level of Polymorphism	Low	Medium	Medium	High	High	High	Low	Low	Low	High	High
Reproducibility	High	High	Low	High	High	High	High	High	Medium	High	Medium
Dominant (D) / Codominant (C)	C	C	D	D	C	D	C	D	C	C	C
Sequence information required	NA	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No / Yes
Amenability to automation	No	No	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes

RLFP- Restriction Fragment Length Polymorphism, RAPD- Random Amplified Polymorphic DNA, AFLP- Amplified Fragment Length Polymorphism, SSR- Simple Sequence Repeat, DArT- Diversity Array Technology, CAPS- Cleaved Amplified Polymorphism Sequence, STS- Sequence Tagged Site, SSCP- Single Strand Conformation Polymorphism, SNP Array- Single Nucleotide Polymorphism array, GBS- Genotyping By Sequencing

nucleotides, compound repeats where two or more SSRs are found adjacent to one another or a combination of these three variants (Duran et al., 2009). SSRs are highly polymorphic, abundant and co-dominant. The markers are highly reproducible, amenable to automated screening, require little DNA for screening and can identify multiple alleles at a single or multiple loci. Primers designed for SSRs in one species show a limited degree of transferability between related species enabling comparative genomic analysis.

2.11.2 Diversity Array Technology (DArT®)

High-throughput and low-cost Diversity Array Technology (DArT®) is a microarray hybridization based technology that enables simultaneous genotyping of polymorphic loci spread over the genome without prior sequence information (http://www.diversityarrays.com; Jaccoud et al., 2001; Wenzl et al., 2006). By scoring the presence versus absence of specific DNA fragments in samples of genomic DNA digested with specific restriction enzymes, DArT® marker analysis can generate whole genome fingerprints rapidly and repeatedly. The markers generated are dominant in nature and therefore heterozygotes cannot be identified which is a major limitation of this technique.

2.11.3 SNP-based markers

Single nucleotide polymorphisms (SNPs) are the most abundant molecular markers that are widely distributed throughout genomes. SNPs represent a single nucleotide base difference between two individuals at a defined position in a DNA fragment. These DNA variants are represented by three different categories: transitions (a substitution of a purine for a different purine or a pyrimidine for another pyrimidine (C / T or G / A)), transversions (the substitution of a pyrimidine for a purine (T / A, C / A, T / G or G / C)) and small insertions / deletions (indels). SNPs can in principle be bi, tri or tetra-allelic at any site, but are usually bi-allelic as tri and tetra-allelic sites are rare (Doveri et al., 2008). SNPs are evolutionarily stable, relatively abundant and can be used as direct markers as the sequence information can provide exact information about the allele location and possibly function. SNPs are more prevalent in non-coding regions of a genome. Within coding regions SNPs can be either synonymous, not altering amino acid sequences in proteins or non-synonymous and therefore altering amino acid sequences (Sunyaev et al., 1999). Occasionally, synonymous SNPs lead to modification of mRNA splice sites causing

phenotypic changes (Richard and Beckman, 1995). The majority of SNP genotyping are based on either one or a combination of techniques; allele specific hybridization, invasive cleavage, oligonucleotide ligation, primer extension or direct sequencing (Syvänen, 2001).

2.11.3.1 Cleaved Amplified Polymorphic Sequences (CAPS) markers

Cleaved Amplified Polymorphic Sequences (CAPS) polymorphisms are locus-specific oligonucleotide primers produced PCR amplicons which contain SNPs or indels that contain a variable site effecting restriction endonuclease recognition site(s) leading to different fragment lengths after digestion. The advantage of CAPS markers are their repeatability, co-dominant nature, low cost and medium throughput. A major disadvantage of CAPS markers are that detailed prior sequence knowledge is needed to design the oligonucleotide primers and to identify if a variable restriction site is present within the amplicon.

2.11.3.2 Genotyping-by-sequencing (GBS)

Barley (*Hordeum vulgare* L.) is diploid with a large haploid genome of 5.1 gigabases (Gb). In 2012, the International Barley Genome Sequencing Consortium (Mayer et al., 2012) sequenced 4.98 Gb of the Morex genome and physically mapped 3.90 Gb anchored to a high density genetic map of barley. It was found that approximately 84 % of the genome consists of either, mobile elements or repeat structures, the majority being long terminal repeat retrotransposons. There is reduced repetitive DNA content within the terminal 10 % of the physical map of each barley chromosome arm. Gene containing BACs show a depletion of retrotransposons. Annotation efforts revealed 24,154 high confidence genes and lead to estimates that the entire barley transcriptome consists of 30,400 genes. On average it was estimated that there are five genes per Mb. This gene density increases to 13 genes per Mb at the proximal and distal ends of the chromosomes (Mayer et al., 2012). When compared to genetic maps, the generated physical map reveals reduced recombination frequency at the pericentromeric and centromeric regions of the barley chromosomes, a feature that compromises exploitation of genetic diversity and negatively impacts genetic studies and plant breeding.

Significant advancements in DNA sequencing technology during the last decade have drastically reduced the cost of DNA sequencing (Delseny et al., 2010). The recently developed Next Generation Sequencing (NGS) technology is high throughput, not limited to expressed

sequences or prior sequence information and generates data that can be quickly incorporated to physical maps (Mammadov et al., 2012). Therefore, SNP discovery using NGS technology offers an advantage of needing no prior sequence knowledge and sequencing efficiency for genotypes scales directly with genetic diversity (Elshire et al., 2011). Therefore, NGS is being used for Genotyping by Sequencing (GBS) to enrich the repertoire of SNP markers for crop improvement.

The major limitation in the utilization of GBS for marker discovery in large, complex genomes, such as wheat and barley, is to avoid highly-repetitive sections of the genome and ensure that each individual is sampled at similar (homologous) regions (Mammadov et al., 2012; Peterson et al., 2012). Early on, it was realised that sequence specificity of restriction endonucleases could be used to accomplish both of these goals. GBS targets the genomic sequence flanking restriction sites. Using methylation-sensitive restriction enzymes (RE) repetitive regions of genomes can be avoided and lower copy regions can be targeted with higher efficiency (Elshire et al., 2011). Originally, GBS approach used a single RE to capture the genomic regions between restriction sites but the method has been recently modified to incorporate a two RE approach termed double digestion Restriction-site Associated DNA (ddRAD) sequencing (Poland et al., 2012a). The ddRAD sequencing approach uses one "rare-cutter" and one "frequent-cutter" enzyme to further reduce genomic complexity allowing for libraries with a suitable and uniform complexity which greatly simplifies quantification of the library prior to sequencing (Poland et al., 2012b). The ddRAD sequencing method eliminates random shearing of the genome and allows for greater size selection. These two features reduce duplicate region sampling which reduces by almost 50 % the number of reads needed to produce high confidence sampling of a SNP associated with a given RE site. Secondly, region representation bias favouring fragments closest to the average size selection increases the likelihood of recovering similar genomic regions across all individuals, even those with read counts recovery below saturation (Poland et al., 2012b).

2.12 Genetic mapping of quantitative trait loci (QTL)

A quantitative trait locus (QTL) is a genomic region that contributes to a trait value. QTL mapping estimates the genomic regions, the number of regions, their effect on phenotypic variation and modes of gene action of individual determinants contributing to the inheritance of a

continuously variable trait (Paterson et al., 2002). Three basic criteria are needed for the genetic mapping of QTL: 1) reliable phenotype data for the population 2) a population showing segregation for the target trait and 3) genetic markers with adequate diversity to represent genotypic data for the population. Genetic mapping places molecular genetic markers in linkage groups based on their co-segregation in a population. Genotyping of various barley mapping populations using genetic markers has resulted in many genetic linkage maps with increasingly high precision over the years. Traditionally, genetic maps are prepared by analysing populations derived from crosses of genetically and phenotypically diverse parents. Estimating the recombination frequency between genetic loci can statistically estimate marker distances, which are reported in centi-Morgan (cM) units. Population size, genetic diversity and marker density influence map resolution. Genetic maps should not be confused with physical genetic maps which represent complete sections of sequenced nucleotides and are measured by nucleotide bases (Young, 2001). Recently, a functional physical map consisting of a cumulative length of 4.98 Gb, representing 96 % of the barley genome was compiled (Mayer et al., 2012).

2.12.1 Determination of (1,3;1,4)-β-glucan

(1,3;1,4)-β-glucan concentration can be determined by enzymatic means. Samples are suspended and hydrolysed in an appropriate buffer and subjected to lichenase digestion. Lichenase (EC 3.2.1.73) is a specific, *endo*-(1,3;1,4)-β-D-glucan 4-glucanhydrolase which cleaves β -(1,4) linkages on the reducing end of a 3-*O*-linked β-D-glucopyranosyl residue in (1,3;1,4)-β-glucan. Digestion results in the release of cellotriose to larger cellodextrin subunits. An aliquot of the filtered sample containing the cellodextrin subunits is then hydrolysed to completion and total glucose is determined in the sample by a colorimetric means (McCleary and Codd, 1991).

Several dyes specifically label (1,3;1,4)-β-glucan. A flourometric evaluation of (1,3;1,4)-β-glucan concentration using the Carlsberg calcofluor method, developed by Aastrup and Jørgersen (1988), is a fast and reliable method. The fluorochrome Calcofluor has been shown to create a dye complex formation with (1,4)-β-glucan in the cell wall. The calcofluor (1,3;1,4)-β-glucan complex has a maximum absorption band at 363 nm and an emission band at 420 nm. When used in aqueous solutions derived from flour this can be used to reliably determine (1,3;1,4)-β-glucan concentration. Calcoflour staining can also be used during microscopy to

visualise (1,3;1,4)-β-glucan within the cell wall. A simple, low-cost and semi-automated method for determination of (1,3;1,4)-β-glucan can also be achieved using Congo Red, a dye which reacts with high-molecular weight (1,3;1,4)-β-glucan. The Congo Red (1,3;1,4)-β-glucan complex can be measured by spectrometer at an absorbance at 545 nm. Congo Red may overestimate (1,3;1,4)-β-glucan concentrations as it has been shown to weakly interact with starch (Wood and Fulcher, 1983). Aniline Blue can also be used for (1,3;1,4)-β-glucan detection though results have shown the specificity for (1,3;1,4)-β-glucan is lower than that for both Calcofuor and Congo Red (Wood and Fulcher, 1983).

A monoclonal antibody generated against (1,3;1,4)-β-glucan-BSA conjugate is specific for (1,3;1,4)-β-glucans (Meikle et al., 1994). The antibody shows no cross-reactivity against (1,3)-β-glucan and a very weak cross-reactivity against cellopentaose-BSA and (1, 4)-β-oligoglucosides. The antibody affinity for ligands containing (1,3;1,4)-oligoglucosides is at least two orders of magnitude higher than those containing only (1,4)-β-linkages and has no cross-reactivity against cellulose. The optimum binding epitope consists of at least a hexa-saccharide with the structure Glu (1,4) Glu (1,4) Glu (1,3) Glu (1,4) Glu (1,4) Glu (R). Quantitation of (1,3;1,4)-β-glucan by a sandwich ELISA gives a near linear response in the range of 1 - 10 ng / ml. Due to the specificity of the antibody it is useful in quantifying (1,3;1,4)-β-glucan in solutions such as beer or wort but it is most useful in light and electron microscopy to measure quality and location of (1,3;1,4)-β-glucan in developing grains (Meikle et al., 1994).

Wood et al. (1994) were the first to successfully use high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) to analyze the oligosaccharides from lichenase-hydrolyzed barley (1,3;1,4)-β-glucan. However, the quantification of oligosaccharides by HPAEC-PAD is limited by the knowledge of weight response factors (Wood et al., 1994). In general, the sensitivity of PAD decreased rapidly from DP2 to DP6, while, for higher oligosaccharides (DP 7 − 17), the decrease in the sensitivity of PAD is minimal (Timmermans et al., 1994). The measurement error for DP3 / DP4 ratios determined by HPAEC-PAD is around ± 10 % (Wood et al., 1994). Despite these drawbacks, HPAEC-PAD is the most used technique for oligosaccharide quantification of (1,3;1,4)-β-glucan fine structure (Collins et al., 2010).

2.12.2 Heritability

Heritability is the proportion of observed variation in a progeny that is inherited or the ratio of variation caused by genetics to total variation (Acquaah, 2007). The phenotypic variance (V_P) of a trait can be expressed mathematically as follows:

$$V_P = V_G + V_E + V_{GE}$$

V_G, V_E and V_{GE} represent genetic variance, environmental variance and variations caused by genotype by environment interactions, respectively. Using this equation we can calculate broad sense heritability (h^2), which is an estimation of heritability on the basis of all genetic effects. Broad sense heritability is calculated using total genetic variance and is expressed as follows:

$$h^2 = V_G/V_P$$

Narrow sense heritability can be calculated if additive genetic effects are accounted for. To find additive genetic effects genetic variance can be further divided into additive genetic variance (V_A), dominance variance (V_D) and the non-allelic or epistasis variance (V_I).

$$V_G = V_A + V_D + V_I$$

Genetic variances caused by genes are additive components and produce linear additive effects. The resemblance of parents to offspring is largely attributed to additive genetic effects. The dominance effects on the expression of quantitative characters are generally small compared with additive effects and epistatic effects are smaller than additive effects. Epistatic effects are generally ignored in calculating heritability. Narrow sense heritability can be expressed as follows:

$$h^2 = V_A/V_P$$

For homozygous individuals broad sense heritability and narrow sense heritability are equal to each other. Heritability estimates below 0.30 are considered low, 0.30 - 0.60 are considered moderate and estimates above 0.60 are considered high (Ayele, 2011). The methods of estimating heritability are based on portioning observed variation of a quantitative character into genetically and environmentally controlled components. The common method for estimating heritability is variance component method using the analysis of variance and parent-offspring regression method. Form analysis of variance (ANOVA), heritability can be estimated as follows (Singh et al., 1993):

$$h^2 = \sigma^2_G / (\sigma^2_G + \sigma^2_{GxE} + \sigma^2_e)$$

$$\sigma^2_G = (M_G - M_{GxE})/ (bL),$$

$$\sigma^2{}_{GxE} = (M_{GxE} - M_e)/ b,$$
$$\sigma^2 e = M_e$$

Where: M_G is the mean sum of squares (MSS) for genotype; M_{GxE} is the MSS for genotype by environment interaction; M_e is the MSS error; b represent number of replications and L is the number of environments used in the study. Calculation of heritability is an important first step genotype selection in various plant breeding approaches and is an important consideration when considering genetic mapping.

2.12.3 Single marker analysis

The simplest form of marker association is single marker analysis based on *t*-test, analysis of variance (ANOVA) and simple linear regression. The advantages of single marker analysis are that linkage maps are not required and they can be performed using basic statistical software. Either ANOVA or linear regression is most commonly used for single marker association. Linear regression can be used because the coefficient of determination (R^2) calculated from the marker can explain the amount of phenotypic variation in the QTL (Collard et al., 2005). ANOVA can be used to estimate the phenotypic variation, determine a favorable allele and calculate the additive and dominant effects of a marker. The major limitation of single marker analysis is the further a QTL is from a marker the more likely it is to be underestimated or go undetected entirely (Tanksley, 1993). Other weaknesses include a failure to provide an accurate QTL location or recombination frequency between the QTL and the marker (Doerge, 2002).

2.12.4 Simple interval mapping

Simple interval mapping (SIM) is an improvement on single marker analysis. SIM uses an estimated genetic map as a framework to locate and tests for QTL presence between each pair of adjacent markers (Lander and Botstein, 1989). The use of linked markers for analysis is statistically more powerful compared to single-point analysis because it compensates for recombination between markers and the QTL (Lander and Botstein, 1989; Liu, 1998). At each interval the SIM calculates a Logarithm of Odds (LOD) score by computing the likelihood of the observed distributions with and without fitting a QTL effect. The LOD scores are plotted along a linkage map and those that exceed a threshold significance level indicates the region associated

with a QTL. The most likely QTL position is indicated by the peak LOD score in that region. SIM uses a model which considers one QTL at a time and therefore can bias identification and estimation. When multiple QTLs are located on the same linkage group the additional QTLs can also contribute to sampling variance (Lander and Botstein, 1989; Zeng, 1994).

2.12.5 Composite interval mapping

Composite interval mapping (CIM), described independently by Zeng (1994) and Jansen and Stam (1994), was developed to overcome some of the limitations of SIM. The method combines interval mapping with multiple regressions as an interval test that attempts to separate and isolate individual QTL effects. By controlling genetic variation in other regions of the genome, CIM reduces background variation that can effect QTL detection. The analysis software incorporates cofactors into the model. These cofactors may be a set of markers that are significantly associated with the trait and may be located anywhere in the genome. They are typically identified by forward or backward stepwise regression, with user input to determine the number of cofactors and other characteristics of the analysis.

2.12.6 Association mapping

Association mapping (AM) is a natural population-based survey approach that identifies trait-marker relationships based on linkage disequilibrium (LD). Originally, AM was used extensively to dissect human diseases, and in the last decade has emerged as a powerful tool to identify QTLs in plants (Flint-Garcia et al., 2003). LD is caused by non-random association of alleles at different loci. The statistical association among a set of loci decays more or less quickly depending on the amount of recombination events that have occurred during meiosis (Dawson, 2000). AM is a cost effective alternative to traditional QTL mapping, in that specific populations do not need to be generated. Mapping populations are usually limited in the number of recombination events that have occurred. AM can be performed on collections of diverse genotypes or within breeding materials, and does not require a prior knowledge of performance. Further savings can be achieved by using historical phenotypic datasets to detect marker-trait associations (Sneller et al., 2009). Association mapping often uses natural populations of very divergent origins to capture the maximum number of ancient recombination events.

The resolution of association mapping is determined by linkage disequilibrium (LD) or invesrley by the amount of linkage decay. The level of LD is dependent upon the amount and distribution of genetic diversity, the mating system, selection regemes and the amount of recombination events in the ancestry of the genotypes. High LD can be attributed to inbreeding, population structure, admixture, low recombination rate, intense selection pressure (natural or artificial) and small population size. LD is usually measured as the difference between the observed and the expected frequency of the haplotype (D or D') but is usually calculated as the correlation between a pair of loci (r or frequently r2) (Zhao, et al., 2007). The reason for this is |D'| is biased according to sample size (Weiss and Clark, 2002), therefore the squared value of the correlation between markers (r2) is favored for association mapping. Within barley reported levels of LD vary depending on the genetic material used, the size of the population and the scale on which it is studied (whole genome or chromosomal level).

There are two main strategies in AM. The first one is genome-wide association mapping, or genome scan, which surveys genetic variations in the whole genome to detect indications of association for various complex traits (Risch and Merikangas 1996). In classic QTL mapping, a few hundred markers are usually sufficient for QTL mapping experiments, whereas genome-wide association studies (GWAS) typically require in the order of tens of thousands of genetic markers to achieve adequate coverage (Nordborg and Weigel, 2008). The second strategy is a candidate-gene association approach, which relates polymorphisms in selected candidate genes that have putative roles in determining phenotypic variation for specific traits (Gore et al., 2009). The limitation of this strategy is that candidate-gene studies rely on having predicted the identity of the correct gene relative to the phenotype studied.

Detailed knowledge of phylogenetic relationships of the population structure is required to add power to AM and reduce the likelihood of false associations (Hubisz et al., 2000). When phylogenetic information is limited, this problem can be overcome by accounting for population structure by genetic analysis (Buckler and Thornsberry, 2002). Bayesian clustering approach can be used to infer the number of subpopulations (K) and to assign individuals to subpopulations based on membership proportion in each subpopulation (Q-matrix). Bayesian clustering operates by minimizing the Hardy-Weinberg and linkage disequilibrium that would result if individuals from different, randomly-mating populations were incorrectly grouped into a common population. Three of the most popular Bayesian based software programs for inferring

subpopulations are STRUCTURE (Pritchard et al., 2000; Hubisz et al., 2009), BAPS (Corrander et al., 2006) and PARTITION (Dawson and Belkhir, 2001).

Ordination is another approach used to reveal population structure. Ordination is commonly used to reduce complex multi-locus data sets into two or three dimensional scatter plots that represent genetic structure spatially, with putative subpopulations forming distinct clusters of points. The most common methods used in genetic studies involve Principal component analysis (PCA) or principal coordinate analysis (PCoA). PCA transforms a similarity matrix, a set of possibly correlated variables, into a smaller number of uncorrelated variables called principal components. The first principal component accounts for as much variability in the data as possible with each succeeding component accounting for as much of the remaining variability as possible. PCoA uses a distance matrix between a set of variables (ie. genetic distance) to assign each item a location in a low dimensional space. Much like PCA, the major axes of variation are then located within the multidimensional data set. Each successive axis explains proportionately less of the total variation, such that when there are distinct groups, the first two or three axes will typically reveal most of the separation among them. PCA is used for similarities and PCoA for dissimilarities. However, binary measures (such as genetic alleles) are distance measures and, therefore PCoA should be used (Zuur et al., 2007).

2.13 (1,3;1,4)-β-glucan QTLs in barley

Even where grain (1,3;1,4)-β-glucan is measured by similar means in multiple studies, large differences due to environmental conditions could mean that the most critical loci for this character vary between diverse sites. Due to functional variations or differences in genetic mechanisms between the parental genotypes, QTLs associated with quality parameters can vary considerably between barley populations (Thomas, 2003). In some cases, QTLs coincide with functional genes (Molina-Cano et al., 2007; Islamovic et al., 2013), but in most cases, QTLs are located on genomic regions containing no clear candidate genes (Mather et al., 1997; Li et al., 2008). Some of the malt (1,3;1,4)-β-glucan loci are co-incident with QTL for malt β-glucanase activity or other quality parameters and not reflective of (1,3;1,4)-β-glucan synthesis. Despite these limitations, all seven barley chromosomes are associated with the (1,3;1,4)-β-glucan concentration in either grain or malt (Figure 2.6).

QTL and association mapping studies have shown areas on 1H may be involved in regulation of (1,3;1,4)-β-glucan accumulation. A meta-analysis of the Morex / Steptoe, Morex / Harrington and Harrington / TR306 populations, showed a common major QTL for (1,3;1,4)-β-glucan near marker Ica1 (BIN 6) in all three populations (Igartua et al., 2002). Grain (1,3;1,4)-β-glucan concentration assessed in a doubled haploid (DH) population, derived from cross between two row winter feed variety Nure and two row malting variety Tremois, identified two QTLs for (1,3;1,4)-β-glucan on Chromosome 1H in the Bmac0399 - Bmag0211 and Cor1 - Bmag0382 intervals, respectively (Laido et al., 2009). This region has also been associated with (1,3;1,4)-β-glucan concentration in barley grain and wort of Steptoe / Morex (Han et al., 1995), Arapiles / Franklin, Alexis / Sloop populations (Panozzo et al., 2007), and two association mapping populations (Houston et al., 2014; Shu and Rasmussen, 2014). Three genes within this 1H region have been identified as likely candidates affecting (1,3;1,4)-β-glucan accumulation. One of the proposed causative factors for the QTL is a *CslF9* gene, which is highly expressed during early grain development (Burton et al., 2008). However, only speculative links to (1,3;1,4)-β-glucan production have been made for *CslF9*. An alternative candidate gene for 1H QTL may be Starch Synthase IIIa (*SSIIIa*), which is involved in starch biosynthesis. Barley containing the *amo1* mutant locus, which affects SSIIIa function, accumulate higher concentrations of (1,3;1,4)-β-glucan in the endosperm in addition to altered starch structure (Li et al., 2011). A third candidate gene for 1H QTL has been proposed based on a syntenic region in rice corresponding to rice gene *Os05g01020* encoding a histone deacetylase complex (HDAC) protein. The genomic region was identified from analysis of the Falcon / Azul mapping population associated with amylose content (Islamovic et al., 2013), but the authors speculate this gene also regulates (1,3;1,4)-β-glucan synthesis through transcriptional repression of targeted genes via histone deacetylation. Using a genetic mapping population derived from a cross between Beka and Logan, Molina-Cano et al. (2007) found a QTL on 1H near the EST marker Ctig8484 (synonym scssr04163) positioned at 183 cM. This marker is found in the 5' UTR region of the UDP-Glc-4 epimerase 1 (*HvUGE1*) gene (Moralejo et al., 2004). UGE1 catalyses the inter-conversion of UDP-Gal and UDP-Glc. UDP-Glc nucleotide sugars act as activated sugar donors for the biosynthesis of cell wall polysaccharides such as (1,3;1,4)-β-glucan.
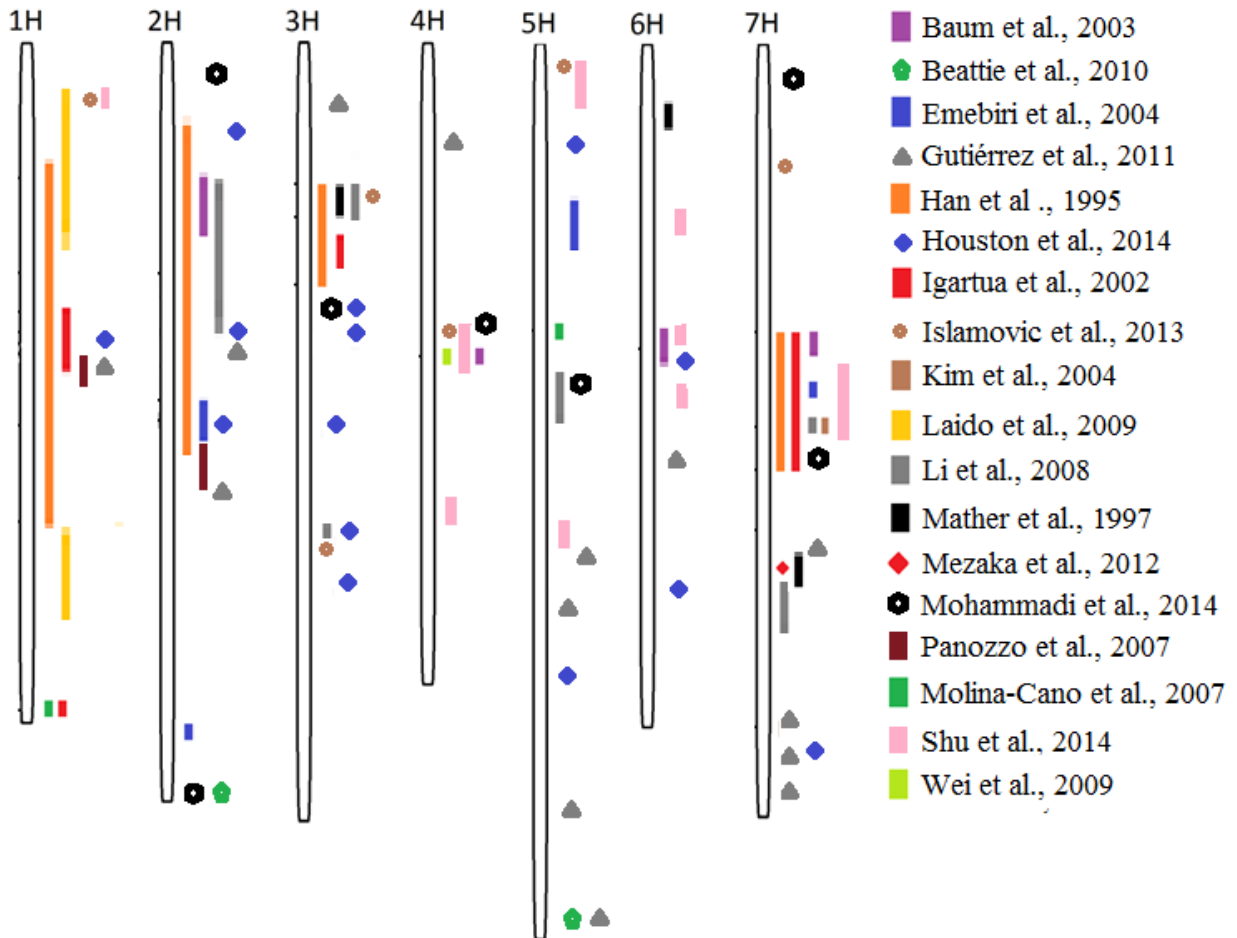
Figure 2.6. Summary of genomic regions associated with barley (1,3;1,4)-β-glucan. QTLs associated with both grain and / or wort concentrations.

Han et al. (1995) noted that the largest effect on barley grain (1,3;1,4)-β-glucan, in the cross Steptoe / Morex, is associated with a locus on chromosome 2H in the interval of ABG019 - ABC162 near the centromere. Burton et al. (2008) identified a cluster of four *CslF* genes in this region and speculated that the QTLs seen in the Morex / Steptoe population could be attributed to these genes. In an Arta / *H. spontaneum* mapping population, one QTL on 2H near HVBKASI explained 10 % of the phenotypic variation in total (1,3;1,4)-β-glucan (Baum et al., 2003). The HVBKASI marker is located in the gene B-ketoacyl-acyl carrier protein synthase I isoenzyme, part of the plant fatty acid synthesis pathway. These synthase proteins catalyze the condensation of acetate units to a growing acyl-ACP leading to the synthesis of palmitoyl-ACP (Kauppinen, 1992). Two QTLs were identified in the VB9524 / ND11231*12 population on chromosome 2H (Emebiri et al., 2003). The first was located near the centromere (EBmac0850) and the second near the telomere of the long arm (P14M55 - 156). A region near EBmac0684 is also associated with wort (1,3;1,4)-β-glucan in the Alexis / Sloop doubled haploid population accounting for 12 % of the variation (Panozzo et al., 2007). Regions on chromosome 2H have been identified in four association mapping studies associated with grain and wort (1,3;1,4)-β-glucan concentration (Beattie et al., 2010; Shu and Rasmussen, 2014; Mohammadi et al., 2014; Houston et al., 2014). Houston et al. (2014) and Shu and Rasmussen (2014) speculated that the cluster of *CslF* genes were responsible for the differences in (1,3;1,4)-β-glucan content. Mohammadi et al. (2014) identified a (1,3;1,4)-β-glucan QTL region closer to the telomere of 2HL in a panel of six-row barley. A glucan endo-1,3-β-glucosidase 14 gene was proposed to be the underlying gene at the QTL.

Several barley mapping populations have shown QTLs on 3H. Li et al. (2008) showed an inconsistent QTL (32 cM) in a TR251 / Bold mapping population, which explained 22.6 % of the phenotypic value. Emebiri et al. (2003) showed a QTL at Bmac0067, where the additive allele was contributed by ND11231_12. Harrington / TR306 doubled haploid population identified one region on chromosome 3H (near marker Ugp2) that affected extract (1,3;1,4)-β-glucan and extract viscosity (Mather et al., 1997). In the Falcon / Azul hulless, waxy population, two QTLs were found; one near Pilot OPA marker 11_20639 and the other near Pilot OPA marker 11_20650 (Islamovic et al., 2013). The marker 11_20639 is near a syntenous region in rice in close proximity to only UTP-1-phosphate uridyltransferase gene (*Ugp2*) in barley. The encoded enzyme is responsible for synthesis and pyrophosphorolysis of UDP-Glc, the key precursor of

(1,3;1,4)-β-glucan and cellulose. The same region has been identified in one association mapping study using the Oregon barley Coordinated Agricultural Project populations (Gutiérrez et al., 2011). A region spanning 55 to 99 cM was associated with grain (1,3;1,4)-β-glucan concentration (Houston et al., 2014; Mohammadi et al., 2014). Houston et al. (2014) speculated that a family of glycosyl transfrease and glycoside hydrolase enzymes may account for the variability of grain (1,3;1,4)-β-glucan concentration explained by this region.

Three bi-parental studies have reported QTLs on chromosome 4H are associated with grain (1,3;1,4)-β-glucan concentration. Wei et al. (2009) reported that bPb-2305 - Bmac0186 interval of chromosome 4H accounted for 11 % of the total variance in a CM72 / Gairdner doubled haploid population. In an Arta / *H. spontaneum* mapping population, a QTL near e32m49 - 06 contributed to 4.3 % of the variation in (1,3;1,4)-β-glucan (Baum et al., 2003). The Falcon / Azul population identified three QTLs associated with grain (1,3;1,4)-β-glucan concentration in the same genomic region. Glucose-6-phosphate 1-dehydrogenase gene (Os03g20300) and endo-1,3-β-glucosidase (*Eng1*) (Os03g18520) were identified as a putative candidate genes present in a syntenous region in the rice genome. Os03g20300 is a rate limiting enzyme in the pentose phosphate pathway and may control (1,3;1,4)-β-glucan synthesis through carbon partitioning while Eng1 is a (1,3;1,4)-β-glucan hydrolyzing enzyme which may control (1,3;1,4)-β-glucan synthesis and degradation (Islamovic et al., 2013). Association mapping studies have also shown regions near the telomere of 4HS (Gutiérrez et al., 2011) and the centromere (Shu and Rasmussen, 2014; Mohammadi et al., 2014) are associated with grain (1,3;1,4)-β-glucan concentration. Shu and Rasmussen (2014) speculated that glucan synthase-like 3 (EC 2.4.1.34, *HvGSL3*) may be a putative candidate gene in this chromosomal region.

Chromosome 5H has been identified as a contributor to variation in wort or grain (1,3;1,4)-β-glucan in several genetic mapping populations. Igartua et al. (2002) noted that the Harrington / TR306 mapping population had two "hot-spots" at each end of the chromosome, but these were contributing to a vast array of phenotypes including kernel plumpness, protein content and higher malt extract. In the Dicktoo / Morex population, a QTL for (1,3;1,4)-β-glucan content was found in the interval from apAdh-apt59a near the telomere of the short arm of chromosome 5H (Ozeil et al., 1996). Islamovic et al. (2013) reported a similar QTL in Bin 2 near the OPA marker 11_21365 and suggested that *CesA2* is a putative candidate gene contributing to (1,3;1,4)-β-glucan concentration variation. In the Beka / Logan population, a major QTL,

Bmag337 was located near the centromere (Molina-Cano et al., 2007). Emebiri et al. (2003) described a similar QTL near XP14M51-203 in this population in association with wort (1,3;1,4)-β-glucan content. Five association mapping studies have found marker associations on 5H for either grain or wort (1,3;1,4)-β-glucan concentration (Beattie et al., 2010; Gutiérrez et al., 2011; Mohammadi et al., 2014; Houston et al., 2014; Shu and Rasmussen, 2014). A region near the telomere of 5HS coincide in two association mapping studies (Shu and Rasmussen, 2014; Houston et al., 2014) though neither study identifies a putative candidate gene in this region. Houston et al. (2014), Mohammadi et al. (2014) Gutiérrez et al. (2011) and Shu and Rasmussen (2014) identify markers in the interval between cM 63.3 and cM 128.7. MLOC_44777 (*HvCel3*; a member of the endo-(1,4)-β-glucanase gene family) and MLOC_65914 (an orthologue to *AtCslE6*) were identified as likely candidate genes affecting grain (1,3;1,4)-β-glucan concentration (Houston et al., 2014). Gutiérrez et al. (2011) and Beattie et al. (2010) identify a similar region near cM190 on the 5H chromosome in associated with wort (1,3;1,4)-β-glucan concentration.

QTL information on chromosome 6H is the least reported and least consistent. TR251 / Bold population had a weak QTL on chromosome 6H that explained 22 % of the variation in (1,3;1,4)-β-glucan but this was inconsistent, only appearing in two of three years analysed (Li et al., 2008). Mather et al. (1997) and Baum et al. (2003) have both reported weak QTLs near the telomere of the short arm of this chromosome. In only one out of six environments studied, a strong QTL with an LOD score of 7.3 was found near the marker MK_4313 - 482 (Islamovic et al., 2013). Gutiérrez et al. (2014), Houston et al. (2014) and Shu and Rasmussen (2014) have independently identified similar regions (30cM to 75cM) on the 6H chromosome in their respective populations using association mapping. The QTLs on chromosome 7H are the most consistent and widely reported for their association with (1,3;1,4)-β-glucan in barley. Originally described in the Steptoe / Morex mapping population described by Han et al. (1995), a large QTL for malt (1,3,1,4)-β-glucan is found in the interval between ABC455 and ABC156D. In this population, Steptoe provides the allele responsible for increased (1,3;1,4)-β-glucan. This large interval had two peaks: one near Brz and the second near Amy2. While this QTL did show an increase in the amount of wort (1,3;1,4)-β-glucan, it did not contribute to the variation found in barley grain (1,3;1,4)-β-glucan. In other studies this region has been associated with increased grain (1,3;1,4)-β-glucan. Molina-Cano et al. (2007) noted that the most significant QTL for grain

(1,3;1,4)-β-glucan was found in 7H near the marker Ctig5200, close to the centromere. Emebiri et al. (2003) and Baum et al. (2003) both report significant QTLs near the centromere in their populations. A QTL for grain (1,3;1,4)-β-glucan and grain weight in the Derkado / B83-12/21/5 population was also reported near centromere region of chromosome 7H near the Amy2 locus (Igartua et al., 2002). The *sex6* locus, which is thought to affect Starch Synthase IIa, is also located in this region. Li et al. (2011) showed that plants containing the *sex6* mutant locus accumulate more (1,3;1,4)-β-glucan the grain. Within our own group we have identified a major QTL near the centromere on chromosome 7H (Li et al., 2008). *CslF6* has also been shown to be in this region (Burton et al., 2008).The *nud* locus has been associated with (1,3;1,4)-β-glucan concentration in barley grain in an association mapping study looking at hulled and hulless two row Latvian spring varieties (Mezaka et al., 2011). Recently, AM studies in a panel of 3069 elite breeding spring barley lines have also identified two genomic regions associated with grain (1,3;1,4)-β-glucan concentration (Mohammadi et al., 2014). The first was present near the telomere of 7HS and a second near 83.4 cM. Houston et al. (2014) identified one marker on the 7H chromosome, SCRI_RS_23061, which was speculated to be near the Sucrose Synthase II (*HvSuSyII*) gene that had been previously suggested as a putative candidate enzyme participating in (1,3,1,4)-β-glucan synthesis (Urbanowicz et al., 2004).

2.14 Hypothesis

The end use of barley grain is greatly influenced by its concentration of (1,3;1,4)-β-glucan, that is a quantitative trait influenced by both genotype and environment. While a number of QTLs have been identified for (1,3;1,4)-β-glucan concentration in grain and wort many of these QTL span large areas on the genome making them of limited use in marker assisted selection for barley grain improvement. Although, every chromosome in barley has been associated with (1,3;1,4)-β-glucan grain or wort concentration but no information exists on the heritability or chromosomal regions affecting (1,3;1,4)-β-glucan fine structure.

Genomic regions near the centromere on 7H play a role in (1,3;1,4)-β-glucan grain concentrations and fine structure.

CHAPTER 3

GENETIC MARKERS FOR *CslF6* GENE ASSOCOATED WITH (1,3;1,4)-β-GLUCAN CONCENTRATION IN BARLEY GRAIN

3.1 Study 1*

In this study *HvCslF6* was sequenced from two barley lines, five genetic markers were developed and validated as significantly associated with (1,3;1,4)-β-glucan concentration in barley grain.

---

*Cory AT, Båga M, Anyia A, Rossnagel BG, Chibbar RN (2012). Genetic markers for *Csl*F6 gene associated with (1,3;1,4)-β-glucan concentration in barley grain. Journal of Cereal Science. 56: 332-339

3.2 Abstract

The amount of (1,3;1,4)-β-D glucan [(1,3;1,4)-β-glucan] accumulated in barley (*Hordeum vulgare* L.) cell walls is an important consideration for grain end-use. One of the major genes responsible for (1,3;1,4)-β-glucan biosynthesis is *HvCslF6*, which was analyzed in this study to determine the allelic variation between low (1,3;1,4)-β-glucan (~ 3.3 %) cultivar CDC Bold and high (1,3;1,4)-β-glucan (~ 5.2 %) line TR251. The CDC Bold *HvCslF6* allele showed 16 single nucleotide polymorphisms (SNPs) and two indels when genomic region downstream of the ATG start codon was compared to TR251 allele. Both indels added 16 nucleotides to *HvCslF6* first intron of CDC Bold and a single SNP in the third exon altered alanine 590 codon in the CDC Bold sequence to a threonine codon in TR251 allele. Genetic markers were developed for five polymorphic sites and confirmed useful to select low and high (1,3;1,4)-β-glucan lines in a previously characterized CDC Bold / TR251 mapping population and a novel F_5 recombinant inbred line (RIL) population derived from a Merit / H93174006 (4.8 and 5.3 % (1,3;1,4)-β-glucan) cross. An analysis of parental lines of six populations segregating for (1,3;1,4)-β-glucan concentration validated association between the TR251 *HvCslF6* haplotype and high (1,3;1,4)-β-glucan concentration in populations showing a (1,3;1,4)-β-glucan quantitative trait locus (QTL) on chromosome 7H.

3.3 Introduction

Beta-glucan ((1,3;1,4)-β-glucan) is a mixed linkage polymer [(1,3)-(1,4)-β-D-glucan] produced by grasses, bryophytes, certain fungi and algae (Fincher, 2009). The molecules are abundant in cell walls of endosperm and aleurone in cereal grains (Gibeaut and Carpita, 1991). Barley (*Hordeum vulgare*) and oat (*Avena sativa*) grain have a relatively high (1,3;1,4)-β-glucan concentration when compared to other cereals (Nemeth et al., 2010). Normally, three to six percent (1,3;1,4)-β-glucan accumulates in barley kernels, but concentrations up to 19.8 % are present in certain genotypes (Munck et al., 2004). Barley lines with very low (2.0 %) or no (1,3;1,4)-β-glucan also exist (Munck et al., 2004; Tonooka et al., 2009). The β-1,4-linked molecules form cellotriose (DP3) and cellotetraose (DP4) units, which are randomly joined by β-1,3 bonds producing kinks in the molecule. The DP3:DP4 ratio affects polymer solubility and

varies from 1.5 to 4.5 depending on the genotype. The highest content of soluble (1,3;1,4)-β-glucan is associated with DP3:DP4 ratios in the 1.5 - 2.5 range (Burton et al., 2010).

The amount of (1,3;1,4)-β-glucan accumulated in grains is a major factor determining barley end-use. Low (1,3;1,4)-β-glucan barley is preferred by the feed, malting and brewing industries as high (1,3;1,4)-β-glucan concentrations reduce feed conversion and cause filtration problems during brewing. For human nutrition, (1,3;1,4)-β-glucan has become a desirable food ingredient as it can lower serum low density lipoprotein-cholesterol (Hecker et al., 1998) and postprandial glucose levels (Cavallero et al., 2002). Thus, depending on the amount of soluble (1,3;1,4)-β-glucan, a barley-rich diet may reduce the risk of developing coronary heart disease or type II diabetes (Poppitt et al., 2007). For future development of barley cultivars for feed, malting or food purposes, the selection of lines with specific (1,3;1,4)-β-glucan concentration and / or composition will benefit from efficient genetic markers for the trait.

Initial mapping studies of (1,3;1,4)-β-glucan content in barley grain showed that the trait is controlled by three to five genetic loci (Powell et al., 1985), but also affected by environmental conditions such as drought (Perez-Vendrell et al., 1996). Later genetic mapping studies implicated all seven barley chromosomes in (1,3;1,4)-β-glucan accumulation in grain or malt. Major quantitative trait loci (QTL) are often reported on chromosomes 1H, 2H, 5H and 7H, less frequently on chromosomes 3H and 4H and occasionally on chromosome 6H (Han et al., 1995; Mather et al., 1997; Oziel et al., 1996; Panozzo et al., 2007; Li et al., 2008; Wei et al., 2009). The 7H QTL, first described for malt (1,3;1,4)-β-glucan in the Steptoe / Morex mapping population (Han et al., 1995), is the most consistently reported (1,3;1,4)-β-glucan QTL in barley.

The biosynthesis of β-1,4 linked polymers in plants is catalyzed by enzymes belonging to the cellulose synthase A (CesA) / cellulose synthase like (Csl) super-family. Certain CSL enzymes are implicated in (1,3;1,4)-β-glucan biogenesis as demonstrated by transgenic expression of a rice *CSL* gene in Arabidopsis, which causes accumulation of (1,3;1,4)-β-glucan in leaf cells, which normally are (1,3;1,4)-β-glucan-free (Burton et al., 2006). A cluster of *CSL* genes underlies the (1,3;1,4)-β-glucan QTL on chromosome 2H in barley and a *Csl*F6 gene (*HvCslF6*) is positioned at the (1,3;1,4)-β-glucan QTL on chromosome 7H (Han et al., 1995; Li et al., 2008). The involvement of *HvCslF6*, *HvCslF4* and *HvCslH1* in (1,3;1,4)-β-glucan synthesis is supported by transgenic expression of the genes in barley (Burton et al., 2011) wheat (Nemeth et al., 2010) and Arabidopsis (Doblin et al., 2009). In addition, an EMS-induced

mutation in *HvCslF6* allele of barley line Nishinohoshia is associated with loss of (1,3;1,4)-β-glucan production (Tonooka et al., 2009). Although there is ample evidence for a role for *HvCslF6* in (1,3;1,4)-β-glucan biosynthesis, the extent of natural genetic variation for this gene has not been studied. To obtain an initial assessment of *HvCslF6* diversity, we determined the DNA sequence of *HvCslF6* carried by a low (1,3;1,4)-β-glucan line CDC Bold (3.2 % (1,3;1,4)-β-glucan) and high (1,3;1,4)-β-glucan line TR251 (5.0 % (1,3;1,4)-β-glucan) to identify nucleotide sequence differences that could possibly explain *HvCslF6* expression or functional differences. Eighteen polymorphic sites were identified within the transcribed region of *HvCslF6* and the two alleles were found to be significantly associated with (1,3;1,4)-β-glucan concentration in mapping populations carrying a (1,3;1,4)-β-glucan QTL on chromosome 7H.

3.4 Materials and methods

3.4.1 Establishment of a Recombinant Inbred Line (RIL) population

The two-row, spring and malting genotypes Merit and H93174006 were used to produce a RIL mapping population of barley. Merit was developed by Busch Agricultural Resources LLC (Fort Collins, Co, USA), whereas H93174006 (TR05671) is derived from a H92076F1 x TR238 cross produced at the Field Crop Development Center, Lacombe, Alberta. The Merit / H93174006 population was advanced by single seed descent through the F3 and F4 generations in a greenhouse at Lacombe during 2007 and 2008. The F5 seeds from each F4 plant were bulked and advanced to produce seeds for F4:6 generation field trials.

3.4.2 Field trial

One hundred and eighty-four F6 RILs and parental lines Merit and H93174006 were planted at Vegreville Alberta, Canada (53 ° 31' N, 112 ° 6' W, 639 m altitude, with the Malmo series of an Eluviated Black Chemozem) in 2009. The trial used a randomized complete block design with three replicates of each F4:6 RIL and the parent lines. Weeds were controlled by Round-up Weathermax application before seeding and by Achieve 40DG and Buctril M spraying later in the growing season.

3.4.3 Determination of (1,3;1,4)-β-glucan concentration in grain

Samples of 10 g grain were milled to flour using an Udy-Mil Cyclone sample grinder (UDY Corporation, Fort Collins, CO, USA) equipped with a 0.5 mm sieve. The total (1,3;1,4)-β-glucan concentration was determined for duplicate 100 mg flour samples using a (1,3;1,4)-β-glucan (mixed linkage) Kit (Megazyme, Wicklow, Ireland). Predetermined samples of oat (8.8 % (1,3;1,4)-β-glucan) and barley flour (4.4 % (1,3;1,4)-β-glucan) were used as internal controls.

3.4.4 Isolation of genomic DNA

Plants were grown in growth chambers maintained at 20 °C and a 16 - h light period with 320 mmol $m^{-2}s^{-1}$. Leaves were harvested at the 5 - 10 leaf stage, frozen in liquid N_2 and stored at -80 °C before DNA was extracted and quantified as described (Li et al., 2008).

3.4.5 Production of *HvCslF6* contig

Oligonucleotide primers for amplification of *HvCslF6* fragments were designed using Primer 3 software (Rozen and Skatetsky, 2000) and the *HvCslF6* cDNA sequence of Morex (GenBank accession EU267181) as template. Amplicons were approximately 800 bp long with a minimum overlap of 150 bp to facilitate assembly of *HvCslF6* contig. The PCR reactions consisted of 100 ng template DNA, 5 pmol of each primer (Table S1), 10 mM Tris–HCl pH 8.3, 3.5 mM $MgCl_2$, 25 mM KCl, 20 mM dNTP, 1 U of Red Taq Polymerase (Sigma–Aldrich, St. Louis, MO, USA) in a total volume of 50 μL. Amplifications were performed using an Eppendorf AG Cycler (Eppendorf, Hamburg, Germany) programmed for an initial denaturation at 94 °C for 5 min followed by 34 cycles of 30 s denaturation at 94 °C, 20 s annealing at optimized temperature and 1 - 4 min extension at 72 °C. In the final cycle, the 72 °C step was extended by 2 min. A minimum of three independent PCR reactions were performed per DNA fragment.

An amplicon covering the 3' end was obtained by genome walking using the Genome Walker® Universal Kit according to manufacturer's instructions (Clontech, Mountain View, CA, USA). Five barley genomic DNA (2.5–5.0 μg) samples were digested at 37 °C overnight with restriction enzymes *Dra*I, *Eco*RV, *Pvu*II, *Sca*I and *Stu*I, respectively. The digested DNA samples were purified by phenol / chloroform extractions and ethanol precipitated before being ligated to Genome Walker adapters supplied with the kit. Primers for PCR amplification of

adapter-ligated fragments were AP1 and AP2 supplied with the kit and *HvCslF6*-specific 5022-F and 5153-F corresponding to nucleotides within the third exon of *HvCslF6* (Table S1)*. 5022-F / AP1 were used in the primary PCR, and 5153-F / AP2 were used in the nested PCR. Only libraries constructed using *Dra*I and *Stu*I generated identical 2,000 bp products, of which 400 bp overlapped with the Morex *HvCslF6* cDNA sequence. Primer pairs 5024-F, 5965-R and 5538-F, 6440-R were designed, which confirmed 1,200 bp of this sequence and extended the 3' genomic sequence by 779 bp.

Generated PCR products were separated by 1 % (w / v) agarose gel electrophoresis, visualized by ethidium bromide staining, excised and purified using Qiaquick gel extraction kit (Qiagen, Hilden, Germany). The DNA sequence of each PCR product was determined in forward and reverse orientations by Sanger sequencing conducted by DNA Sequencing Facility, Robarts Research Institute, London, Ontario, Canada. The DNA sequence of a *HvCslF6* fragment was considered complete when four of six high quality reads were in consensus. Analysis of generated sequences and assembly of *HvCslF6* contig was done using the Geneious 5.4.5 bioinformatics software (Biomatters Ltd; Auckland, New Zealand).

3.4.6 Genotyping

HvCslF6-specific primers were designed based on the contig sequence determined for CDC Bold allele. The genotyping reactions consisted of 100 ng template DNA, 5 pmol of forward and reverse primers, 10 mM Tris–HCl pH 8.3, 3.5 mM $MgCl_2$, 25 mM KCl, 20 mM dNTP, 0.5 U of Taq Polymerase (Thermo Fisher Scientific, Waltham, MA, USA) in a total volume of 25 µL. Amplifications were done in a Eppendorf AG Cycler (Eppendorf, Hamburg, Germany) programmed for an initial denaturation at 94 °C for 5 min followed by 40 cycles of 45 s denaturation at 94 °C, 20 s annealing at optimized temperature and 20 s extension at 72 °C. PCR products were separated by 1 % w / v agarose gel electrophoresis and visualized by ethidium bromide staining. For Cleaved Amplified Polymorphic Sequences (CAPS) markers, the PCR products were digested with restriction enzyme *Bgl*I or *Mn*Il *(*New England Biolabs, Ipswich, MA, USA), respectively, followed by analysis of digestion products by 2 % (w / v) agarose gel electrophoresis.

Genome-wide genotyping of populations were done using simple sequence repeat (SSR) markers available in the public domain (Graingenes; www.wheat.pw.usda.gov) and diversity

arrays technology (DArT) markers analyzed by Triticarte Pty Ltd. (Canberra, Australia) as described (Wenzl et al., 2006).

3.4.7 Genetic mapping

Significant differences between single alleles were determined using unpaired t-tests. The genotype data for mapping populations was analyzed using the Joinmap3 software (van Ooijen and Voorrips, 2001). Recombination frequencies were converted into centiMorgan map distances using the Kosambi mapping function. Genetic markers were assembled into linkage groups with the likelihood ratio statistic (LOD) 9.0 and assigned to seven barley chromosomes based on previous published microsatellite and DArT marker maps (Li et al., 2008; Wenzl et al., 2006). The MapQTL5 software (van Ooijen, 2004) was used for non-parametric analysis of variance (Kruskal Wallis test) and interval mapping. The significant genome-wide LOD threshold at p-value of 0.05 was determined by 1,000 permutation tests.

3.4.8 Protein sequence analysis

Protein sequences for *AtCesA1* and *HvCslF* and *HvCesA* families were obtained from the publically available NCBI protein database. Sequence alignment was done using the alignment function and Blosum62 matrix of Geneious 5.4.5 software (Biomatters Ltd; Auckland, New Zealand). Putative phosphorylation sites were searched using the PhosPHat 3.0 (Durek et al., 2009) and Netphos 2.0 (Blom et al., 1999) applications. A predetermined cut-off score of 0.8 was used as it identifies the known phosphorylation sites within the class specific region of AtCesA1 (Chen et al., 2010).

3.5 Results and discussion

3.5.1 DNA sequence analysis of two *HvCslF6* alleles

The analysis of *HvCslF6 DNA sequence* was conducted on the parent lines for the TR251 / CDC Bold population, which shows a major QTL for grain (1,3;1,4)-β-glucan concentration on chromosome 7H (Li et al., 2008). Alignment of QTL position with mapped position for *HvCslF6* (Burton et al., 2008) showed good agreement (data not shown), which supported *HvCslF6* involvement in TR251 / CDC Bold trait variation. To generate DNA fragments for *HvCslF6*, overlapping segments of the coding sequences and introns carried by high (1,3;1,4)-β-glucan

genotype TR251 and low (1,3;1,4)-β-glucan genotype CDC Bold were generated by PCR. The 3' sequence was obtained by genome walking; however, attempts to generate 5' sequences covering the promoter region were unsuccessful. Amplicons of *HvCslF6* were analyzed by DNA sequencing and assembled into a 6.4 kb contig for each parent. The contigs stretched from the translational start codon to 1,000 bp downstream of the translational stop codon and encompassed three exons and two introns as outlined in Figure 3.1A. The DNA and amino acid sequences determined for CDC Bold are presented in supplementary Figure S2.

A DNA sequence alignment of the two *HvCslF6* alleles, revealed 16 SNPs and two indels (Table 3.1; Figure 3.1A). An additional SNP was identified within the second exon by alignment to *HvCslF6* cDNA sequence from Morex (GenBank accession EU267181). The two indels and 13 of the SNPs were positioned within introns or non-coding regions and none of the polymorphic sites affected sequences at exon / intron borders. Both indels were positioned within the first intron, which was 16 bp longer for CDC Bold than for TR251 (Figure 3.1B). Four of the SNPs within exons were silent, but SNP_4105 in the third exon converted an alanine codon in *HvCslF6* of CDC Bold and Morex to a threonine codon in the TR251 sequence. The alteration of the encoded HvCSLF6 peptide occurred at the 590th amino acid of the 948 residue long protein.

3.5.2 Development of *HvCslF6* genetic markers

The markers CSLF6_1028T and CSLF6_1028G, which uses different forward primers (1013-F and 1013G-F) and a common reverse primer (1318-R) were initially developed to target SNP_1029 in intron 1 (Figure 3.1B; Table 3.2). The CSLF6_1028T marker specific for CDC Bold *HvCslF6* allele produces a 306 bp product, whereas the TR251 CSLF6_1028G marker gives a 292 product (Table 3.2). To extend the marker analysis to indel_1178 - 1191 and SNP_1029, the CSLF6_1028T and CSLF6_1028G markers were converted to CAPS markers by utilizing two *Mnl*I restriction sites overlapping the two polymorphic sites (see Figure 3.1B). Upon *Mnl*I digestion, the CAPS marker CSLF6_1028T generated 136-bp and 168-bp fragments for the CDC Bold allele, whereas 208-bp and 84-bp fragments were produced from the TR251 marker CSLF6_1028G (Table 3.2). The CSLF6_1028T and CSLF6_1028G CAPS markers were routinely used in subsequent screening of allele variants for SNP_1029, indel_1178 - 1191, and SNP_1229. A second set of primers (1454-F and 1576-R) was designed to create marker CSLF6_1532-1534 for second indel in the first intron (Figure 3.1C; Table 3.2).

Table 3.1. Summary of *HvCslF6* nucleotide variations.

Polymorphism[1]	Region	Morex[2]	CDC Bold	TR251	Amino acid change
SNP_215	Exon 1	G	G	A	
SNP_453	Intron 1		C	A	
SNP_661	Intron 1		G	C	
SNP_1029	Intron 1		T	G	
Indel_1178-1191	Intron 1		GCCATGAGAAGAG	-	
SNP_1229	Intron 1		T	C	
Indel_1534-1535	Intron 1		TA	-	
SNP_1571	Intron 1		G	C	
SNP_2130	Exon 2	A	G	G	
SNP_2809	Intron 2		T	C	
SNP_2999	Intron 2		C	T	
SNP_3111	Intron 2		A	G	
SNP_3174	Intron 2		C	T	
SNP_3205	Intron 2		T	C	
SNP_3303	Intron 2		G	A	
SNP_4105	Exon 3	G	G	A	A590T
SNP_4842	Exon 3	T	T	C	
SNP_5475	Exon 3	C	C	A	
SNP_6121	3' untranslated		T	C	

[1] Nucleotide positions refer to xx *HvCslF6* DNA sequence (Figure S1). [2] cDNA sequence EU267181.
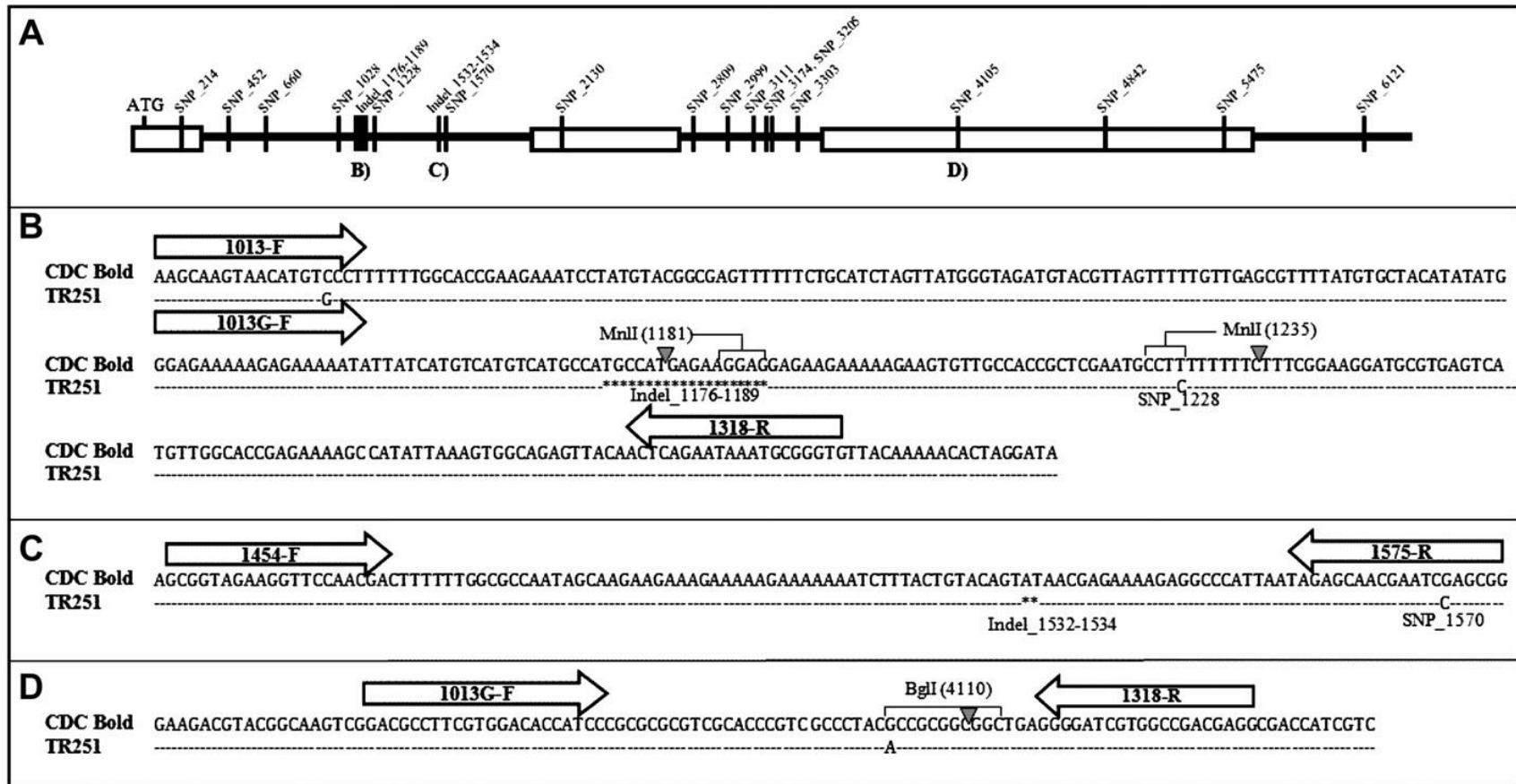
Figure 3.1. Location of *HvCslF6* polymorphism and marker development. A) Schematic illustration of *HvCslF6* with exon sequences represented by horizontal filled bars and black lines illustrating introns and 3' untranslated sequences. Vertical lines show locations of SNPs and indels identified for CDC Bold and TR251 *HvCslF6* (Table 3.1). B) Position of genetic markers developed for *HvCslF6* first intron (Table 3.2). Sequence similarities are represented by dots in the TR251 sequence. SNPs are represented by letters and indels are represented by dashes, location of the two diagnostic *MnlI* restriction site used for CAPS markers are shown. C) Primer position for marker CSLF6_1533-1534 D) Position of marker CSLF6_4105 developed for exon 3 with the *BglI* restriction site indicated (Table 3.2).

Table 3.2. Developed genetic markers for *HvCslF6*.

Marker	Polymorphism targeted	Forward	Reverse	Tm	Enzyme	CDC Bold	TR251
CSLF6_1028T	SNP_1029	1013-F	1318-R	56 °C		306 bp	-
CSLF6_1028G	SNP_1029	1013G-F	1318-R	56 °C		-	292 bp
CAPS CSLF6_1028T	SNP_1029 Indel_1178-1191 SNP_1229	1013-F	1318-R	56 °C	MnlI	136+168 bp	-
CAPS CSLF6_1028G	SNP_1029 Indel_1178-1191 SNP_1229	1013G-F	1318-R	56 °C	MnlI		208+84 bp
CSLF6_1532-1534	Indel_1534-1535	1454-F	1576-R	57 °C		124 bp	122 bp
CSLF6_4105	SNP_4105	4055-F	4136-R	60 °C	BglI	81 bp	55+26 bp

Three of the four SNPs found within the coding regions were silent (SNP_215, SNP_2130 and SNP_4842) and therefore, unlikely to underlie variation in (1,3;1,4)-β-glucan concentration in the CDC Bold / TR251 population. In contrast, SNP_4105 resulting in an A590T amino acid variation in HvCSLF6 was considered a potential cause for trait variation. The SNP_4105 ablated a *Bgl*I restriction site in CDC Bold *HvCslF6* allele (Figure 3.1D), which was utilized to produce CAPS marker CSLF6_4190 (Table 3.2). Oligonucleotide primers 4055-F and 4136-R designed to amplify an 81 bp fragment encompassing SNP_4105 were used in the PCR reactions and upon *Bgl*I digestion of PCR products generated two fragments of 55 bp and 26 bp for the TR251 allele, whereas CDC Bold product remained undigested (81 bp).

The developed CAPS markers CSLF6_1028T, CSLF6_1028G and CSLF6_4190 were used for genotyping 190 lines of the CDC Bold x TR251 population, for which (1,3;1,4)-β-glucan QTLs are known (Li et al., 2008). None of the analyzed lines showed any recombination between the different CSLF6 markers and high (1,3;1,4)-β-glucan lines were preferentially associated with the TR251 allele and low (1,3;1,4)-β-glucan lines frequently carried the CDC Bold allele (Figure 3.2A). The CSLF6 markers were mapped between Bmac0031 and E32M48.3 loci on chromosome 7H map constructed for the population (Li et al., 2008). As expected, the map location coincided with QTL peak for grain (1,3;1,4)-β-glucan concentration explaining 39.1 % of (1,3;1,4)-β-glucan variation in the population.

3.5.3 Validation of *HvCslF6* markers on new mapping population

To test the efficacy of developed CSLF6 markers for prediction of high and low (1,3;1,4)-β-glucan lines, we selected a *de-novo* mapping population, H93174006 / Merit composed of 186 RILs for analysis. The lines and parents were grown in three replications at Vegreville in 2009 and (1,3;1,4)-β-glucan concentration in produced grain was determined. The parental line H93174006 consistently showed a higher (1,3;1,4)-β-glucan concentration (5.13 - 5.42 %) than Merit (4.63 - 5.14 %) for all three replications. Among the lines, the (1,3;1,4)-β-glucan concentrations ranged from 3.8 % to 7.2 % over the three trials and a significant correlation (p < 0.001) existed between the replications (Rep1 / Rep2, 0.57; Rep1 / Rep3, 0.40; Rep2 / Rep3, 0.45). The population demonstrated a near normal distribution (Figure 3.2A), but transgressive
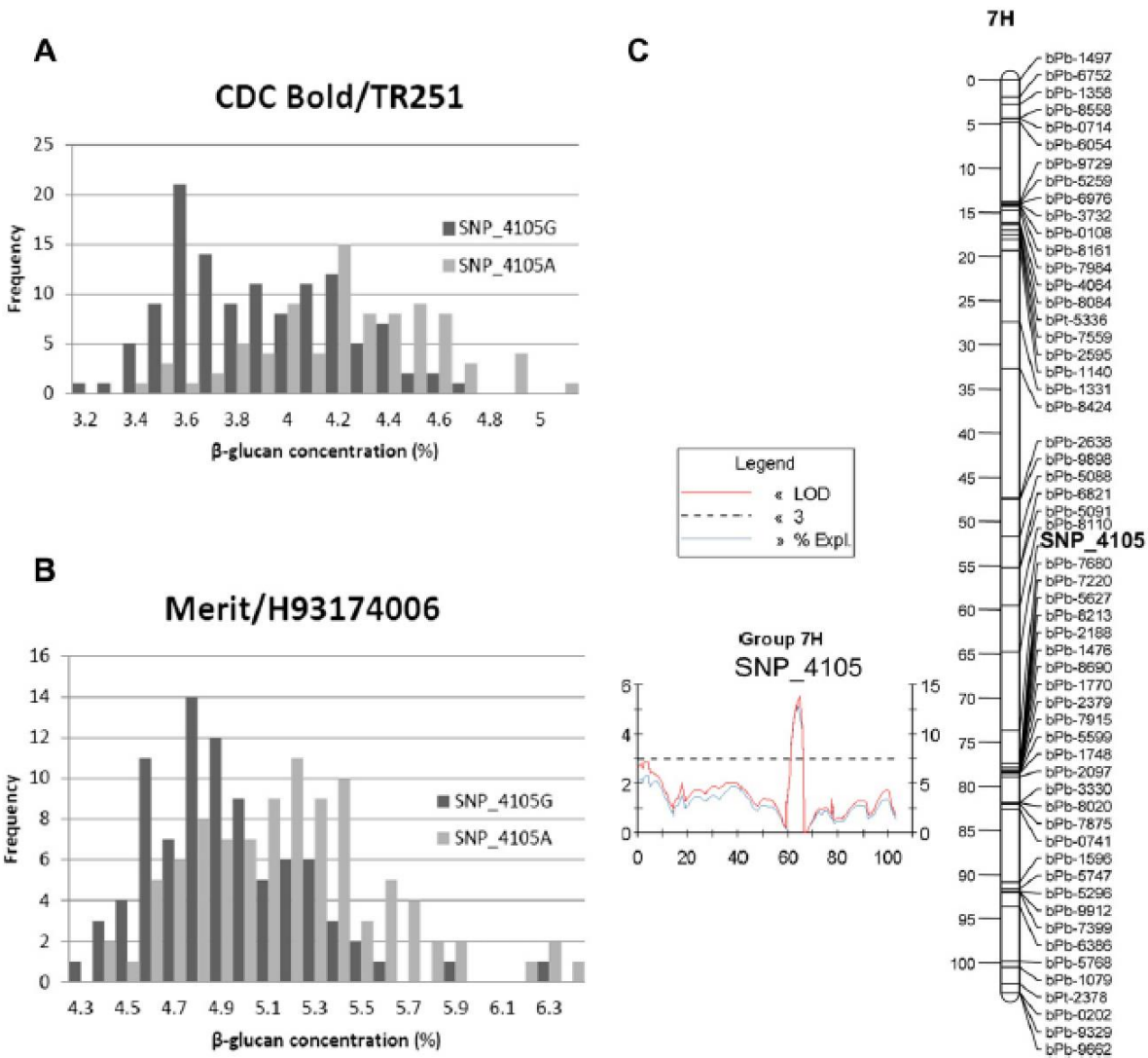
Figure 3.2. Association between Ala / Thr 590 variation and (1,3;1,4)-β-glucan concentration in grain. A) Frequency distribution of (1,3;1,4)-β-glucan concentration and *HvCslF6* haplotypes within the TR251 / CDC Bold and B) H93174006 / Merit populations. C) Location of (1,3;1,4)-β-glucan QTL on chromosome 7H in H93174006 / Merit mapping population.

segregation was evident with most lines having either higher or lower (1,3;1,4)-β-glucan concentrations than the parents. Genotyping of parents using the CSLF6 markers showed that the low (1,3;1,4)-β-glucan parent Merit carried the same allelic differences for *HvCslF6* as CDC Bold, whereas the high (1,3;1,4)-β-glucan parent H93174006 shared *HvCslF6* haplotype with TR251 (Table 3.3). When the population was split by *HvCslF6* allele, a significant difference ($p<0.001$) between (1,3;1,4)-β-glucan concentrations was revealed by a paired t-test.

To validate the CSLF6 markers on the Merit / H93174006 population, genotyping was expanded to 28 SSR markers and 626 DArT markers showing polymorphism between the parent lines. The linkage groups assembled from genotype data represented about 70 % of the barley genome when aligned to the 2006 DArT marker map (Wenzl et al., 2006). The generated map of chromosome 7H showed good representation of markers and covered 115 cM of the chromosome with marker order similar to previously published barley maps (Li et al., 2008; Wenzl et al., 2006). The CSLF6 markers were mapped between pBp-8110 and Bmac0031 (Figure 3.2B), which corresponded well to position on CDC Bold / TR251 map. When the phenotype and genotype data were tested by a non-parametric analysis of variance (Kruskal-Wallis test), the highest association between (1,3;1,4)-β-glucan concentration and genotype was found for the CSLF6 markers showing K values above 23 ($p < 0.001$). Similar to the CDC Bold / TR251 population, a major QTL for (1,3;1,4)-β-glucan concentration was mapped to the CSLF6 loci, where a LOD peak of 5.8 was obtained (Figure 3.2B). The 7H QTL accounted for over 13 % of the trait variation within the population.

3.5.4 Mapping populations with CDC Bold / TR51 polymorphism for *HvCslF6* show (1,3;1,4)-β-glucan QTL on chromosome 7H

To determine the universality of the CSLF6 markers, we genotyped parents from six mapping populations, for which QTLs for the (1,3;1,4)-β-glucan trait are known. The TR306 / Harrington population shows three (1,3;1,4)-β-glucan QTLs, including one on chromosome 7H, where the TR306 allele associates with increased grain (1,3;1,4)-β-glucan content (Mather et al., 1997). The Steptoe / Morex population shows a 7H QTL associated with higher wort (1,3;1,4)-β-glucan contributed by Steptoe alleles; however high grain (1,3;1,4)-β-glucan content is associated with Morex alleles on chromosome 1H and 2H (Han et al., 1995). Alexis / Sloop,

68

Table 3.3. *HvCslF6* analysis of barley genotypes used in mapping and expression studies.

Mapping population	7H QTL	Parent	SNP_1028	Indel_1176 - 1189	SNP_1228	SNP_4105
CDC Bold / TR251[a]	Yes	TR251[i]	G	-	C	A
		CDC Bold	T	GCCATGAGAAGAG	T	G
Merit / H93174006[b]	Yes	H93174006[i]	G	-	C	A
		Merit	T	GCCATGAGAAGAG	T	G
Steptoe / Morex[c]	Yes	Steptoe[k]	G	-	C	A
		Morex	T	GCCATGAGAAGAG	T	G
TR306 / Harrington[d]	Yes	TR306[k]	G	-	C	A
		Harrington	T	GCCATGAGAAGAG	T	G
Alexis / Sloop[e]	No	Sloop[k]	T	GCCATGAGAAGAG	T	G
		Alexis	T	GCCATGAGAAGAG	T	G
Arapiles / Franklin[e]	No	Franklin[k]	T	GCCATGAGAAGAG	T	G
		Arapiles	T	GCCATGAGAAGAG	T	G
CM27 / Gaidner[f]	No	Gairdner[i]	T	GCCATGAGAAGAG	T	G
		CM27	T	GCCATGAGAAGAG	T	G
Dicktoo / Morex[g]	No	Dicktoo[i]	T	GCCATGAGAAGAG	T	G
		Morex	T	GCCATGAGAAGAG	T	G
Expression study[h]		Himalayan[i]	T	GCCATGAGAAGAG	T	G
		Sloop	T	GCCATGAGAAGAG	T	G

[a] Li et al., 2008; [b] current study; [c] Han et al., 1995; [d] Mather et al., 1997; [e] Panozzo et al., 2007; [f] Wei et al., 2009; [g] Ozeil et al., 1996; [h] Burton et al., 2008. [i] Parent with higher grain β–glucan concentration. [k] Parent with higher wort β–glucan concentration.

Arapiles / Franklin (Panozzo et al., 2007), CM27 / Gairdner (Wei et al., 2009), and Dicktoo / Morex (Oziel et al., 1996) populations have all revealed QTLs on chromosomes 1H, 2H, 4H and 5H but not on chromosome 7H. When the parental lines of the various populations were genotyped using the CSLF6 markers, we found that parents of populations showing a QTL on 7H followed a similar pattern to that of TR251 and CDC Bold. Thus, each parent contributing towards a higher (1,3;1,4)-β-glucan concentration shared *HvCslF6* haplotype with TR251, whereas parents with lower (1,3;1,4)-β-glucan showed the CDC Bold haplotype. For the four populations which do not display a QTL on chromosome 7H, all parental lines shared *HvCslF6* haplotype with CDC Bold.

The 7H QTL for the Steptoe / Morex population relates to (1,3;1,4)-β-glucan concentration in wort rather than grain, suggested that 7H QTL may have a higher effect on (1,3;1,4)-β-glucan solubility in this population. This may be explained by higher *HvCslF6* expression in Steptoe, as high *HvCsl*F6 expression is known to increase soluble (1,3;1,4)-β-glucan levels in transgenic barley (Burton et al., 2010) and wheat (Nemeth et al., 2010). Conversely, *RNAi* mediated down-regulation of *HvCsl*F6 transcription in wheat significantly reduces the hot water extractible (1,3;1,4)-β-glucan (Nemeth et al., 2010). On the other hand Morex is an elite malting barley selected for and these differnces may be due to differences in beta-glucanase activity in the wort.

To study if any of CSLF6 markers could be associated with *HvCslF6* expression levels, the barley lines Sloop and Himalayan were selected for genotyping. The Himalayan line shows higher *HvCslF6* expression and (1,3;1,4)-β-glucan concentration than Sloop (Burton et al., 2008). However, genotyping of the Sloop and Himalayan lines did not reveal any polymorphism for CSLF6 markers (Table 3.3). Both lines shared the CDC Bold haplotype, which suggest that higher level of *HvCslF6* expression in Himalayan barley endosperm is not due to any allele differences found in this study. As these polymorphic sites did not explain the differences in expression levels seen between Himalayan and Sloop, they were considered less likely to cause variation in *HvCslF6* expression between TR251 and CDC Bold. A preliminary evaluation of *HvCslF6* transcript levels has not indicated any major difference in expression levels between CDC Bold and TR251 during kernel development (unpublished results). However, further studies are needed to validate this initial observation.

3.5.5 Canadian barley genotype screening

To determine the relative abundance of the *HvCslF6* alleles in a wider context, 150 advanced breeding lines from Alberta Agriculture Food and Rural Development were screened using the CSLF6_4190, CSLF6_1028T, CAPS CSLF6_1028T and CAPS CSLF6_1028G markers. Only seven lines did not conform to the CDC Bold and TR251 pattern of alleles and showed recombination between markers. The majority of lines (76 %) were consistent with the CDC Bold haplotype. This over-representation probably reflects strong selection for low (1,3;1,4)-β-glucan concentration in the development of malting barleys in Canada. The large number of genetic variations within this gene between TR251 and CDC Bold makes it likely that other mutations could be present elsewhere in the *HvCslF6* gene in this larger barley population.

3.5.6 Putative phosphorylation motif precedes CSLF6_A590T substitution

HvCslF6 belongs to the CesA / Csl superfamily composed of members with eight predicted transmembrane domains, a cytoplasmic CesA domain, three conserved aspartate residues and a QxxRW signature motif within the catalytic domain (Chen et al., 2010; Nühse et al., 2004). The variant A590T amino acid difference between CDC Bold and TR251 sequences was located within the class specific region (Chen et al., 2010), which falls between the second and third conserved aspartate residues (Figure 3.3). The corresponding region in Arabidopsis contains two functional phosphorylation sites, both of which regulate enzyme activity and are important for cell wall development (Nühse et al., 2004) by regulating microfibril development (Chen et al., 2010). A test for possible phosphorylation sites within the two *HvCslF6* variants using PhosPhat 3.0 indicated seven phosporylation sites with a score > 0.8 for both CDC Bold and TR251 *HvCslF6* variants, whereas 30 sites with a probability score > 0.8 were identified using Netphos 2.0. One of the high-scoring sites (Y589 PhosPhat3.0 score 1.1; S587 NetPhos2.0 score 0.98) was found immediately N-terminal of the A590T substitution site and carried a SHPSPY motif (Y589 [A / T]; score = 1.1 [A] score = 1.2 [T]). *In silico* analysis of the HvCSLF6 proteins showed conservation of this SHPSPYAA site within barley, rice and wheat (data not shown). Also, *Hv*CSLF1, 4, 7, 8, and 9 as well as *Hv*CSLH1 were predicted to carry a phosphorylation site within this region, although the SHPSPY[T / A]AAA motif was not conserved in the paralogous proteins. Whether differences in HvCSLF6 phosphorylation status caused by A590T.
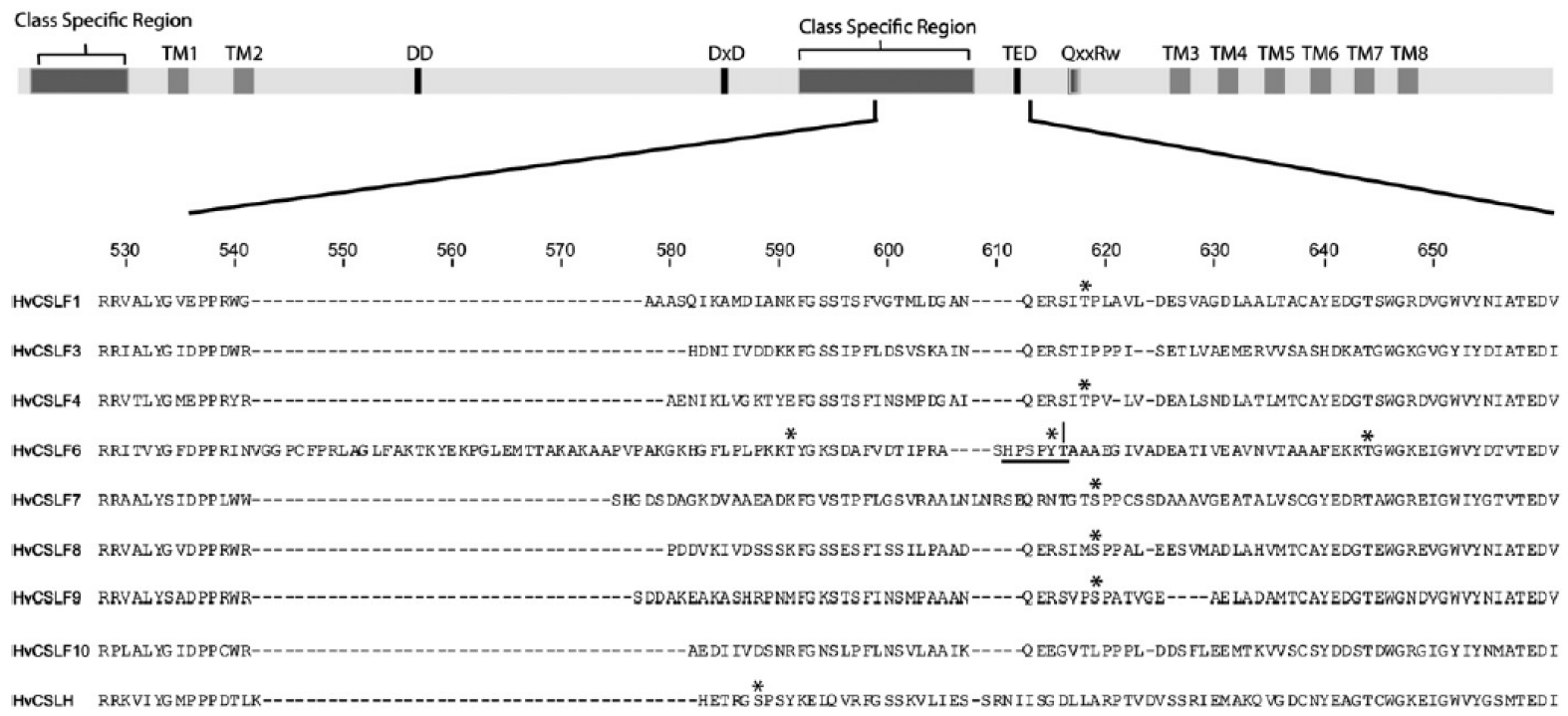
Figure 3.3. Alignment of HvCSLF and HvCSLH class specific regions. Top: Position of class specific region in relation to the transmembrane domains (TM), conserved aspartic acid residues (DD, DxD, TED) and catalytic domain (QxxRW) in generic CESA / CSL proteins. Bottom: Sequence alignment of class specific regions with third conserved aspartic residue as landmark (TED). Predicted phosphorylation sites are highlighted with an asterisk. The position of the amino acid variant described in HvCslF6 is indicated by a line.

variation is the underlying cause for 7H (1,3;1,4)-β-glucan QTL remains to be confirmed by further studies.

3.6 Conclusions

The *HvCslF6* allele variants described in this study resulted in the development of six markers related to grain (1,3;1,4)-β-glucan concentration (Table 3.2). The most significant allelic variation was SNP_4105 predicted to cause differences in *Hv*CSLF6 phosphorylation status which could affect enzymatic activity. Whether this variation is the underlying cause for 7H (1,3;1,4)-β-glucan QTL remains to be determined by further studies. However, this amino acid variation could affect the fine structure of (1,3;1,4)-β-glucan, which makes it valuable to breeders. Selection for the TR251 *HvCslF6* allele may increase the amount of soluble fiber in barley making it more desirable to the increasingly health conscious consumer. Equally, a malting barley breeder would benefit from CSLF6 markers in being able to select for reduced concentration of soluble (1,3;1,4)-β-glucan.

CHAPTER 4

GENOTYPE, ENVIRONMENT AND GxE INTERACTION INFLUENCE (1,3;1,4)-β-GLUCAN FINE STRUCTURE

4.1 Study 2*

In this study, HPAEC-PAD was used to phenotype 91 RIL6 lines produced in two environments in order to determine heritability of (1,3;1,4)-β-glucan fine structure and an associated genomic region.

4.2 Abstract

Ninety-one RIL5 lines were phenotyped for (1,3;1,4)-β-glucan grain content, cellotriose content (DP3), cellotetraose content (DP4) and cellotriose:cellotetraose (DP3:DP4) ratio in two environements. DP3, DP4, (1,3;1,4)-β-glucan and total DP3+DP4 were strongly positively correlated to each other suggesting no preference for DP3 or DP4 subunit production in high or low (1,3;1,4)-β-glucan lines. DP3:DP4 ratio showed no strong correlation with any other measured trait. Significant affects arising from genotype and environment were associated with (1,3;1,4)-β-glucan, DP3 and DP4 and DP3:DP4 ratio. Only DP3:DP4 ratio showed a significant GxE interaction. Single marker analysis showed an association between marker CSLF6_4105 and (1,3;1,4)-β-glucan fine structure in Vegreville but not Castor supporting significant GxE interaction in (1,3;1,4)-β-glucan fine structure. Association mapping of candidate markers in 119 barley genotypes of diverse origin grown in greenhouse conditions shows that CSLF6_4105 is associated with (1,3;1,4)-β-glucan concentration, Bmac273e is associated with (1,3;1,4)-β-glucan concentration and DP3:DP4 ratio and Bmac504 and Bmac211 are associated with DP3:DP4 ratio. This study suggests that DP3:DP4 ratio is strongly affected by genotype and may be influenced by selective breeding.

4.3 Introduction

(1,3;1,4)-β-glucan is a glucan polymer that accumulates in the cell walls of grasses. (1,3;1,4)-β-glucan is a mixed linkage linear polymer made up by β-D-glucopyranosyl molecules forming either β-1,3 or β-1,4 linkages. The β-1,4-linked molecules are constituted primarily of non-randomly arranged cellotriose (DP3) and cellotetraose (DP4) units present in ratios ranging from 2.3:1 to 4.8:1 depending on genotype (Collins et al., 2010). More than 90 % of (1,3;1,4)-β-glucan consists of DP3 and DP4 subunits joined together by 1,3 linkages and only a small percentage are represented by higher order cellodextrin (DP5+) units (Lazaridou et al., 2004; Wood and Fulcher, 1983).

(1,3;1,4)-β-glucan is a desired functional, bio-active ingredient of interest in human nutrition (Cui and Wood, 2000). The health benefits of including high viscosity (1,3;1,4)-β-glucan in human nutrion include reduced plasma cholesterol and a better control of postprandial serum glucose levels (Lazaridou and Biliaderis, 2007; Wood, 2007; Bhatty, 1999). DP3:DP4

ratios influence the viscosity, gelling properties and rheological behavior of (1,3;1,4)-β-glucan (Wood, 2010). The structural features of (1,3;1,4)-β-glucan are important determinants of their their physiological responses when they are considered as ingredients in cereal based foods and other formulated products (Cui et al., 2000; Lazaridou et al., 2003).

The fine structure of (1,3;1,4)-β-glucan is influenced both by genetic and environmental factors (Wood, 2007; Doehlert and Simsek, 2012). DP3:DP4 ratio differ between grass species; barley has the highest and most varied DP3:DP4 ratio ranging from 2.8 - 3.3:1, rye is close to barley with 3.0 - 3.2:1 and oats show the lowest and least divergent ratios with a value of 2.1 - 2.3:1 (Wood, 2007). In oats it was found that wet environments were associated with a higher DP3:DP4 ratio (Doehlert and Simsek, 2012). Hulled, hulless, long awn, short awn and low amylose (waxy) barley genotypes showed no effect of environment on (1,3;1,4)-β-glucan fine structure, but waxy starch phenotype increases the DP3:DP4 ratio (Wood et al., 2003). Over expression of Cellulose synthase like F4 (*HvCslF4*) in Golden Promise increases DP3:DP4 ratio from 2.8:1 to 3.1:1 while overexpression of *HvCslF6* in the same genetic background decreases the DP3:DP4 ratio to 2.1:1 (Burton et al., 2011). RNAi mediated silencing of *TaCslF6* in wheat (*Triticum aestivum*) did not affect the DP3:DP4 ratio but reduces both the molecular weight and total concentration of grain (1,3;1,4)-β-glucan (Nemeth et al., 2010).

The main objective of this study was to determine the interaction of environment and genotype on (1,3;1,4)-β-glucan fine structure in two row, spring barley genotypes with normal starch characteristics. A recombinant inbred line (RIL) population (sixth generation) was grown at two locations in Alberta, Canada. The seeds from two biological replicates at each site were used to determine (1,3;1,4)-β-glucan grain content and its fine structure. These results suggests that DP3, DP4, DP3:DP4 ratio and (1,3;1,4)-β-glucan are significantly affected by genotype and environment.

4.4 Materials and methods

4.4.1 Plant material

A two-row, spring barley (*Hordeum vulgare* L.) genotype Merit (Busch Agricultural Resources LLC; Fort Collins, Co, USA) and H93174006 (derived from a H92076F1 x TR238 cross made at the Field Crop Development Center, Lacombe, Alberta, Canada) were crossed to produce a F1 hybrid and subsequent RIL population of barley (CDC Lacombe, Alberta, Canada).

The Merit / H93174006 population was advanced by single seed descent through the F3 and F4 generations in a greenhouse at Lacombe during 2007 and 2008. The F5 seeds from each F4 plant were bulked and advanced to produce F4:6 seeds for the study.

A set of 184 F5 RILs and the two parental lines were planted under rain-fed conditions in 2009 at Vegreville (53 ° 31' N, 112 ° 6' W, 639.3 m altitude), with the Malmo series of an eluviated black chemozemic soil, and at Castor (52 ° 8' N, 111 ° 54' W, 807.7 m altitude), with a dark brown chemozemic soil, at both locations in Alberta, Canada. The two sites were characterized by distinct soil moisture conditions. The average annual precipitation and within season rainfall (June to August) from 1977 to 2007 was 340 mm and 172 mm at Castor, compared with Vegreville which had 382 mm and 193 mm, respectively (AgroClimatic Information Service (ACIS) 2009; Environment Canada 2009). The year of 2009 had higher rainfall in Castor (283 mm) compared to Vegreville (275 mm) according to weather stations situated at Vegreville and Halkirk (http://agriculture.alberta.ca/acis/alberta-weather-data-viewer.jsp). Weeds were controlled by Round-up Weathermax application before seeding and by Achieve 40DG and Buctril M spraying later in the growing season. Grains were harvested and stored at room temperature until used for analysis.

A total of 119 barley genotypes were grown to assess candidate markers associated with (1,3;1,4)-β-glucan grain concentration and DP3:DP4 ratio. The genotypes were grown in the University of Saskatchewan greenhouse under a night / day temperature range of 19 - 28 °C with an 18 hour photoperiod with an average photosynthetically active radiation of 385 μmol $^{-2}$s^{-1}. Each genotype was grown in triplicate. Barley lines were grown over 150 days and harvested at maturity. Grains were stored at room temperature until used for analysis.

4.4.2 Determination of (1,3;1,4)-β-glucan grain concentration

(1,3;1,4)-β-glucan grain concentration was determined on the parents and eighty-nine randomly selected RILs from two biological replications in both environments. Grain samples (10 g) were milled to flour using an Udy-Mil Cyclone sample grinder (UDY Corporation, Fort Collins, CO, USA) equipped with a 0.5 mm sieve. To determine total (1,3;1,4)-β-glucan grain concentration, flour samples (100 mg) were subjected to a lichenase digestion similar to the AACCI method 32 − 23.01 as described in (1,3;1,4)-β-glucan (mixed linkage) Kit manual (Megazyme, Wicklow, Ireland) with minor modifications as follows. Flour (100 mg) was placed

in 15 ml screw cap tubes and was wetted with 200 μl of 50 % (v / v) ethanol. 4 ml of 20 mM sodium phosphate buffer (pH 6.5) was added to the wetted flour sample. The flour suspension was boiled for 5 min with brief vortexing every 60 s to keep the material well suspended and allow effective inactivation of native enzymatic activity. The solution was vortexed and allowed to equilibrate to 50 °C for 5 min. Lichenase (5 U, 200 μl; Megazyme, Wicklow, Ireland) was then added and the solution was incubated at 50 °C with constant shaking for 120 min. 5 ml of 200 mM sodium acetate (pH 4) was added to stop lichenase digestion. The digested samples were centrifuged at 1500 g for 15 min, and the supernatant was collected and filtered through a C18 column (Thermos Scientific. Bellefonte, PA, USA). Aliquots of the filtered supernatant containing the lichenase digested (1,3;1,4)-β-glucan were saved for determination of (1,3;1,4)-β-glucan grain concentration and fine structure. The lichenase digest used for (1,3;1,4)-β-glucan grain concentration was incubated with beta-glucosidase for 20 min at 50 °C and the released glucose concentration was analysed as described (AACC method 32 – 23.01). Barley genotypes included in the association mapping panel were ground as described for the RIL population but (1,3;1,4)-β-glucan was determined by the calcofluor flow injection method (Aastrup and Jørgersen, 1988) using 25 mg of flour. Each sample was analysed in triplicate, and pooled averages were used in subsequent statistical analysis.

4.4.3 HPAEC-PAD determination of DP3 and DP4

Freshly prepared lichenase digested (1,3;1,4)-β-glucan containing filtered supernatant was diluted to 1:10 in degassed distilled water before samples were analysed by HPAEC-PAD. The filtered, diluted lichenase digest solution were injected into a CarboPac PA1 column using a Dionex ICS 5000 system equipped with an auto-sampler (Dionex, Sunnyvale, CA, USA). Samples were eluted at 0.8 ml ∕ min with 200 mM sodium hydroxide for 2 min followed by a 0 to 250 mM sodium acetate gradient in 200 mM sodium hydroxide over 15 min. The gradient was followed by a 10 min flush with 200 mM sodium hydroxide. Sample blanks were included after every fifth run to ensure absence of column contamination. (1,3:1,4)-β-gluco-triose (Megazyme, Wicklow, Ireland cat. No O-BGTRIB) and (1,3:1,4)-β-gluco-tetraose (Megazyme, Wicklow, Ireland cat. No O-BGTETB) were included as controls under the same chromatographic conditions and in each set of the two. Each sample was analysed in duplicate and values reported are the average value obtained. Standard curves were calculated for each run. Samples of oat and

barley flour supplied with the Megazyme β-glucan (mixed linkage) kit were used as internal controls. The performance of optimized and accuracy of the method was evaluated by calculating the coefficient of determination (R^2), level of detection (LOD), level of quantification (LOQ) and intermediate precision. The suitability of chromatography column was assessed by determining peak resolution and peak asymmetry. Calibration curves using five concentrations [7.8, 15.6, 31.25, 62.5, 125 μM injection (volume 10 μL)] for each standard was prepared to develop a regression equation and calculate R^2. On the basis of calibration curves, LOD and LOQ scores were calculated using formulae found in ICH harmonized tripartite guidelines (2005).

4.4.4 Isolation of genomic DNA

Leaves of each barley plant were harvested at the 5 to 10 leaf stage, frozen in liquid nitrogen and stored at -80 °C before DNA was extracted and quantified as described (Li et al., 2008). Primers used for CSLF6_4015 consisted of a forward primer 4055F - GACGCCTTCG TGGACACCATCC and reverse primer 4136R – CTCGTCGGCCACGATCCCCT. The genotyping reactions for marker CSLF6_4105 consisted of 100 ng template DNA, 5 pmol of forward and reverse primers, 10 mM TriseHCl pH 8.3, 3.5 mM $MgCl_2$, 25 mM KCl, 20 mM dNTP, 0.5 U of Taq Polymerase (Thermo Fisher Scientific, Waltham, MA, USA) in a total volume of 25 μL. Amplifications were done in an Eppendorf AG Cycler (Eppendorf, Hamburg, Germany) programmed for an initial denaturation at 94 °C for 5 min followed by 40 cycles of 45 s denaturation at 94 °C, 20 s annealing at 60 °C and 20 s extension at 72 °C. PCR products were digested with restriction enzyme BglI (New England Biolabs, Ipswich, MA, USA), respectively, followed by analysis of digestion products by two percent (w / v) agarose gel electrophoresis.

4.4.5 Statistical analysis

General linear model (GLM) was applied to calculate analysis of variance (ANOVA) using Minitab 16 statistical software (Minitab Inc., State College, PA, USA). Mean sum of squares (MSS) from ANOVA was used to calculate broad sense heritability (h^2) (Singh et al., 1993), categorized as low (≥ 30 %), moderate (30 to 60 %) and high (≤ 60 %) (Robinson et al., 1949). Variation due to genetics, environment, GxE interaction and error were calculated by dividing the individual components least square by the total adjusted least squares. Correlations

were determined using Pearson correlation function. Single marker association was done by one way ANOVA using marker class as the fixed factor and phenotype as response variable.

4.5 Results and discussion
4.5.1 Plant material and site selection

The Merit / H93174006 RIL5 population was selected due to the wide range (1,3;1,4)-β-glucan grain concentration (see below) and normal starch charecteristics. The RIL5 population was planted under rain fed conditions in 2009 at Vegreville, and at Castor, both locations in Alberta, Canada. The two sites are characterized by distinct soil moisture conditions with the Castor site traditionally dryer than the site at Vegreville. However, the year of 2009 had timely and higher rainfall in Castor (283 mm) compared to Vegreville (275 mm).

4.5.2 (1,3;1,4)-β-glucan grain concentration

Grain (1,3;1,4)-β-glucan concentration was consistently higher in H93174006 (5.57 ± 0.23 %) compared to Merit in the Castor site (4.58 ± 0.02 %). The same trend was seen in Vegreville with H93174006 having higher (1,3;1,4)-β-glucan grain concentration (5.34 ± 0.09 %) than Merit (4.89 ± 0.26 %). Mid parent values across both environments were 5.45 ± 0.21 % for H93174006 and 4.78 ± 0.11 % for Merit. In the RILs the (1,3;1,4)-β-glucan values were normally distributed across both environments and in all biological replications (Figure 4.1). Transgressive segregation was observed in all replications and environments. The widest range of (1,3;1,4)-β-glucan concentrations observed was in the Castor environment (4.16 ± 0.10 % to 6.22 ± 0.01 %; mean 5.11 ± 0.48 %) but the Vegreville (1,3;1,4)-β-glucan concentrations had a higher mean value (5.21 ± 0.38 %) but a smaller range (4.40 ± 0.20 % to 6.35 ± 0.23 %). The higher (1,3;1,4)-β-glucan grain concentrations found in Vegreville, the dryer environment, was expected as a negative correlation between grain (1,3;1,4)-β-glucan concentration and moisture has been previously reported (Perez-Vendrell et al., 1996).
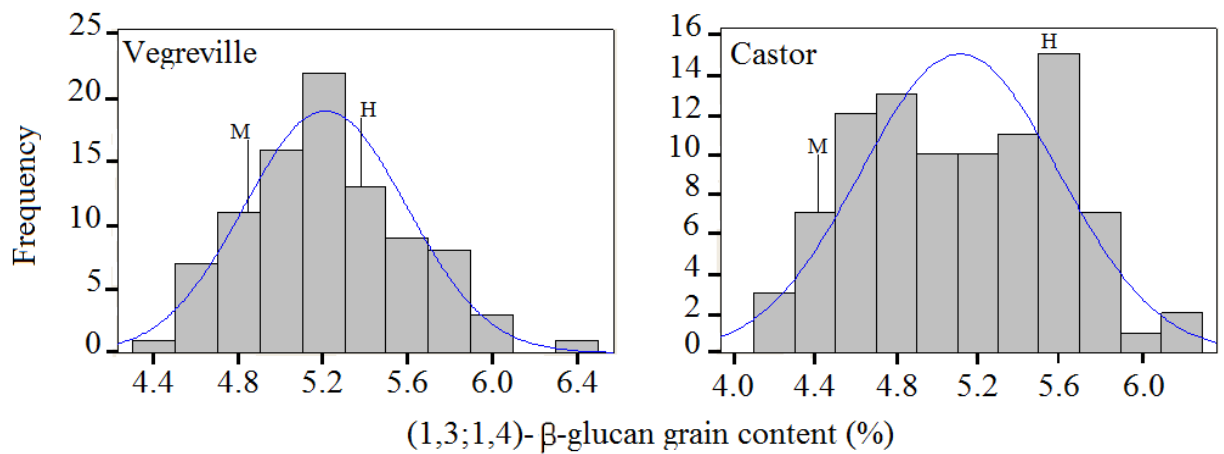
Figure 4.1. Frequency distribution of (1,3;1,4)-β-glucan of 91 barley RILs. Parental values are indicated as M (Merit) and H (H93174006).

4.5.3 HPAEC-PAD separation of DP3 and DP4 fractions from lichenase digestion

HPAEC-PAD analysis of the carbohydrate fractions showed excellent separation of several small oligosaccharides released from lichenase digestion. (1,3:1,4)-β-gluco-triose standard showed a symmetric peak eluting at 12 min, while the (1,3:1,4)-β-gluco-tetraose standard showed a symmetric peak eluting at 13 min (Figure 4.2). To quantify DP3 and DP4 peaks a serial dilution of the DP3 and DP4 standards was used to identify a range of concentrations corresponding to a linear response curve. Standards showed linear response to concentrations between 3.9 μM to 125 μM for both DP3 and DP4 with higher concentrations showing a nonlinear response. The concentration range showing linear response corresponds to 0.20 mg / 100 mg to 6.31 mg / 100 mg for DP3 and 0.26 mg / 100 mg to 8.33 mg / 100 mg for DP4. Intermediate precision was calculated to be 6.83 % for DP3 and 7.37 % for DP4. The quantitation limit (LOQ) was calculated at 0.35 mg / 100 mg for DP3 and 1.08 mg / 100 mg for DP4 which is lower than all concentrations determined for the barley lines subjected to lichenase digestion (Table 4.1). HPAEC-PAD analysis of the lichenase digest revealed several peaks that could be identified as sucrose, raffinose, maltose, (1,3;1,4)-β-gluco-triose and (1,3;1,4)-β-gluco-tetraose (Figure 4.2). Other peaks were putatively identified as cellopentose, cellohexose and higher order cellobiose units. Due to the unavailability of standards and relatively low measurable areas at this dilution these peaks were considered uninformative.

4.5.3.1 Cellotriose (DP3) concentrations

The two parents significantly differed in DP3 concentration in both environments (Figure 4.3). In the Castor environment DP3 concentration was higher in H93174006 (3.48 ± 0.15 mg / 100 mg) compared to Merit (3.18 ± 0.04 mg / 100 mg). Vegreville was similar in respect to DP3 levels with H93174006 having higher DP3 concentration (3.89 ± 0.07 mg /100 mg) than Merit (3.46 ± 0.02 mg / 100 mg). Mid parent DP3 concentrations across both environments were 3.68 ± 0.23 mg / 100 mg for H93174006 and 3.32 ± 0.16 mg / 100 mg for Merit. DP3 concentrations were normally distributed in both Castor and Vegreville. Mean DP3 concentrations were higher for the Vegreville environment compared to Castor (3.58 ± 0.30 mg / 100 mg Vegreville, 3.46 ± 0.36 mg / 100 mg Castor) but Castor had a wider range of DP3 concentrations (2.81 - 4.31 mg / 100 mg; Vegreville, 2.62 - 4.17 mg / 100 mg Castor). The DP3 concentrations showed transgressive segregation similar to the (1,3;1,4)-β-glucan grain concentration (Figure 4.3).
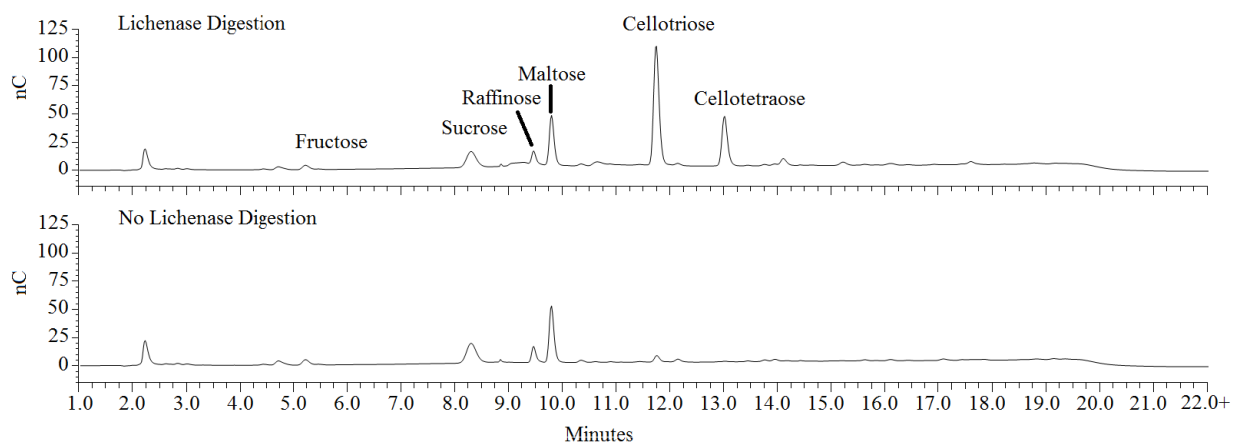
Figure 4.2. HPAEC-PAD chromatograph of treated and untreated barley flour. HPAEC-PAD chromatograph showing separation of fructose, sucrose, raffinose, maltose, cellotriose and cellotetraose on a lichenase digested sample (Top) and no lichenase control (Bottom).

Table 4.1. HPAEC-PAD response to chromatographic conditions.

Standard	t_R (min)	R^2	LOD (μM)	LOQ (μM)	IP (%)	PA	PR
DP3	11.82	0.997	1.03	1.99	6.83	1.13	6.86
DP4	13.07	0.997	3.13	6.02	7.37	1.20	6.86
N	5	5	5	5	5x5	5	5

t_R retention time, R^2 coefficient of determination, LOD level of detection, LOQ level of quantification, IP intermediate precision, PA peak asymmetry, PR peak resolution.
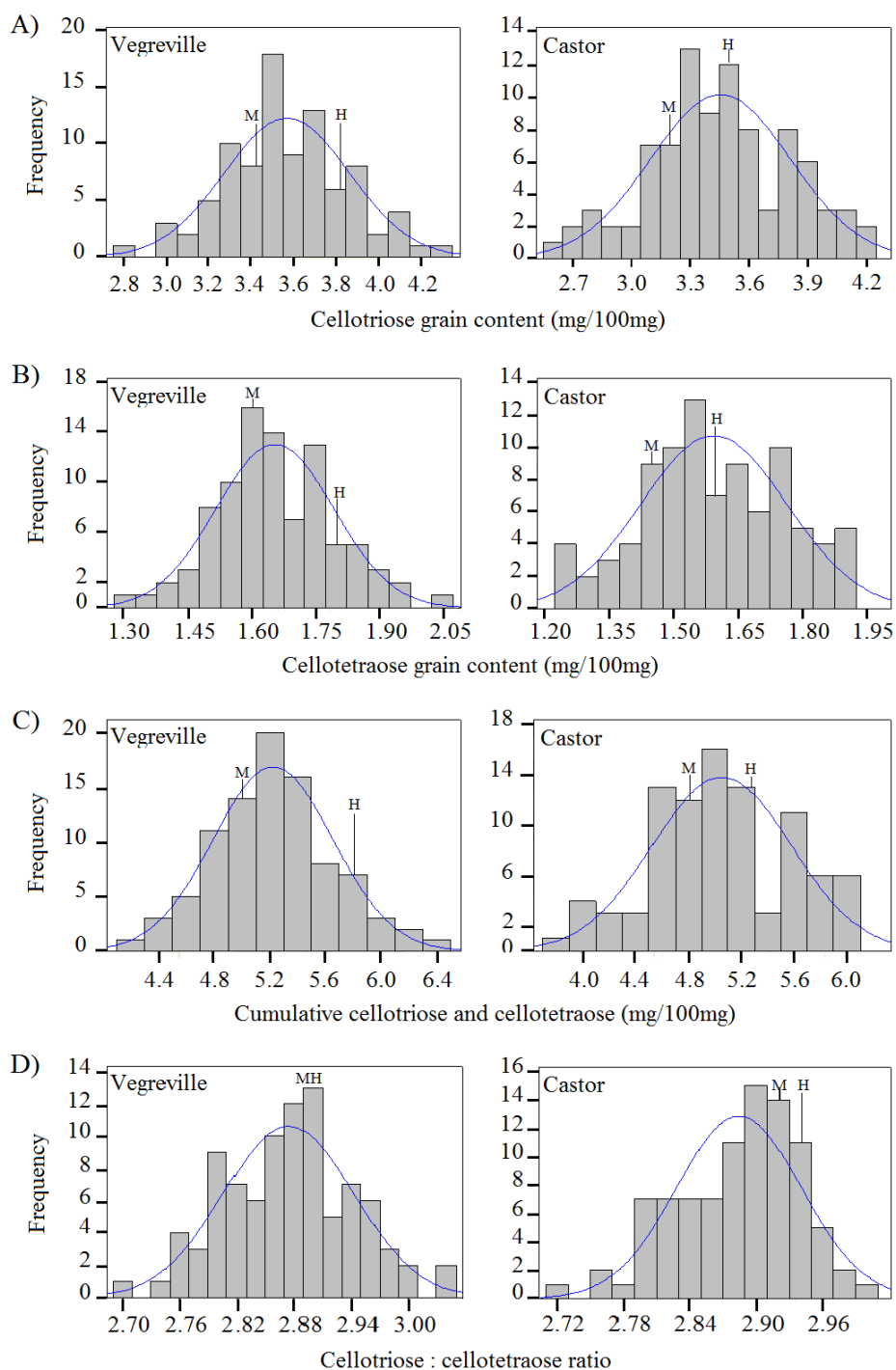
Figure 4.3. Frequency distribution of HPAEC-PAD characterized traits. A) Cellotriose concentration B) Cellotetraose concentration C) Cumulative cellotriose and cellotetraose concentration D) Cellotriose to cellotetraose ratio.

4.5.3.2 Cellotetraose (DP4) concentrations

DP4 concentraions also showed significant differences between the two parents in both environments (Figure 4.3). The Castor environment showed higher concentrations of DP4 in H93174006 (1.57 ± 0.07mg / 100mg) compared to Merit (1.45 ± 0.02 mg / 100 mg). H93174006 also showed higher levels of DP4 in the Vegreville environment (1.79 ± 0.02 mg / 100 mg) compared to Merit (1.60 ± 0.04 mg / 100 mg). Mid parent DP4 concentrations across both environments were 1.68 ± 0.03 mg / 100 mg for H93174006 and 1.53 ± 0.02 mg / 100 mg for Merit. DP4 concentrations were normally distributed in both the Castor and Vegreville environments. The range and mean DP4 concentrations were greater in Vegreville (1.31 ± 0.09 - 2.00 ± 0.06 mg / 100 mg; mean 1.65 ± 0.14 mg / 100 mg) compared to Castor (1.24 ± 0.09 -1.91 ± 0.03 mg / 100 mg; mean 1.59 ± 0.17 mg / 100 mg).

4.5.3.3 Cumulative cellotriose and cellotetraose (DP3+DP4) concentrations

Cumulative cellotriose and cellotetraose (DP3+DP4) were calculated as the sum of the DP3 and DP4. Mid parent cumulative DP3+DP4 concentrations across both environments were 5.36 ± 0.35 mg / 100 mg for H93174006 and 4.84 ± 0.24 mg / 100mg for Merit. H93174006 had higher concentrations of cumulative DP3+DP4 in Castor (5.04 ± 0.22 mg / 100 mg) compared to Merit (4.63 ± 0.05 mg / 100 mg). In Vegreville, H93174006 had higher cumulative DP3+DP4 (5.68 ± 0.24 mg / 100 mg) than that of Merit (5.06 ± 0.46 mg / 100 mg). Vegreville DP3+DP4 concentrations ranged from 4.11 ± 0.18 - 6.36 ± 0.23 mg / 100 mg and Castor ranged from 3.86 ± 0.25 - 6.09 ± 0.03mg / 100 mg. Mean concentrations of cumulative DP3+DP4 were significantly higher in Vegreville (5.22 ± 0.43 mg / 100 mg) compared to Castor (5.05 ± 0.53 mg / 100 mg).

Means for DP3, DP4, DP3+DP4 and (1,3;1,4)-β-glucan were all significantly higher in Vegreville compared to Castor. Castor received more precipitation, earlier, compared to Vegreville during the growing season of 2009 making it a more favourable growing environment. While no relationship has been previously shown for DP3 or DP4 subunits and environment, the results were not unexpected as > 90 % of (1,3;1,4)-β-glucan consists of these two subunits (Lazaridou et al., 2004; Wood and Fulcher, 1983).

4.5.3.4 DP3:DP4 ratio

DP3:DP4 ratio was determined by dividing the peak area of DP3 by the peak area of DP4 and multiplying by a factor of 1.321 which correlates to the molecular weight ratio of (1,3:1,4)-β-gluco-triose and (1,3:1,4)-β-gluco-tetraose (Burton et al., 2011). Ratios were confirmed by comparing calculated values of DP3 divided by DP4 and multiplied by the molecular constant; both calculations correlated with an r value greater than 0.99. The parents showed significantly different values for DP3:DP4 ratio in Vegreville and Castor with H9317004 having a higher ratio (2.90 ± 0.03 and 2.94 ± 0.00, respectively) compared to Merit (2.89 ± 0.01 and 2.92 ± 0.01, respectively). DP3:DP4 ratios were higher in Castor (2.89 ± 0.06) compared to Vegreville (2.87 ± 0.07). The ranges between the environments were similar with 2.69 ± 0.01 − 3.05 ± 0.00 in Vegreville and 2.76 ± 0.03 - 3.04 ± 0.00 in Castor. The Vegreville environments show normal distribution while the Castor environment is skewed towards higher values (Figure 4.3).

The higher precipitation environment of Castor was associated with a higher DP3:DP4 ratio which is similar to finding in oats (Doehlert and Simsek, 2012). In oats it was found that moist environments were associated with a higher frequency of DP3 subunits. *In vitro* studies suggest the (1,3;1,4)-β-glucan synthetic machinery is located within Golgi membranes and utilizes uridine diphosphoglucose (UDP-Glc) from the cytosol (Gibeaut and Carpita, 1994). *In vitro* synthesis of (1,3;1,4)-β-glucan from maize coleoptiles requires Golgi vesicles, UDP-Glc, and either Mn^{2+} or Mg^{2+} as cofactors (Gibeaut and Carpita, 1993; Becker et al., 1995). Analysis of *in vitro* synthesised products has led to a model suggesting three sites on the (1,3;1,4)-β-glucan complex synthesize three β-(1,4) linkages at a time, followed by a β-(1,3) linkage in optimum conditions (Buckeridge et al., 1999). If substrate concentration is below a certain threshold and unable to fill the three spaces available, the complex will lengthen the β-(1,4) series and skip the β-(1,3) linkage creating an irregular pattern of β-(1,3) and β-(1,4) linkages. This proposed mechanisms for (1,3;1,4)-β-glucan synthesis inherently suggests environment would play a role in the fine structure of (1,3;1,4)-β-glucan (Buckeridge et al., 1999). Findings in this study may support this model. The results indicate that wet, more favorable growing conditions in Castor could be associated with higher UDP-Glc concentrations in the cell leading to a more favourable environment for DP3 production. Conversely the frequency of DP4

subunits may be increased in dry environments due to reduced availability of UDP-Glc in the cell.

4.5.4 Correlations

Pearson correlation was used to analyse the relationship between DP3, DP4, DP3+DP4, (1,3;1,4)-β-glucan concentrations and DP3:DP4 ratio averaged across all environments. The DP3, DP4 and DP3+DP4 correlated significantly with (1,3;1,4)-β-glucan concentrations determined colormetrically (r= 0.658, 0.682 and 0.670; $p < 0.001$ respectively) (Table 4.2). Correlations suggest that as total (1,3;1,4)-β-glucan increases; neither the DP3 nor DP4 subunits are preferentially increased. The contribution of higher MW subunits (DP5+) has been shown to account for up to approximately nine percent of the content of (1,3;1,4)-β-glucan (Lazaridou et al., 2004). Differences in methods used to determine (1,3;1,4)-β-glucan vs individual subunits and the exclusion of higher molecular weight subunits may explain differences in values between grain (1,3;1,4)-β-glucan concentrations and DP3+DP4 concentrations. Calculated DP3+DP4 values determined by the HPAEC-PAD method showed values that were slightly higher than expected by colormetric evaluation. Both the (1,3:1,4)-β-gluco-triose and (1,3:1,4)-β-gluco-tetraose standards have a purity of > 95 % which may contribute to a slight overestimation of DP3 and DP4 compared to the colormetric evaluation.

DP3 and DP4 concentrations are positively correlated with DP3+DP4 concentrations (r = 0.996, 0.985 respectively) suggest that the subunits are not preferentially produced in higher or lower (1,3;1,4)-β-glucan lines in this population. These correlations are the same when analysed on an environmental or biological replicate basis. DP3 and DP4 are highly correlated to each other (r = 0.967); also suggesting that subunit production is not competitive (Table 4.2). DP4 concentrations and DP3:DP4 ratio are negatively correlated (r = -0.147; p = 0.164). When analysed on an environment basis, Castor had significant negative relationships between DP4 concentrations and DP3:DP4 ratio (r = -0.25, p = 0.02) while Vegreville showed no significant relationship (r = -0.08, p = 0.49). This may suggest the efficiency of DP4 subunit production affects (1,3;1,4)-β-glucan fine structure in higher rainfall environments. No significant correlation is observed for DP3 concentrations and DP3:DP4 ratio when analysed across environments (0.09, p = 0.40). These results suggest that (1,3;1,4)-β-glucan fine structure is

Table 4.2. Correlations between cellotriose, cellotetraose, cumulative cellotriose and cellotetraose, (1,3;1,4)-β-glucan and DP3:DP4 ratio in 91 barley RIL barley.

	DP3	DP4	DP3+DP4	Ratio
DP4	0.967***	-		
DP3+DP4	0.996***	0.985***	-	
Ratio	0.089^{ns}	-0.147^{ns}	0.014 ^{ns}	-
(1,3;1,4)-β-glucan	0.658***	0.682***	0.670***	-0.050^{ns}

*** $p < 0.001$, **$p < 0.01$ * $p < 0.05$, ^{ns} $p > 0.05$

affected by environment by a decrease in the frequency of DP4 subunits in higher rainfall environment is in accordance with the model proposed by Buckeridge et al. (1999).

4.5.5 Genotype, environment, GxE and heritability

Genotype and environment show a significant effect for all (1,3;1,4)-β-glucan characteristics analysed. Broad sense heritability (h^2) was calculated for each of the traits studied and they ranged from low to moderate values. DP3, DP4, DP3+4 and (1,3;1,4)-β-glucan concentrations had similar h^2 of 24.1 %, 21.3 %, 23.1 % and 30.9 % respectively. DP3:DP4 ratio had the highest heritability score at 43.1 %. All values showed a significant effect of environment. Only DP3:DP4 showed a significant interaction between genotype and environment (Table 4.3). Analysis of variance showed environment explained 68.5 % of the variation in DP3 concentration, 73.3 % of the variation in DP4 concentration, 69.6 % of the variation in DP3+DP4 content, 49.1 % of the variation in (1,3;1,4)-β-glucan and 46.6 % of the variation in DP3:DP4 ratio. Analysis showed genotype is significant for DP3, DP4, DP3+DP4 and (1,3;1,4)-β-glucan concentration (Table 4.3). This indicates that there is a low to moderate genetic influence associated with these characteristics. Environment was also a factor with mean values for all characteristics being higher in Castor compared to Vegreville. DP3, DP4 and DP3+DP4 have low heritability scores (~ 0.22) indicating they are unlikely to be a direct target for modification through breeding. The lower h^2 of the DP3, DP4 and DP3+DP4 may reflect the cumulative effect of environment on the total accumulation of (1,3;1,4)-β-glucan in the cell wall as well as the effect of environment on the ratio of DP3:DP4 ratio.

4.5.6 Association of marker CSLF6_4105 by single marker analysis

DP3:DP4 ratio has a higher heritability than all other measured (1,3;1,4)-β-glucan characteristic which suggests a stronger genetic control to maintain the ratio of DP3 to DP4 in barley (Table 4.3). Overexpression of *CslF6* in barley influences (1,3;1,4)-β-glucan fine structure (Burton et al., 2011). It has been previously demonstrated in the Merit / H93174006 population (Cory et al., 2012) the marker CSLF6_4105 can explain 13 % of the variation in grain (1,3;1,4)-β-glucan concentration. Single marker analysis revealed the CSLF6_4105 is associated with (1,3;1,4)-β-glucan concentration in both environments attributing to 16.8 % of the variation in Vegreville and 8.5 % of the variation in Castor (Table 4.4). DP3:DP4 ratio was also associated

Table 4.3. General linear model results showing the significance of genetics, environment, GxE interaction and broad sense heritability of four measured traits.

Trait	Genotype (G)	Environment (E)	GxE	Replication	Heritability(h^2)
DP3	0.294***	1.170**	0.137 ^{ns}	6.90x10^{-3ns}	24.1 %
DP4	6.40x10^{-2}***	0.332**	3.18x10^{-2ns}	7.77x10^{-3ns}	21.3 %
DP3+DP4	0.627***	2.67**	0.298^{ns}	2.14x10^{-2ns}	23.1 %
BG	0.553***	0.873*	0.198^{ns}	1.99x10^{-2ns}	30.9 %
Ratio	1.18x10^{-2}***	9.51x10^{-3}*	3.76x10^{-3}***	5.40x10^{-3ns}	43.1 %

*** $p < 0.001$, * $p < 0.05$, ^{ns} $p > 0.05$

BG - (1,3;1,4)-β-glucan

with the CSLF6_4105 marker (2.90 vs 2.85 p = 0.00) in Vegreville explaining 8.9 % of the variation but was not significantly associated with DP3:DP4 ratio in Castor. CSLF6_4105 allele from Merit is associated with reduced (1,3;1,4)-β-glucan concentration and lower DP3:DP4 ratio whereas the allele from H93042007 is associated with increased (1,3;1,4)-β-glucan concentration and higher DP3:DP4 ratio. Despite this there is no correlation between (1,3;1,4)-β-glucan and DP3:DP4 ratio in the Merit / H93174006 RIL6 population as a whole or when the population is split by the CSLF6_4015 marker.

4.5.7 Association mapping

A panel of 119 diverse two-row spring genotypes were selected to perform association mapping. (1,3;1,4)-β-glucan grain content was normally distributed ranging from 6.41 ± 0.60 % to 2.89 ± 0.51 % with a mean value of 4.46 ± 0.64 %. DP3:DP4 ratio was also normally distributed ranging from 3.28 to 2.43 with a mean value of 2.83 ± 0.17 (Figure 4.4). Similar to the RIL population, there was no correlation between (1,3;1,4)-β-glucan and DP3:DP4 ratio (0.097, p = 0.296). To further explore the relationship between DP3:DP4 ratio in barley the panel of 119 two-row spring genotypes was subject to association mapping using 24 markers. TASSEL 2.0 was used to perform general linear model for marker association (Table 4.5). The (1,3;1,4)-β-glucan grain concentration was associated with markers CSLF6_4105 and Bmac273e. The CSLF6_4105 marker was previously shown to be a functional marker predicting (1,3;1,4)-β-glucan grain content (Cory et al., 2012). DP3:DP4 ratio was associated with two different genomic regions, Bmac273e on 7H and Bmac504 and Bmac211 on 1H. The region on 1H was shown to be a pleiotropic QTL controlling acrospire growth, grain (1,3;1,4)-β-glucan and wort viscosity (Laido et al., 2009). A putative candidate gene identified was identified in a syntenous region in rice, Os05g01020, which is thought interact with Histone Deacetylase Complex (HDAC) and may control chromosome methylation and gene expression (Islamovic et al., 2013).

4.6 Conclusion

The marker CSLF6_4105 is an indicator of (1,3;1,4)-β-glucan but its utility varies due to environment. It seems to be a weak indicator in wet environments. All phenotypes studied are significantly influenced by genetics and environment. The results show that DP3:DP4 ratio does have a GxE interaction and wetter environments have a higher DP3:DP4 ratio indicating support

for the (1,3;1,4)-β-glucan synthesis model proposed by Buckeridge et al. (1999). DP3:DP4 ratio is significantly associated with genotype and can be a target for genetic selection. The CSLF6_4105 marker is associated with DP3:DP4 ratio in Vegreville but not in Castor. CSLF6_4105 has been shown to be a functional marker in a larger association mapping panel and regions on 7H and 1H are associated with DP3:DP4 ratio.
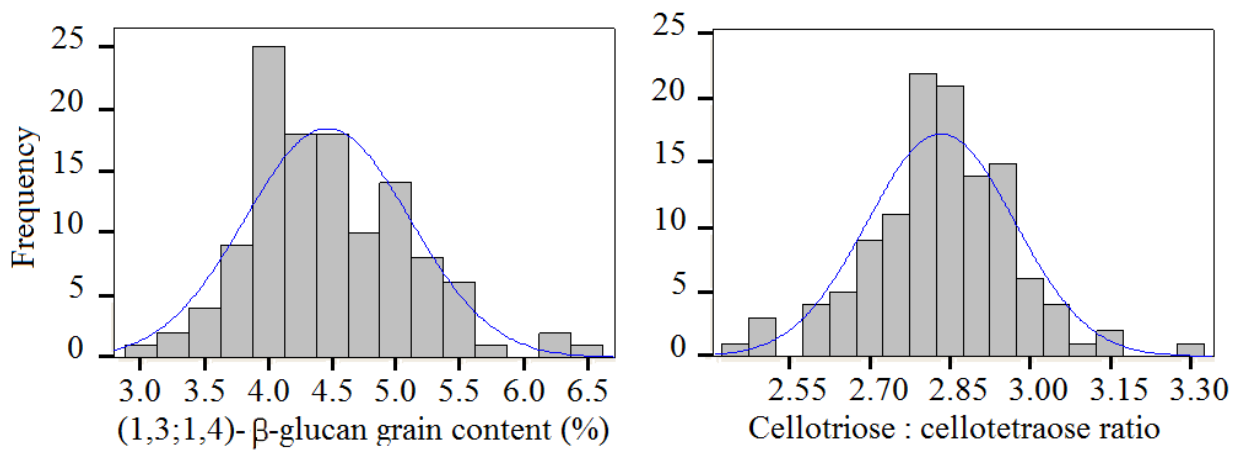
Figure 4.4. Frequency distribution of barley grain (1,3;1,4)-β-glucan concentration and cellotriose to cellotetraose ratio in 119 genotypes used for association mapping.

Table 4.4. Single marker analysis of CSLF6_4105 by one way ANOVA.

Phenotype	p value	Mean CSLF6_A	Mean CSLF6_G	R^2 (%)
Average DP3	0.064	3.57	3.47	3.81
Average DP4	0.160	1.64	1.61	2.21
Average DP3+DP4	0.091	5.22	5.07	3.17
Average BG	0.000	5.33	5.03	15.85
Average Ratio	0.02	2.89	2.87	4.98
Vegreville BG	0.000	5.38	5.07	16.76
Vegreville Ratio	0.002	2.90	2.85	8.88
Castor BG	0.005	5.27	4.98	8.51
Castor Ratio	0.094	2.89	2.88	0.98

BG- (1,3;1,4)-β-glucan

Table 4.5. Candidate marker association by general linear model with grain (1,3;1,4)-β-glucan content and DP3:DP4 ratio in 119 barley genotypes.

Locus	Chrom	df_M	df_E	p BG	R^2 BG	p Ratio	R^2 Ratio	Population
Bmag504	1H	7	111	0.035	0.124	3.31e-4	0.212	Laido et al., 2009
Bmag211	1H	5	113	0.034	0.100	1.32e-4	0.197	Laido et al., 2009
Bmac0350a	1H	7	111	0.781	0.034	0.881	0.026	Panozzo et al., 2007
Bmac0090	1H	2	116	0.839	0.003	0.392	0.016	Li et al., 2011
EBmac0501	1H	2	116	0.008	0.079	0.161	0.031	Panozzo et al., 2007
Bmag0382	1H	3	115	0.204	0.039	0.677	0.013	Laido et al., 2009
scssr004163	1H	6	112	0.388	0.054	0.395	0.053	Molina-Cano et al., 2007
EBmac0684	2H	5	113	0.014	0.116	0.225	0.059	Baum et al., 2003
Ebmac0850	2H	3	115	0.050	0.065	0.545	0.018	Emebiri et al., 2004
Bmag0749	2H	3	115	0.703	0.012	0.425	0.024	Emebiri et al., 2004
Bmac0067	3H	6	112	0.572	0.041	0.052	0.104	None
Bmag0013	3H	9	109	0.209	0.101	0.153	0.111	None
Bmag0023	3H	1	117	0.298	0.009	0.578	0.003	Mather et al., 1997
Bmag0603	3H	6	112	0.223	0.070	0.303	0.061	Mather et al., 1997
Bmac0186	4H	4	114	0.386	0.036	0.055	0.078	Wei et al., 2009
Bmag0337a	5H	2	116	0.704	0.006	0.099	0.039	Molina-Cano et al., 2007
Bmag0751	5H	5	113	0.892	0.015	0.171	0.065	Molina-Cano et al., 2007
HMV14	6H	5	113	0.235	0.058	0.022	0.108	Baum et al., 2003
Bmag341	7H	4	114	0.025	0.093	0.098	0.066	Li et al., 2008
Bmac0273a	7H	4	114	0.072	0.072	0.525	0.027	Li et al., 2008
CSLF6_4105	7H	1	117	2.14E-5	0.144	0.077	0.027	Cory et al., 2012
Bmag516	7H	7	111	0.004	0.166	0.292	0.072	Molina-Cano et al., 2007
Bmac0273e	7H	3	115	6.97E-5	0.173	1.27e-4	0.164	Li et al., 2008
Bmac0156	7H	3	115	0.197	0.040	0.943	0.003	Kim et al., 2011

df_M – Degrees of freedom model, df_E – Degrees of freedom error, R^2 - proportion of variance explained by the model.

CHAPTER 5

ANALYSIS OF ddRAD SEQUENCING BY TWO BIOINFORMATICS PIPELINES REVEAL REGIONS ASSOCIATED WITH BARLEY GRAIN (1,3;1,4)-β-GLUCAN CONCENTRATION

5.1 Study 3*

In this study double digestion Restriction-site Associated DNA (ddRAD) sequencing was performed on 94 barley genotypes and analysed by two bioinformatics pipelines which identified several putative candidate genes associated with grain (1,3;1,4)-β-glucan concentration.

5.2 Abstract

A total of 94 two row spring barley genotypes of diverse origins were genotyped by double digestion Restriction-site Associated DNA (ddRAD) sequencing performed using an Illumina sequencer. Two bioinformatics pipelines were used to discover and call SNPs. The SAMtools bioinformatics pipeline identified 9,062 markers and UNEAK identified 3,060 markers, 2,311 of which were identical between both bioinformatics pipelines. Both sets of markers showed excellent coverage of the genome and could be used to split the ninety-four genotypes into two subgroups which could be defined as Canadian and non-Canadian germplasm. The generated marker data was used for association mapping performed using TASSEL 3.0. Grain (1,3;1,4)-β-glucan content was associated with regions on the 2HS and 5HS telomere by markers generated using both UNEAK and SAMtools. Additional marker associations were identified on 1H, 2H 4H and 7H by individual datasets.

5.3 Introduction

In barley (*Hordeum vulgare* L.), (1,3;1,4)-β-glucan accumulates in the cell wall of the endosperm and to a lesser extent in the aleurone layer (Carpita 1984; Gibeaut and Carpita, 1991; Fincher and Stone, 1986). The glucan polymer has great influence on the nutritional value, functionality and uses of barley. (1,3;1,4)-β-glucan influences the rate of endosperm modification during the malting process, the viscosity of wort during brewing and is a major factor determining malting potential and brewing yield (Brennan and Cleary, 2005). In contrast to malting, high (1,3;1,4)-β-glucan concentrations in grain is desirable for barley used in human nutrition, largely due to its acceptance as a functional, bio-active ingredient (Cui and Wood, 2000). (1,3;1,4)-β-glucan derived from barley can be incorporated into widely consumed foods to reduce their glycemic response, making it an attractive natural food additive (Cavallero et al., 2002). Barley (1,3;1,4)-β-glucan reduce serum cholesterol and modulate the glycemic index in hypercholesterolemic and diabetic patients (El Khoury et al., 2012). Due to the proven health benefits of (1,3;1,4)-β-glucan consumption the US Food and Drug Administration (FDA) have allowed whole grain barley and barley-containing products to carry a claim that they reduce the risk of coronary heart disease (FDA News Release, 2005).

The concentration of (1,3;1,4)-β-glucan in grain is inherited as a quantitative trait, influenced by the additive effects of several genetic factors (Powell et al., 1985), but is also

influenced by environment (Morgan and Riggs, 1981; Perez-Vendrell et al., 1996). Normally, (1,3;1,4)-β-glucan concentrations in barley can range from zero (Kato et al., 1995) to twenty-one percent (Munck et al., 2004) but normally vary between three to six percent (Holtekjølen et al., 2006). The amount of (1,3;1,4)-β-glucan in grain or wort is controlled by several QTL located on all seven barley chromosomes. Thus, QTL can be found on 1H (Han et al., 1995; Igartua et al., 2002; Molina-Cano et al., 2007), 2H (Han et al., 1995; Li et al., 2008), 3H (Li et al., 2008; Islamovic et al., 2013), 4H (Igartua et al., 2002; Wei et al., 2009), 5H (Li et al., 2008; Islamovic et al., 2013; Molina-Cano et al., 2007), 6H (Islamovic et al., 2013) and 7H (Igartua et al., 2002; Li et al., 2008; Molina-Cano et al., 2007; Cory et al., 2012). However these QTLwere identified through bi-parental mapping and are often specific to the population studied. This may be due to the limited genetic diversity and limited number of recombination events in bi-parental mapping populations that have occurred which severely limits the number of QTL that can be detected, making QTL from these crosses of limited value (Flint-Garcia et al., 2003).

Association mapping (AM) is an emerging alternative to bi-parental mapping to locate a wider range of marker associations and therefore genes of interest in the genome. AM is a natural population-based approach tha surveys a large amount of genetic diversity for a trait not limited to two parents in a bi-parental cross. The method is based on trait-marker relationships based on linkage disequilibrium caused by non-random association of alleles at different loci. The statistical association among a set of loci will decay more or less quickly depending on the amount of recombination events that have occurred during meiosis (Dawson, 2000). By studying individuals from a wide variety of genetic backgrounds in which a maximum number of ancient meiotic events have occurred, allow for increased mapping resolution. To date, several AM studies have analyzed (1,3;1,4)-β-glucan concentrations in barley. Markers associated with (1,3;1,4)-β-glucan of grain concentrations have been identified on 1H (Houston et al., 2014; Shu and Rasmussen, 2014), 2H (Houston et al., 2014; Mohammadi et al., 2014), 3H (Houston et al., 2014; Mohammadi et al., 2014), 4H (Mohammadi et al., 2014; Shu and Rasmussen, 2014), 5H (Houston et al., 2014; Mohammadi et al., 2014; Shu and Rasmussen, 2014), 6H (Houston et al., 2014; Mohammadi et al., 2014; Shu and Rasmussen , 2014) and 7H (Houston et al., 2014; Mezaka et al., 2011; Mohammadi et al., 2014; Shu and Rasmussen, 2014). Until recently the marker system used for AM has been limited to sequences derived from expressed sequence tags (EST) through the use of DaRT or Affinity Chip technology.

Next Generation Sequencing (NGS) has allowed the discovery of an increasing number of markers at reduced costs to researchers. This technology has the advantages of being high throughput, not limited to expressed sequences or prior sequence data, and markers discovered may be quickly incorporated into physical maps (Mammadov et al., 2012). SNP discovery using NGS technology offers the advantages of needing no prior sequence knowledge and sequencing efficiency for genotypes scales directly with genetic diversity (Elshire et al., 2011). The major limitation in the utilization of GBS for marker discovery in large, complex genomes such as barley, is to avoid highly-repetitive sections of the genome and to ensure marker veracity by sampling homologous regions repeatedly in each individual (Mammadov et al., 2012). To reduce genomic complexity, a two restriction enzyme approach termed double digestion Restriction-site Associated DNA (ddRAD) sequencing uses one "rare-cutter" and one "common-cutter" enzyme can be used allowing for libraries with a suitable and uniform complexity prior to sequencing (Poland et al., 2012a). Region representation bias favouring fragments closest to the average size selection increases the likelihood of recovering similar genomic regions across all individuals (Poland et al., 2012b).

In this study, we have used a diverse set of 94 two-row spring barley genotypes of various backgrounds to perform AM. Coupling NGS with AM for marker discovery allows us to identify novel markers associated with grain (1,3;1,4)-β-glucan content not restricted to expressed genes or limited by genetic diversity found in bi-parental populations. Two bioinformatics pipelines were used to analyse markers produced using ddRAD sequencing: the *de novo* TASSEL pipeline UNEAK and a reference-based analysis using SAMtools (Li et al., 2009) and bowtie2 (Langmead and Salzberg, 2012). We have compared the marker coverage, substructure identification and marker association and found genomic regions on 1H, 2H, 4H 5H and 7H associated with grain (1,3;1,4)-β-glucan.

5.4 Materials and methods

5.4.1 Plant material

94 two-row barley genotypes were grown in the University of Saskatchewan greenhouse under a night / day temperature range of 19-28 °C with an 18 hour photoperiod with an average photosynthetically active radiation of 385μmol m^{-2} s^{-1}. Each variety was grown in triplicate. Barley lines were grown over 150 days and harvested at maturity.

5.4.2 (1,3;1,4)-β-glucan determination

Grain (10 g) was milled to flour using an Udy-Mil Cyclone sample grinder (UDY Corporation, Fort Collins, CO, USA) equipped with a 0.5 mm sieve. 25 mg samples of flour were used to determine total (1,3;1,4)-β-glucan, by the calcoflour flow injection method (Aastrup and Jørgersen, 1988). Each sample was analysed in triplicate, and pooled averages were used in subsequent statistical analysis.

5.4.3 DNA extraction

Leaves were harvested at the 10 leaf stage, frozen in liquid nitrogen and stored at -80 °C before DNA was extracted and quantified as described (Li et al., 2008). DNA quality was determined using Beckman Coulter DU 800 spectrophotometer (Fullerton, CA, USA) and visually inspected for random shearing on one percent agarose gel. DNA was quantified using Quant-iT™ PicoGreen® dsDNA assay kit (Invitrogen, Molecular Probes, Eugene, OR, USA) using a SpectraMax Gemini XS Flourecence Microplate Reader (Molecular Devices, Sunnyvale, CA, USA). Genomic DNA concentrations were normalized to 10ng / μL and subsequently used for library preparation. Sequencing libraries were prepared by the Université de Laval, Institut de biologie intégrative et des systèmes according to the GBS protocol as per Elshire et al. (2011) except for the use of selective primers. Single-end sequencing was performed on a single lane of an Illumina Genome Analyzer II (at the McGill University-Génome Québec Innovation Center in Montreal, Canada).

5.4.4 Processing of illumina raw sequence read data and SNP calling

Sequence information for the *CslF* genes was obtained from EnsemblPlant *Hordeum vulgare* database (Kersey et al., 2014) corresponding to sequences from MLOC_59289: *HvCslF3*, MLOC_74149: *HvCslF4*, MLOC_57200: *HvCslF6*, MLOC_52689: *HvCslF8*, MLOC_59237: *HvCslF9*, MLOC13463: *HvCslF10*, MLOC_ 19594: *HvCslF11*, and MLOC_7825: *HvCslF12*. Blast search was used to identify as genomic sequence corresponding to *HvCslF13*. No genomic sequence for *HvCslF7* could be identified.

Perl programing language was used to implement a pipeline for the processing of raw Illumina single end reads. Genious R6 (6.1.4) was used to de-multiplex the reads and trim

barcodes. Publically available software tools which were subsequently used to process the single-end sequence reads. 94 FastQ files were generated and subjected to grooming using trimmomatic. Each sequence was groomed using a sliding window of four with a minimum quality rating of 20 (Q20). The minimum length of a read to be accepted was 50 nucleotides. The publically available, unmasked *Hordeum vulgare* L. genome (030312v2) was downloaded from EnsemblPlants (ftp://ftp.ensemblgenomes.org/pub/plants/release-23/fasta/hordeum_vulgare/dna/) on July 1, 2014 (Kersey et al., 2014). The genome was indexed and used as a template for alignment of the trimmed and groomed FastQ files. Bowtie2 was used to align the 94 two-row barley genotypes to the reference genome to generate SAM files. SAMtools was used to convert SAM files to BAM files. The BAM files were then sorted, PCR duplicates were identified and removed and the BAM files were indexed. The indexed BAM files were processed using SAMtools mpileup to generate a variant call file (VCF file). The raw VCF was filtered using VCFtools. Heterozygous calls were masked as missing data and filtered to identify markers with a minimum read depth of six, missing data in genotypes of no more than twenty percent, and a minor allele frequency greater than five percent.

The UNEAK TASSEL GBS pipeline was used to call SNPs in tags prior to alignment on the reference genome 030312v2 (Mayer et al., 2012; Kersey et al., 2014). Initial filtering was used to keep only markers with a minimum read depth of six. Additional filtering was performed using the filtering options in TASSEL 3.0. Markers showing heterozygous calls were masked as missing data, any data with a minor allele frequency less than five percent and markers with more than 20 % missing data were filtered. Using the *de novo* GBS pipeline, 64 bp tags from TASSEL were mapped against the Morex whole-genome shotgun assembly (Mayer et al., 2012) using bowtie2. Only markers mapped to one of the barley chromosomes and meeting the filtering criteria were used for subsequent analysis. Visualisation of the marker coverage was prepared using MapChart 2.2 (Voorrips, 2002).

5.4.5 Kinship

Kinship calculations were performed using SPAGeDi (Spatial Pattern Analysis of Genetic Diversity) software package (Hardy and Vekemans, 2002) using the matrix calculations from Ritland (1996) with 10,000 permutations. Markers with a read depth of six, minimum

missing data of 20 % and a minor allele frequency (MAF) of five percent were used for the UNEAK dataset, whereas a subset of the SAMtools dataset representing an MAF of 10 % was used for calculations as this showed sufficient coverage and represented a similar number of data points to the UNEAK kinship calculations.

5.4.6 Population structure

Two different methodologies and software packages were employed and compared to estimate the number of subgroups in the two-row barley population. For the quantitative assessment of the number of groups in the panel, a Bayesian clustering analysis was performed using the software package STRUCTURE v2.3.4.that uses multi-locus genotypic data to assign individuals to clusters (k) without prior knowledge of their population affinities. The program was run with 3,061 markers from the UNEAK pipeline and a subset of 1,089 SNP markers from the SAMtools pipeline for k-values 1 to 6 (hypothetical number of subgroups), with 100,000 burnin iterations followed by 200,000 MCMC (Markov Chain Monte Carlo) iterations with 5 independent runs for each k. The most probable number of groups was determined by Structure Harvester, implementing the Evanno method to determine the most probable number of clusters (Earl and vonHoldt, 2012). The largest value of an ad hoc statistic ΔK was used as an indicator for the true number of clusters. In a second approach, Principal Coordinate Analysis (PCoA) based on the dissimilarity matrix was performed using GenALEx (Peakall and Smouse, 2012) on all markers in the respective data sets. The number of principal components to include in the linear model was determined by scree plot (Cattell and Vogelmann, 1977).

5.4.7 Association mapping

Association analysis was conducted using Tassel 3.0 standalone software (Bradbury et al., 2007). Analysis was performed using a naive general linear model (GLM), a general linear model with substructure correction using a P-matrix or Q-matrix (GLM+P, GLM+Q), a mixed linear model with kinship alone mode (MLM+K) and a mixed linear model with P- and Q-matrix (MLM+KP, MLM+KQ) using markers from both pipelines. A false discovery rate (FDR; Benjamini and Hochberg, 1995) was used for multiple testing correction of the GWAS results. To assess the impact of population structure control, cumulative distributions of p-values for all models were calculated and compared.

5.4.8 Candidate gene identification

Genes associated with markers were identified using the EnsemblPlant *Hordeum vulgare* L. database release version 23 (Kersey et al., 2014). Marker names correspond to absolute positions within the database. A list of genes was compiled from between marker intervals surrounding significantly associated genes. Information on unnamed genes was collected corresponding to the Gene Ontology database and / or orthologous genes where available.

5.5 Results and discussion

5.5.1 (1,3;1,4)-β-glucan grain content

Analysis of mature grain from 94 barley genotypes grown under greenhouse conditions was analysed by calcoflour flow injection (Aastrup and Jørgersen, 1988) to determine (1,3;1,4)-β-glucan concnetraions. Within the population (1,3;1,4)-β-glucan concnetrations ranged from 2.90 % ± 0.63 % to 6.41 % ± 0.74 % with an average of 4.56 % ± 0.60 %. The frequency distribution was observed to follow a normal distribution when plotted as a histogram (Figure 5.1). When split by origin, the 53 Canadian varieties had an average (1,3;1,4)-β-glucan content of 4.62 % ± 0.60 % and the 41 non-Canadian varieties had a slightly lower mean (1,3;1,4)-β-glucan content of 4.47 % ± 0.71 %. However, analysis using t-test showed the averages were not significantly different.

5.5.2 Genetic markers

Sequencing of the 94 barley genotypes produced 193 million raw reads of which 91 % could be separated by barcode; this is within the expected range reported (Mascher et al., 2013). After processing and filtering for a minimum read depth of six, the UNEAK pipeline identified 52,511 markers across all genotypes. 36,748 (69.22 %) of the markers could be positioned on the barley reference genome. SAMtools pipeline identified 19,918 markers, all of which are aligned
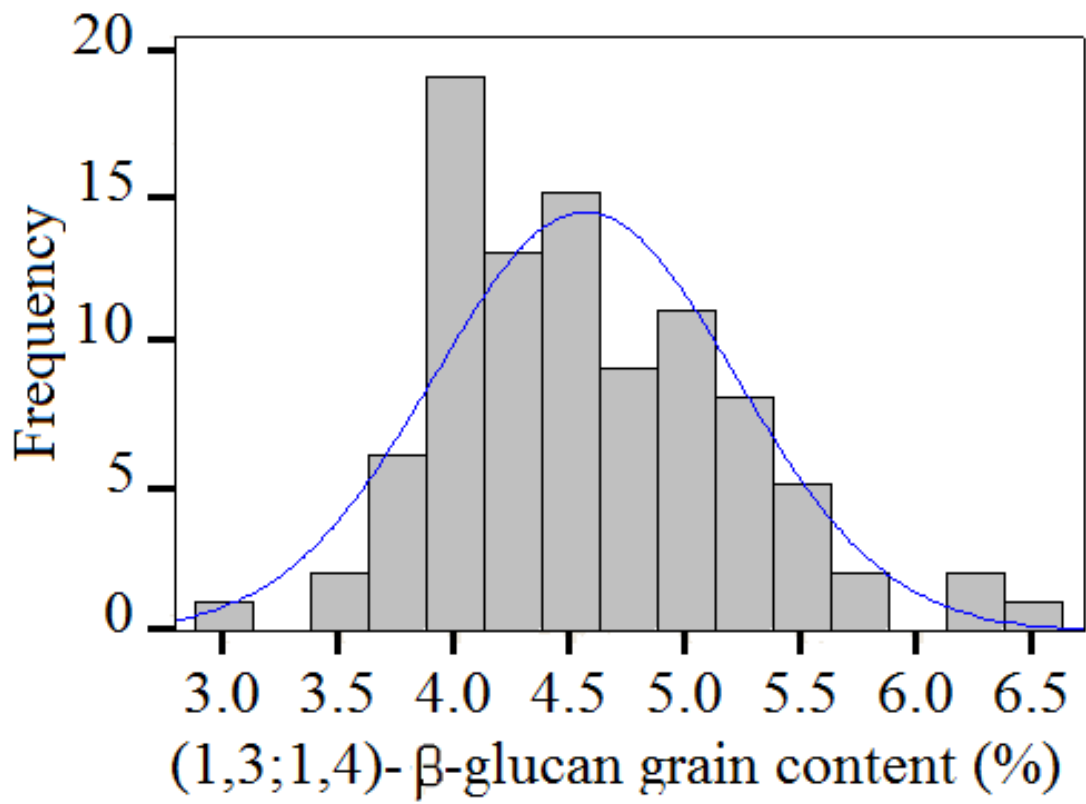
Figure 5.1. Frequency distribution of (1,3;1,4)-β-glucan concentration in 94 two-row barley genotypes.

to the reference genome. SNPs used for subsequent analysis were filtered for the following criteria: 1) heterozygous SNPs were converted to missing data, 2) a minimal read depth of six for each SNP, 3) minimal allele frequency above five percent for each marker and 4) no more than 20 % missing data for each marker. After filtering, 9,062 markers remained from the SAMtools pipeline, and 3,192 remained from UNEAK pipeline. Map positions of the markers generated from both pipelines were aligned in Microsoft Excel and markers were determined to be in common based on the following criteria: i) similar map position, ii) identical nucleotide calls and iii) alleles matched for greater than 95 % of all calls. 2,311 markers were common between the two bioinformatics programs while 749 were unique to the UNEAK pipeline and 6,751 markers were unique to the SAMtools pipeline (Figure 5.2).

SAMtools bioinformatics pipeline identified over twice the number of high quality markers compared to UNEAK. This may be partially due to the initial quality filtering and trimming. Within the UNEAK bioinformatics, the initial filtering and trimming reduces the usable sequence to 64 bp including the invariable cut site. This effectively reduces the useable sequence to 60 bp as compared to a possible 92 bp available using the SAMtools pipeline. Quality control parameters would also allow for more sequence information to be used in the SAMtools pipeline. SAMtools allows for sequence with a quality score of 20 in a sliding window of 4 and any sequence over 50 bp. UNEAK removes any sequence with an N within the 64 base pair trimmed read. These two differences could mean that there might be a reduced set of initial reads kept by UNEAK compared to SAMtools. More unique markers are generated with UNEAK at a minimum read depth of six, therefore this is unlikely to be the cause on its own. UNEAK may filter out more reads from individual lines therefore accumulating more reads that fail to meet a minimum read depth of 6 and an overall missing marker value under 20 %. Another contributing factor to the difference in markers identified could be the initial alignment to the reference genome in the SAMtools pipeline. This initial alignment would act as a quality control filter, removing DNA sequences which are not present in the barley variety Morex. UNEAK may generate more unique reads in individual genotypes which are found in only a small subset of varieties but would not pass the missing data filtering criteria. Initially, UNEAK identified 52,000 markers when filtered for a minimum read depth of six, but after filtering for a minimum missing data of less than 20 % only 3,662 loci remain. After filtering for a minimum read depth,
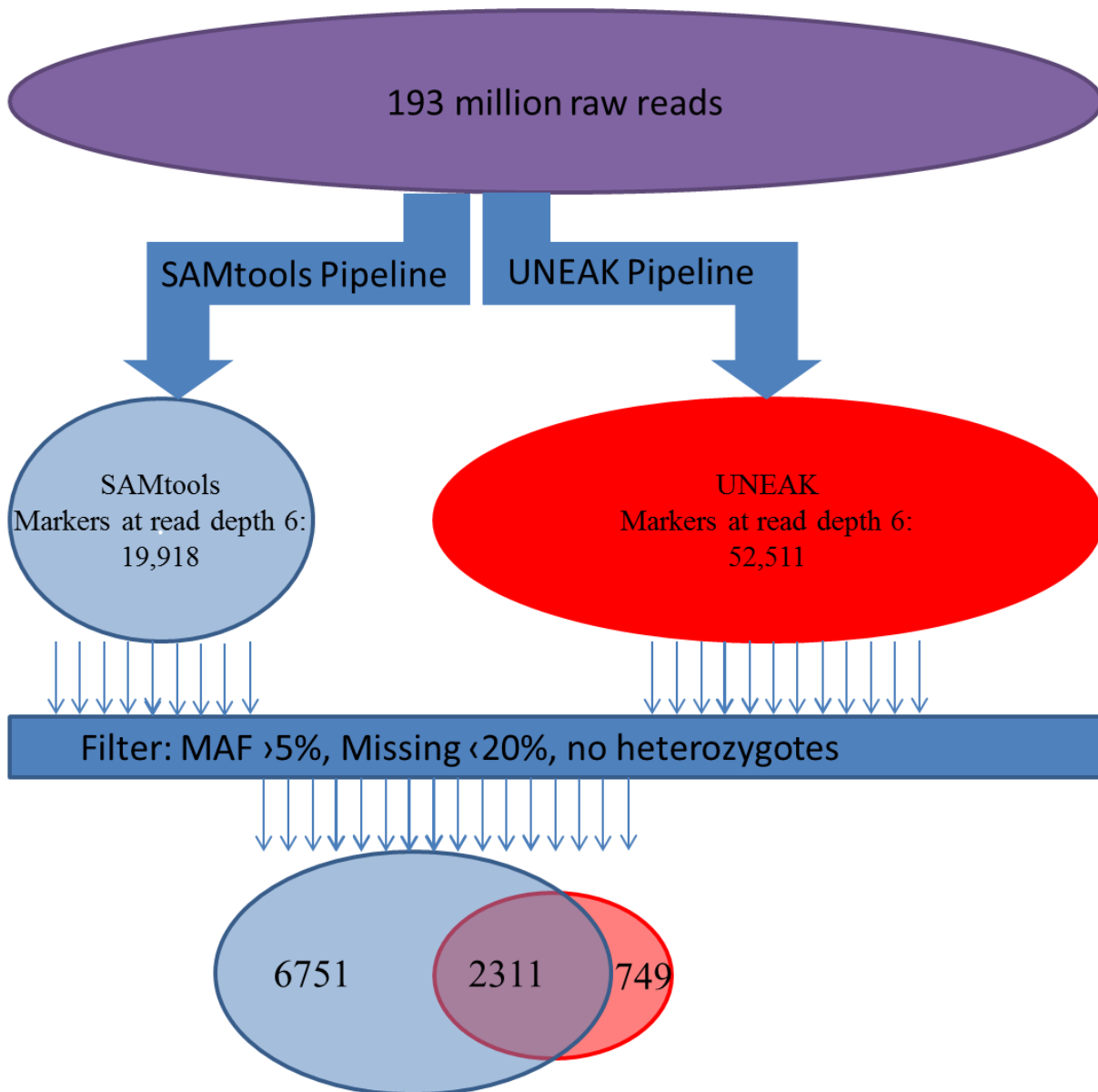
Figure 5.2. A schematic illustration to identify loci using two bioinformatics pipelines. Processing of raw sequences was done using both UNEAK and SAMtools bioinformatics pipelines. At the initial filtering step 52,511 loci were identified in 94 barley genotypes using UNEAK, 19,918 were identified using SAMtools bioinformatics. After filtering for minor allele frequency (MAF) of five percent, no heterozygous calls and no more than 20 % missing data 3060 loci were identified using UNEAK, 9062 using SAMtools with 2311 loci identified by both bioinformatics pipelines.

MAF greater than five percent and alignment on a reference genome, both pipelines identified 2,311 identical loci. This is similar to results found by Mascher et al. (2013), who compared data generated from a RIL6 population 'Morex' and 'Barke' on three different sequencing platforms and analysed using UNEAK and SAMtools pipelines. The study showed the largest difference in identified markers came from the bioinformatics pipeline used.

Mapchart 2.2 was used to visualize the position of all markers in relation to the reference genome. The reference genome consists of seven barley chromosomes generated from 138,000 whole genome shotgun (WGS) contigs labeled 1 - 7 with sequencing information from an additional 355,000 WGS contigs that could be assigned to a chromosome arm. As seen in figures 5.3 and 5.4, markers are arranged in 7 chromosomes and unanchored 13 chromosomes regions. Although higher marker density is seen in the SAMtools dataset compared to UNEAK dataset both show similar distribution throughout the genome. Markers are denser towards the telomeres and less coverage is found toward the centromere in both datasets (Figure 5.3 and Figure 5.4).

5.5.3 Genetic marker distribution

Markers are found throughout the genome but they are unevenly distributed (Figure 5.3 and 5.4). The larger pseudo-molecules show increased marker density near the telomeres and reduced marker density near the centromeres. The marker distribution is likely caused by the arrangement of the reference genome. The reference genome consists of seven barley chromosomes generated from 138,000 whole-genome shotgun (WGS) contigs labeled 1-7 with sequencing information from an additional 355,000 WGS contigs that could be assigned to a chromosome arm but not to a position on the larger pseudo-molecule (Mayer et al., 2012). Markers assigned to the larger pseudo-molecules near the centromere fall on "islands" of sequence surrounded by large stretches of non-sequence (Ns). The smaller pseudo-molecules may fit into these areas. For example, there is little coverage near the centromere of 7H but good coverage on the smaller 7HS pseudo molecule where *Amy2* is located. From consensus sequences *Amy2* would fall within the regions of low marker coverage on the larger 7H chromosome.
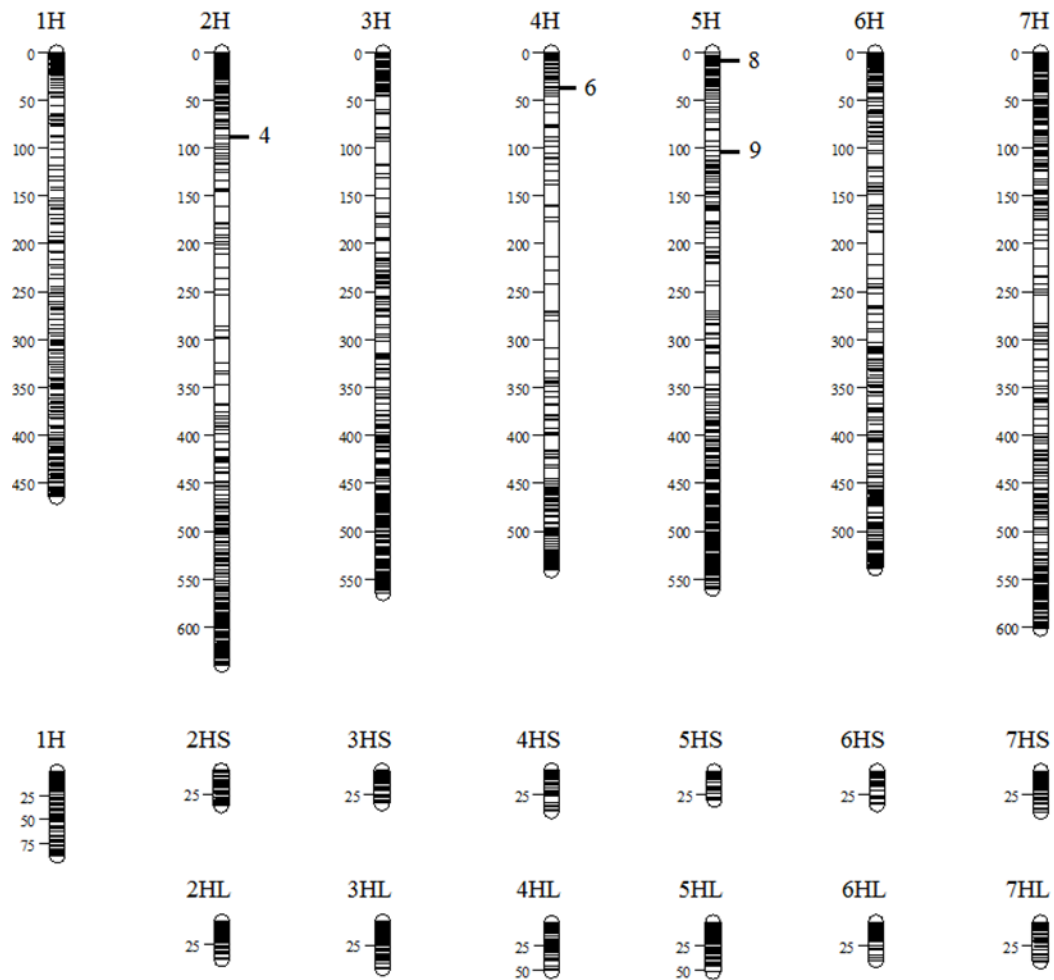
Figure 5.3. Markers generated from the SAMtools bioinformatics pipeline arranged on the *Hordeum vulgare* L. reference genome 030312v2. Indicators on the left represent distance in megabase pairs, markers associate with grain (1,3;1,4)-β-glucan content are represented on the right by numbers that correspond to Table 5.1.
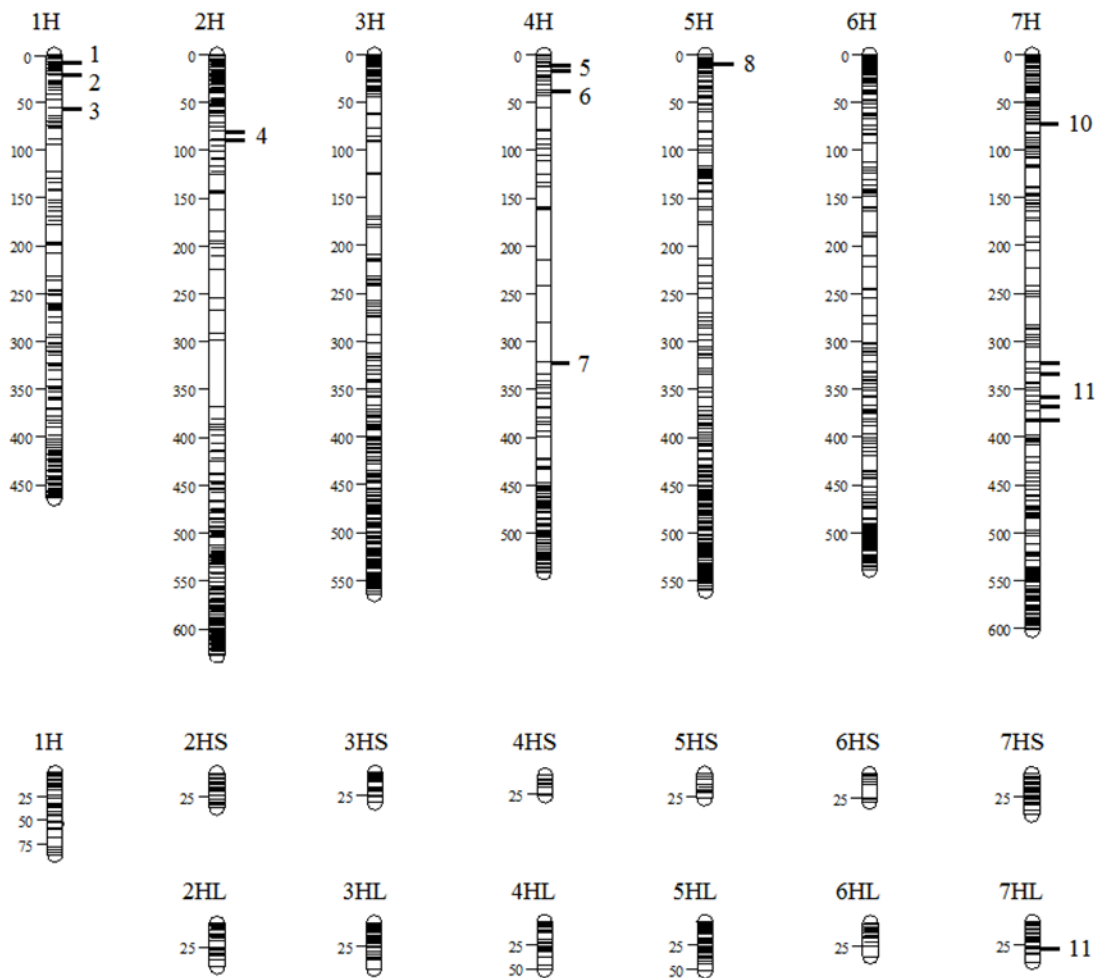
Figure 5.4. Markers generated from the UNEAK bioinformatics pipeline arranged on the *Hordeum vulgare* L. reference genome 030312v2. Indicators on the left represent distance in megabase pairs, markers associate with grain (1,3;1,4)-β-glucan content are represented on the right by numbers that correspond to Table 5.1.

Table 5.1. Marker association by statistical model and dataset.

Figure legend	Model	Dataset	Chrom	Position	F	-log(p)	Df	r2(%)
1	GLM+P	UNEAK	1H	4396096	14.87	3.66	87	12.8
2	MLM+Q	UNEAK	1H	19876373	15.21	3.73	91	16.7
3	MLM+Q	UNEAK	1H	50285585	15.66	3.82	92	17.2
4	MLM+Q	UNEAK	2H	79055395	14.6	3.56	77	19.3
4	MLM+Q	SAMtools	2H	80208254	17.31	4.1	82	19.9
4	MLM+Q	UNEAK	2H	80248225	14.75	3.63	91	16.5
4	MLM+Q	SAMtools	2H	80248284	18.15	4.3	93	19.5
5	GLM+P	UNEAK	4H	8111198	15.28	3.73	86	12.5
5	MLM+Q	UNEAK	4H	8111198	14.47	3.58	89	17.4
5	MLM+Q	UNEAK	4H	12157561	14.93	3.65	84	18.6
5	GLM+P	UNEAK	4H	13266457	14.79	3.65	89	12.4
6	MLM+Q	SAMtools	4H	31395904	17.26	4.31	94	18.1
6	MLM+Q	UNEAK	4H	36996981	13.62	3.41	89	14
7	GLM+P	UNEAK	4H	321179812	14.88	3.66	87	12.8
7	MLM+Q	UNEAK	4H	321179812	17.9	4.24	90	19.6
8	MLM+Q	SAMtools	5H	7406056	17.7	4.21	92	18.7
8	MLM+P	SAMtools	5H	7491839	20.1	4.67	93	17.5
8	MLM+Q	SAMtools	5H	7491839	21.78	3.97	93	22.9
8	GLM+P	UNEAK	5H	7557217	20.73	4.73	80	18
8	MLM+Q	UNEAK	5H	7557217	20.68	8.73	84	26
9	MLM+Q	SAMtools	5H	10176328	18.78	4.37	82	22
10	GLM+P	UNEAK	7H	68292885	18.88	4.42	87	15.7
11	MLM+Q	UNEAK	7H	327853389	12.66	3.21	89	14
11	MLM+Q	UNEAK	7H	332149988	12.66	3.21	89	14
11	MLM+Q	UNEAK	7H	343695470	12.66	3.21	89	14
11	MLM+Q	UNEAK	7H	362218969	12.66	3.21	89	14
11	MLM+Q	UNEAK	7H	383576564	12.66	3.21	89	14
11	MLM+Q	UNEAK	7HL	26607213	12.66	3.21	87	14

Chrom- chromosome, Position- physical position on the physical Morex genome map,
Df- marker degrees of freedom, r2- coefficient of determination converted to percentage.

5.5.4 Population structure

Population substructure was determined using two different methods; STRUCTURE v 2.3.4 software, and Principal Coordinate Analysis (PCoA) using GenAlEx 6.5 software. Similar results for substructure assignment were found using both STRUCTURE v 2.3.4 and PCoA in both datasets. STRUCTURE v 2.3.4 identified two groups ($k = 2$) which roughly split the genotypes into groups from Canadian breeding programs and genotypes from other sources (Figure 5.5). This was similar to the stratification found by PCoA (Figure 5.6). PCoA identified six principal components. Scree plot analysis identified the first three principal components as non-trivial. Collectively these three components accounted for 67.7 % and 67.4 % of the genetic variation using the SAMtools or UNEAK datasets, respectively. Scatter plots of the first two principal components show two distinct groups that separate Canadian varieties from more diverse varieties (Figure 5.6). The SAMtools data set shows better separation of the two groups, which may reflect the larger number of markers used for the calculations.

In genome wide association studies, population structure has been considered an important cause of spurious associations and an explanation of failure to replicate significant predictions, making statistical methods accounting for population structure essential to validate standard association tests (Balding, 2006). Factors including geographic localization, breeding patterns and selective breeding based on agronomically significant traits during crop improvement may lead to strong population structure and familial relatedness within plant populations in association mapping studies (Atwell et al., 2010). In the present study, population structure was demonstrated primarily due to division by geographical origin. Both PCoA and the Bayesian cluster analysis by STRUCTURE v 2.3.4 have identified one main subdivision of this population based on geographical origin. Stratification of samples divides Canadian genotypes from those from other regions of the world. This may reflect the specific selection pressure due to preferences among Canadian breeders or the specific environmental pressures present in Canada. However, two notable exceptions to the groupings are Canadian Thorpe and M98135002. A closer investigation of the pedigree of M98135002 shows that this variety is derived from a Japanese (Kanto Nijo) and Mexican (Arupo / K8755 // Mora) barley genotype. Canadian Thorpe was one of the first barley varieties to be registered in Canada and is originally a selection from the UK Thorpe variety. Three Ethiopian landraces, CIho4961, CI3124 and CI9819, were a distinct subgroup in both datasets indicating their common country of origin. The
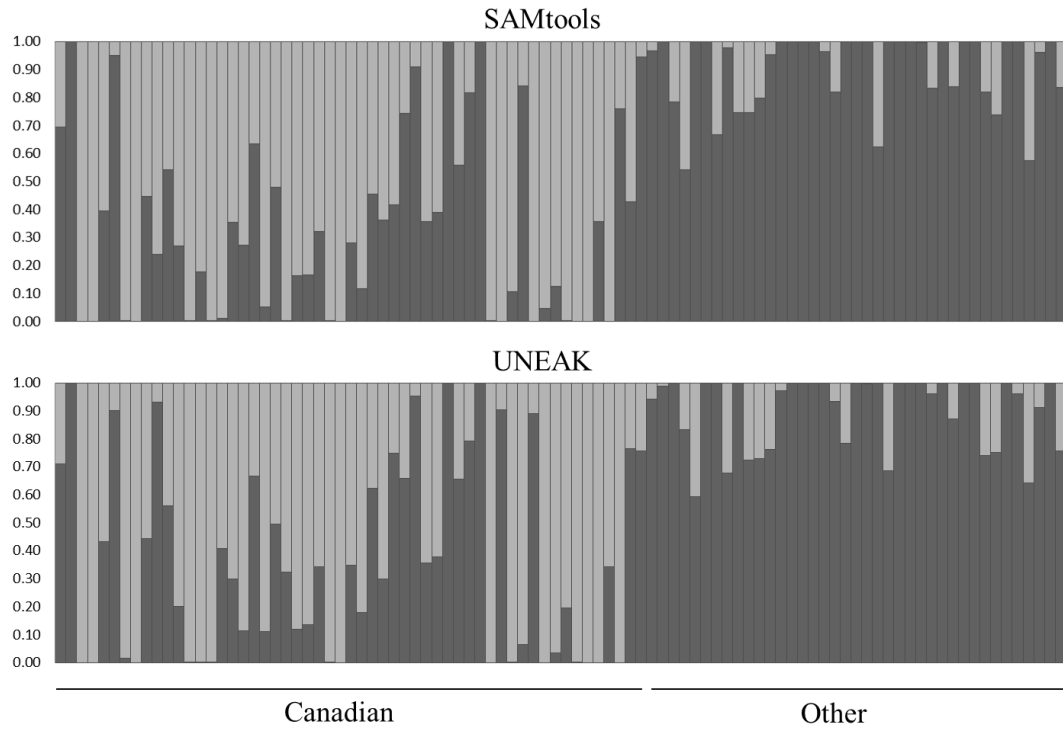
Figure 5.5. Subpopulation assignment of 94 two-row barley genotypes. Analysis using the linkage ancestry model with correlated allele frequencies in STRUCTURE 2.3.4. Each bar represents an individual, the proportion of genomic membership is in each subpopulation is represented by different degrees of shading within the bar.
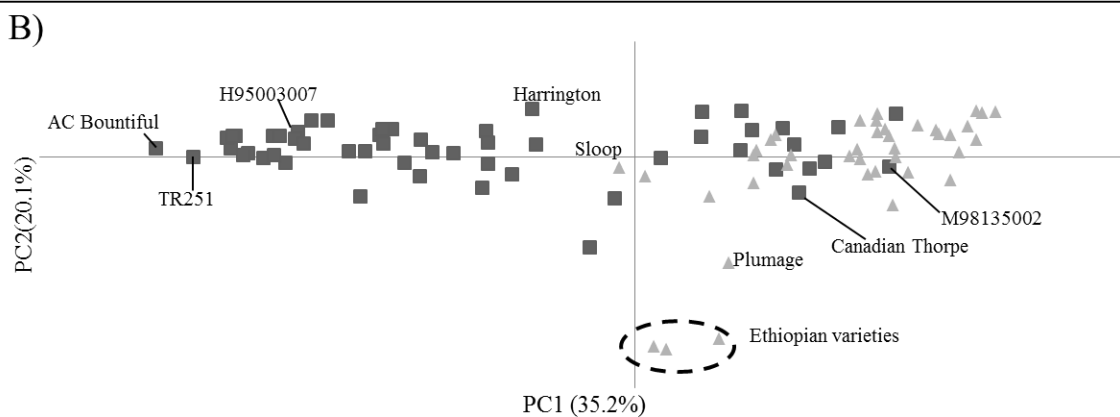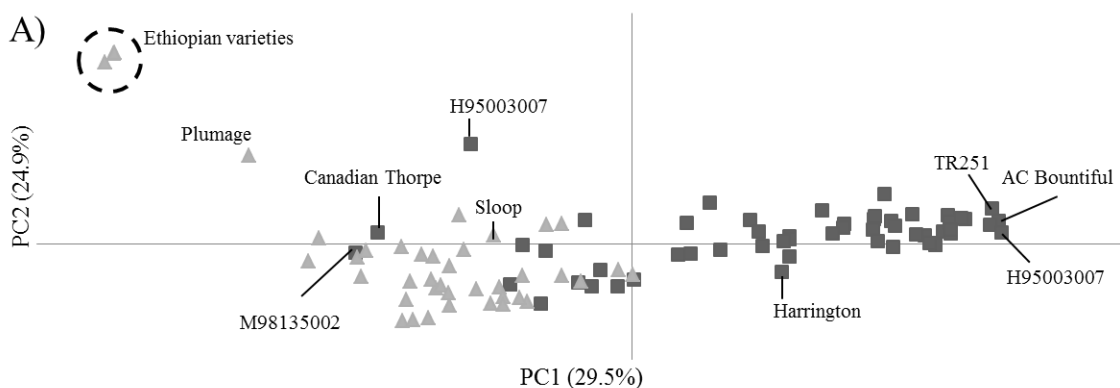
Figure 5.6. Scatter plot of 94 two row barley genotypes based on the first two principal coordinate principal coordinate analysis (PCoA) axes. A) SAMtools dataset B) UNEAK dataset. The percentage of variance explained by each axis is indicated. Canadian varieties are indicated by squares and other varieties are represented by triangles.

The population stratification found in this study is similar to results of other studies. Tondelli et al. (2013) demonstrated population substructure in a two row European barley panel based on old vs. modern barley varieties. Comarand et al. (2009) showed a strong influence of row number and geographical location in a study of 192 barley accessions from the Mediterranean basin, while Beattie et al. (2010), working with 91 two-row varieties, identified stratification due to breeding program.

5.5.5 Marker associations

To decrease the likelihood of false associations and still maintain a strong prediction power, an optimization of the model is essential. Naïve, Q, PCoA, Kinship (K), Q+Kinship (Q+K), and PCoA+Kinship (PCoA+K) were performed and interpreted based on the fitting plot (observed against expected probability, PP-plot; Figure 5.7). Correction for substructure using Q-matrix was the least effective in both data sets. Incorporation of a P-matrix decreased the amount of skew more significantly than the Q-matrix, but kinship and P-matrix reduced the skew the most. The UNEAK data set, the PK model skewed results below expected limits. Quintile-Quintile (QQ) plot showed the actual association results below the expected, indicating over-fitting.

Using the method of Benjamini and Hochberg (1995), a false discovery rate (FDR) was used to adjust for multiple testing when determining significant marker-trait associations. A FDR of 0.10 was selected instead of the normal 0.05 to account for the fact that each marker does not constitute an independent test as many of the markers are found within linkage blocks in both datasets. With the SAMtools and UNEAK data sets, the P+K model showed only one and no markers associated with (1,3;1,4)-β-glucan concentration, respectively. Investigation of both the PP-plot and the QQ-plot showed the UNEAK P+K model was over-fit, with all points falling below the expected values in both graphs. The UNEAK dataset analysed with P-matrix without kinship was the least skewed but not over-fit. The P-matrix model using the UNEAK dataset identified six associated markers in five separate genomic regions. Both data sets identified the same chromosomal region associated with (1,3;1,4)-β-glucan concentration in barley. Both datasets identified markers on the telomere of chromosome 5H. These markers were unique to their respective datasets but fell within 0.65 Mbp of each other according to positioning on the reference genome. MLM+Q showed similar regions associated with grain (1-3,1-4)-β-glucan
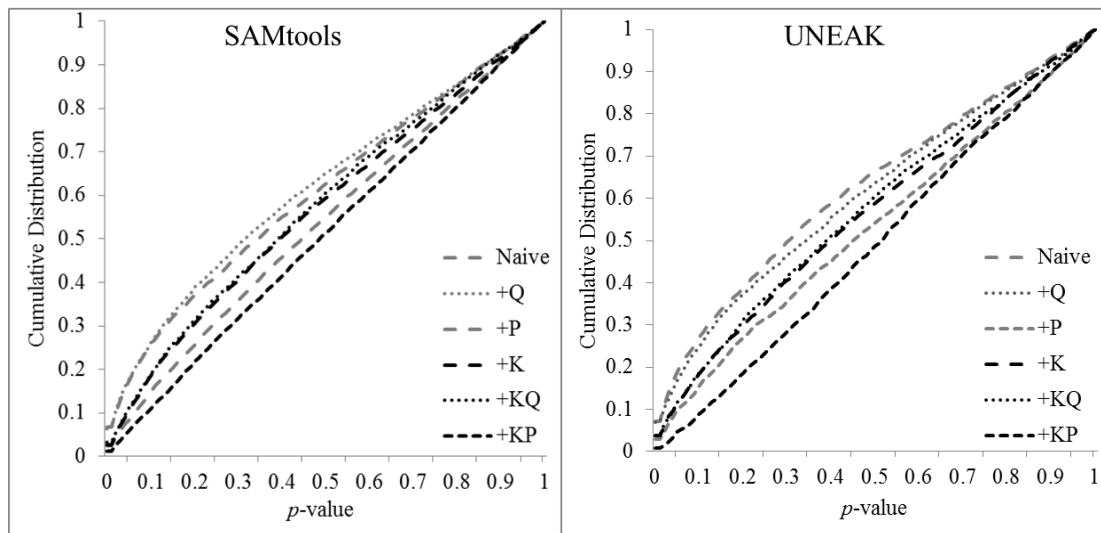
Figure 5.7. Cumulative distribution of *P* values for six linear mixed-models differing in population structure control method. A naïve model with no substructure control is compared to statistical models correcting for kinship (K) and structure determined by principal coordinate analysis (P) or STRUCTURE (Q).

concentration in both datasets with regions in 2H, 4H, and 5H. MLM+Q using the UNEAK pipeline dataset also identified regions on 1H and 7H. These regions were identified by the SAMtools dataset but failed to pass the FDR post hoc test due to the larger number of null hypotheses. A list of genes identified within near markers on 1H, 2H, 4H, 5H, and 7H can be seen in the appendix.

Both data sets identified the same chromosomal region associated with (1,3;1,4)-β-glucan content in barley near the telomere of 5H. Within the SAMtools marker dataset, marker 5_7491839 was identified as significant in all statistical models except for GLM+P which showed a marker, 5_10176328, as significantly associated. Within the UNEAK marker dataset, an independent marker, 5_7557217, was seen as significant. These markers are in a similar region on 5H as that described by Houston et al. (2014), Shu and Rasmussen (2014) and Islamovic et al. (2013). Marker 5_7491839 explained 17.5 % of the variation in the population when using the SAMtools dataset and marker 5_7557217 explained 18.0 % of the variability when using the UNEAK marker dataset. These markers are unique to their respective datasets but fall within 0.65Mbp of each other according to positioning on the reference genome. Within this region only one gene can be identified, MLOC_21074; this is an uncharacterized gene which encodes a 64 amino acid protein. A second likely candidate is the uncharacterized gene MLOC_2781. This gene is 75 % similar to the rice gene Os12g0630500, an alpha-amylase / trypsin inhibitor. This gene may be of particular interest as this region of 5H has been associated with both (1,3;1,4)-β-glucan and amylose (Shu and Rasmussen, 2014).

The UNEAK GLM+P identified five other significant markers. The marker 1_4396096 on 1H is found in a similar region to that described by Laido et al. (2009) and Shu and Rasmussen (2014). Laido et al. (2009) commented that the region was a pleiotropic QTL controlling acrospires growth, grain (1,3;1,4)-β-glucan and wort viscosity. Islamovic et al. (2013) identified a gene in a syntenous region in rice, Os05g01020, putatively involved in Histone Deacetylase Complex (HDAC) associated with amylose content in the Falcon / Azul population. The marker identified in this study is located near a cluster of genes. A promising candidate is the uncharacterized gene MLOC_76007. According to the Gene Ontology database this gene may have protein dimerization abilities and may participate in methyl transfer. Another likely candidate is a putative C2H2 zinc finger protein, MLOC_37432. This class of protein has been shown to be transcription factors controlling gene expression in all eukaryotic organisms (Wolfe

et al., 2000). Three markers on 4H were found associated with grain (1,3;1,4)-β-glucan concentration. Two markers (4_8111198 and 4_13266457) near the end of 4HS encompass a large cluster of genes including MLOC_3791, a putative acid phosphatase protein coding gene. Phosphorylation has been shown to regulate cellulose β-glucan synthesis in Arabidopsis (Chen et al., 2010; Taylor, 2007), pea (Ray, 1973), tobacco (Kaida et al., 2009) and corn (Paliyath and Poovaiah, 1988).

One marker (4_321179812) on 4HL was found in a similar region to a QTL found by Wei et al. (2009). The nearest gene is MLOC_21017, a hypothetical gene encoding a 912 amino acid similar to at5g01310, an APRATAXIN-LIKE, APTX a basic helix-loop-helix transcriptional factor that is involved in regulation of xylan synthesis (TAIR: The Arabidopsis Information Resource). The final marker is 7_68292885 near MLOC_76756, a hypothetical protein of 110 amino acids orthologous to the Rhodanese / Cell cycle control phosphatase superfamily in Arabidopsis which are involved in cell wall biosynthesis based on reviewed computational evidence listed in TAIR. A second gene in this region is MLOC_73315 an uncharacterised gene displaying a UDP-glucuronosyl / UDP-glucosyltransferase domain. UDP-Glc is the main building block for cellulose and (1,3;1,4)-β-glucan and the involvement of a UDP glucosyltransferase has been postulated as a subunit in the (1,3;1,4)-β-glucan synthase complex (Urbanowicz et al., 2004).

MLM+Q model using the SAMtools dataset showed three regions associated with grain (1,3;1,4)-β-glucan content (Figure 5.3). The telomere of 5HS was significantly associated with (1,3;1,4)-β-glucan content at marker 5_7491839 accounting for 22.9 % of the variation. 4HS also had a marker (4_31395904) with significant association which falls within a cluster of three genes. Of these three genes MLOC_53722, a putative serine / threonine-protein kinase is the most likely to affect cell wall biosynthesis. Two markers on 2H (2_80208254 and 2_80248284) were found to be associated and accounted for 19.9 % and 19.4 % of the variation respectively. These markers are found in the middle of 2HS in a region previously associated with (1,3;1,4)-β-glucan content by Baum et al. (2003). MLOC_56623 is a potential candidate gene found in this region. This gene encodes a putative protein containing a target SNARE coiled-coil domain which is involved in vesicle transport. *In vitro* evidence suggests that the (1,3;1,4)-β-glucan synthase complex assembles the (1,3;1,4)-β-glucan in the Golgi vesicles where it is then shuttled to the plasma membrane to be released into the cell wall (Urbanowicz et al., 2004).

The MLM+Q model using the UNEAK dataset identified 15 associated markers in five different regions (Figure 5.4). As with all other models the region on 5H was represented by marker 5_7557217; accounting for 25.2 % of the variation. Regions on 4H were found to be associated with the same markers found to be significant by the GLM+P model. Marker 2_79055395 and 2_80248225 were found associated in a similar region to those found in the SAMtools dataset. A region near the centromere on 7H was also identified as significantly associated with (1,3;1,4)-β-glucan content. This region is the most commonly reported genomic area associated with (1,3;1,4)-β-glucan content. It contains genes such as *Amy2* and *CslF6*. Six different markers were found to be significant in this region. Each marker had a similar R^2 explaining 14 % of the variation in (1,3;1,4)-β-glucan content results are summarised in Table 5.1.

5.5.6 Sequence information from *CslF* gene family

The cellulose synthase F family of genes has been shown to be important for (1,3;1,4)-β-glucan synthesis (Burton et al., 2006; Tonooka et al., 2009). To explore these genes in greater detail FastA files containing the groomed reads associated with each individual line were converted to databases and blasted against the genomic sequence information for *CslF3, 4, 6, 8, 9, 10, 11, 12* and *13* (Figure 5.8). *CslF3*, 10 and 12 showed no sequence coverage. *CslF4* revealed three areas with sequence coverage. The second exon had sequence information for 74 bp in 80 of the genotypes studied. This sequence information falls within the second transmembrane domain and is likely to be highly conserved. Two other sites in the third exon were also observed. The first was 72 bp region from five genotypes and the second was a 94 bp region within the third transmembrane domain starting at L630 found in 73 genotypes studied. *CslF6* had coverage in two areas, in all varieties, covering 116 bp in the first intron and 172 bp in the third intron. The sequence coverage in the first intron starts from nucleotide 690 after the transcriptional start site and encompasses an area in which only two sequence variants in a single genotype have been reported in OUT329, a hulless six-row barley genotype from Turkey (Taketa et al., 2012). The area covered in the third intron falls in a highly conserved domain encompassing the catalytic site and the third transmembrane domain. No polymorphisms were identified in these regions within the 94 genotypes in this study. *CslF8* has sequence coverage for 168 bp in all lines in the promoter region. *CslF9* displayed one region with sequence
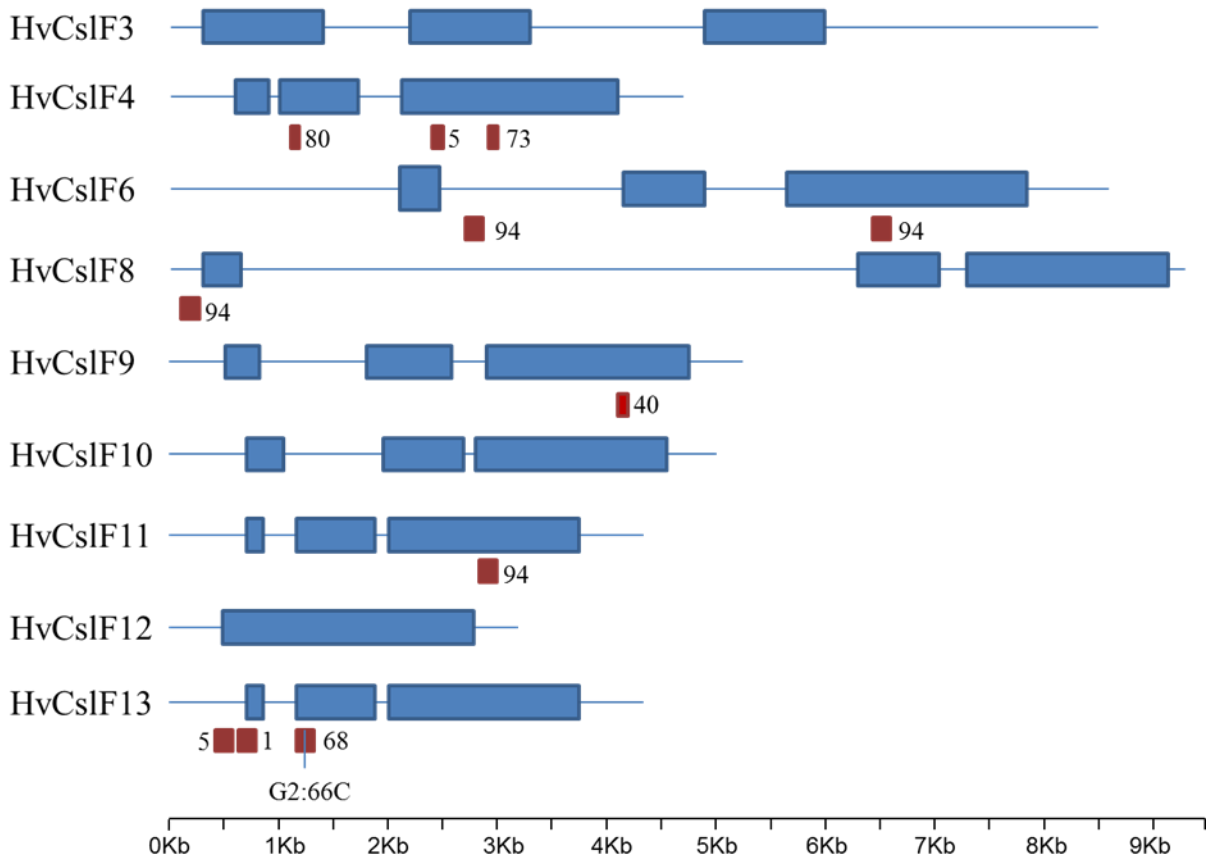
Figure 5.8. Schematic representation of nine *HvCslF* genes. RNA coding regions are represented by blue squares. Non-coding regions are represented by lines. Red squares indicate areas where sequence information was obtained through ddRAD sequencing. Numbers beside the red blocks indicate the number of genotypes in which a minimum of sequence information with a minimum read depth of six was achieved. SNP variant (G2:66C) is denoted as a ratio of each SNP variant.

coverage in the third intron starting at the sequence encoding G742 in 40 genotypes. This region is likely highly conserved as it is part of the sixth transmembrane domain in *CslF9*. *CslF11* had sequence coverage across 157 bp in the third intron. This area covers the catalytic domain and part of the third transmembrane domain. *CslF13* had three areas of coverage. Five genotypes had sequence coverage in the promoter. The hooded forage barley "Stockford" had 180 bp of sequence from the first exon. *CslF13* was the only gene in this group that displayed a variant SNP call within the sequence information gathered. 72 bp was sequenced from 68 individuals in an area corresponding to the first transmembrane domain. The SNP identified would cause an amino acid shift from alanine to glutamine at position 236. The combination of filters used on our dataset would have excluded this SNP from our analysis.

5.6 Conclusion

NGS technology is a powerful tool that has vast potential for association genetics. Large differences were observed in the number of markers identified by the SAMtools and UNEAK bioinformatics pipelines but much of the downstream analysis showed a high level of convergence. Despite the differences between bioinformatics pipelines two distinct subpopulations can be determined. Large differences in the number of markers identified were caused by the different bioinformatics pipeline used but many of the associations were similar between the datasets. All statistical models identify a region on the telomere of 5H that is associated with $(1,3;1,4)$-β-glucan grain concentration. Other candidate genes identified include two putative transcription factors, MLOC_37432 and MLOC_21017, and two putative phosphatase signalling proteins MLOC_3791 and MLOC_53722, which may be involved in cell wall biosynthetic regulation. MLOC_56623 a putative tSNARE domain containing protein involved in vesicle transport and MLOC_73315 an uncharacterised gene displaying a UDP-glucuronosyl / UDP-glucosyltransferase domain.

CHAPTER 6
DISCUSSION

6.1 General discussion

In terms of production, barley is the fourth most important cereal grain after maize, rice and wheat (FAO, 2014). Barley is a rich source of dietary fiber and an increasingly important cereal in human nutrition. One of the major reasons for the renewed popularity of barley in human nutrition is the United States Food and Drug Administration decision that foods containing soluble fiber from barley can be labeled with a health claim related to reducing blood cholesterol in humans (FDA, 2006). The major soluble fiber in barley grains is $(1,3;1,4)$-β-glucan, a mixed linkage polymer made up of repeated units of cellotrioses and cellotetraoses and with very small amounts of higher polymers cellopentoses, cellohexoses or higher orders of glucans. The composition of barley $(1,3;1,4)$-β-glucan is also very important as it influences its viscosity in solutions, a major determinant of the human health benefits associated with barley.

Barley $(1,3;1,4)$-β-glucan concentration is a quantitative trait, influenced by environmental (Morgan and Riggs, 1981; Perez-Vendrell et al., 1996) and genetic factors (Powell et al., 1985). The possibilities of in-depth study of quantitative traits in crops has evolved rapidly during the last decade due to significant advances in DNA sequencing technologies and identification of new molecular markers to detect genotypic differences. Interestingly, barley grain $(1,3;1,4)$-β-glucan concentration has been associated with Quantitative Trait Loci (QTLs) located on all the seven barley grain chromosomes (Han et al., 1995; Baum et al., 2003; Islamovic et al., 2013; Molina-Cano et al., 2007), but no genetic regions have been associated with $(1,3;1,4)$-β-glucan structure elements such as DP3:DP4 ratio or concentrations of cellotrioses (DP3) and cellotetroses (DP4). The work presented in this thesis was based on the hypothesis that *"Genomic regions near the centromere on 7H play a role in (1,3;1,4)-β-glucan grain concentrations and fine structure"*. To test this hypothesis the work was performed as three objectives (Section 1.3).

6.2 Development and validation of *HvCslF6* markers for grain (1,3;1,4)-β-glucan concentration in barley

Genetic mapping of a doubled haploid population developed from a cross between CDC Bold and TR251 identified a major QTL on chromosome 7H, that explained up to 39 % of the variation for grain (1,3;1,4)-β-glucan concentration (Li et al., 2008). The QTL on chromosome 7H contained a cellulose synthase like gene, *HvCslF6,* for which parental alleles were analyzed by DNA sequencing. Nucleotide sequence comparison of 6.4 kb *HvCslF6* contigs derived from CDC Bold and TR251 revealed 16 SNPs and two indels within the transcribed region. Two indels and 12 of the SNPs were positioned within introns or noncoding regions, three of the SNPs within the exons were silent, but SNP-4105 in the third exon converted an alanine codon in *HvCslF6* of CDC Bold to a threonine codon in the TR251 sequence (Chapter 3). Screening of the parents of six independent mapping populations showed that populations with a QTL on 7H for grain or wort (1,3;1,4)-β-glucan concentration had haplotypes identical to TR251 or CDC Bold. Through an association mapping study of 119 barley genotypes it was confirmed that CSLF6_4105 TR251 allele was associated with high and CDC Bold allele with low grain (1,3;1,4)-β-glucan concentration (Chapter 4, Table 4.5).

6.3 Determination of the heritability of (1,3;1,4)-β-glucan fine structure

An HPAEC-PAD based method was optimized to study the fine structure (DP3:DP4) of barley (1,3;1,4)-β-glucan (Chapter 4). The analysis of grain produced by 91 lines of a Merit / H93174006 RIL6 population grown in two different environments indicated DP3:DP4 ratios in (1,3;1,4)-β-glucan were moderately heritable and significantly affected by genotype, environment and GxE interactions. Single marker analysis by ANOVA showed that the CSLF6_4105 marker was significantly associated with both (1,3;1,4)-β-glucan concentration and DP3:DP4 ratio. The effectiveness of the marker differed between the two growth environments, where a stronger association between alleles at 7H QTL and (1,3;1,4)-β-glucan properties was noted for grains produced in the drier environment. An environmental impact on the strength of the 7H QTL for (1,3;1,4)-β-glucan concentration is also present for the TR251 / CDC Bold population (Li et al., 2008). To further analyze the impact of CSLF6_4105 marker on (1,3;1,4)-β-glucan concentration and fine structure, an association mapping panel of 119 barley genotypes were phenotyped and subjected to a general linear model (GLM) analysis. The results indicated

that CSLF6_4105 is significantly associated with (1,3;1,4)-β-glucan concentration but not its fine structure. However, (1,3;1,4)-β-glucan fine structure could be associated with a CSLF6_4105 linked region on chromosomes 7H and an additional marker on chromosome 1H (Chapter 4, Table 4.5).

6.4 Implications of results on the control of (1,3;1,4)-β-glucan synthesis

Buckeridge et al. (1999) proposed a model where multiple sites on the (1,3;1,4)-β-glucan synthase complex synthesize three β-(1,4) linkages at a time, followed by a β-(1,3) linkage in optimum conditions or longer cellodextrin units if the cellular UDP-Glc concentration is less than a certain threshold (Figure 2.4). Unfortunately, Buckridge et al. (1999) was unable to show that varying the amount of substrate produced the same effect *in vivo* leading the authors to speculate the involvement of other regulating factors. Recently, it has been shown that the environment plays a role in determining DP3:DP4 ratios in (1,3;1,4)-β-glucan produced in field grown oats (Doehlert and Simsek, 2012) and barley (Chapter 4).

The cellulose synthase like gene *CslF6* has been extensively studied for its role in barley (1,3;1,4)-β-glucan synthesis. A high *CslF* expression correlates with a relatively high (1,3;1,4)-β-glucan concentration in barley grains of genotype Himalaya and a low *CslF6* expression is seen in the elite malting barley variety Sloop producing low (1,3;1,4)-β-glucan concentration in grain (Burton et al., 2008). In contrast, high lysine mutants such as *lys5f* and *lys5g* with relatively high grain (1,3;1,4)-β-glucan concentrations show reduced *CslF6* expression when compared to their wild type parent Bomi with low grain (1,3;1,4)-β-glucan content (Christensen and Scheller, 2012). In the *lys5g* mutant, reduced *CslF6* expression is accompanied by low (1,3;1,4)-β-glucan synthase activity at 20 days after pollination as compared to Bomi, but the (1,3;1,4)-β-glucan synthase activity in the *lys5f* mutant is unaffected despite low *CslF6* expression. These inconsistencies between *CslF6* expression and (1,3;1,4)-β-glucan production levels may be due to post-transcriptional or epigenetic regulation of genes involved in (1,3;1,4)-β-glucan biosynthesis such as *CslF6*. An argument for the latter was provided from a study by Christensen and Scheller (2012) indicating possible methylation of region surrounding *CslF6* in high lysine *lys3a* genotype showing an extreme suppression of *CslF6* expression (1/1000th of wild type). Despite *CslF6* down-regulation in the *lys3a* mutant, the (1,3;1,4)-β-glucan concentration is relatively high throughout endosperm development. Thus, additional factors acting at the post-

transcriptional level may boost CSLF6 activity levels in *lys3a* mutant to allow efficient (1,3;1,4)-β-glucan production. It would be interesting to test if such regulation involves the putative phosphorylation site (A590T) on CSLF6 identified for the TR251 line.

Transgenic overexpression of *CslF6* leads to increases grain (1,3;1,4)-β-glucan content and lower DP3:DP4 ratio (Burton et al., 2011). In the *lys5g* mutant with low *CslF6* expression, the DP3:DP4 ratios is higher in comparison to its parent Carlsberg II (Mikkelsen et al., 2013; Christensen and Scheller, 2012). In this context it may be reasonable to assume that higher *CslF6* expression increases the amount of active CSLF6 creating a correlation between high (1,3;1,4)-β-glucan concentrations and low DP3:DP4 ratios. This may also be consistent with a model suggested by Urbanowicz et al. (2004; Figure 2.4.), where an increase in CSLF6 proteins intensifies the competition for UDP-Glc substrate molecules and reduces their relative concentration favoring the production of longer cellodextrin subunits. However, the results presented in Chapter 4 do not seem to support this model as the more active *CslF6* allele of TR251 produces higher DP3:DP4 ratio in suboptimal environments than the less active *CslF6* allele of CDC Bold (Chapter 4, Table 4.4). However, the apparent discrepancy between previous reports and this study could be caused by differences in enzyme characteristics. Assuming TR251 CSLF6 590T enzyme has a higher affinity for UDP-Glc than the CSLF6 590A isoform, the difference would lead to a higher DP3:DP4 ratio in CSLF6 590T lines as compared to CSLF6 590A lines when grown under conditions with reduced concentration of free UDP-glucose. In barley grown at Vegreville, approximately 8.9 % of the variation could be explained by the CSLF6_4105 marker as the mean ratio for the CSLF6 590T allele was significantly higher than for the CSLF6 590A allele. According to this hypothesis, the CSLF6 590T enzyme may have a higher specific activity for substrate than the CSLF6 590A isoform.

Phosphorylation has been shown to regulate β-glucan synthesis in Arabidopsis (Chen et al., 2010; Taylor, 2007), pea (Ray, 1973), tobacco (Kaida et al., 2009) and corn (Paliyath and Poovaiah, 1988). Mutation studies of *AtCesA1* have shown that mimicking constitutive phosphorylation of serine residues within the class specific domain increases cellulose production (Chen et al., 2010). Putative phosphorylation sites are also present on barley CSLF6 as suggested by bioinformatics analysis using NetPhos 2.0 and PhosPhat 3.0 software. One site of particular interest is located immediately N terminal of the variant A590T site (Chapter 3). If functional, threonine at this putative phosphorylation site may either enhance or mimic

phosphorylation, resulting in higher (1,3;1,4)-β-glucan synthesis. The sequence of the SHPSPY[T/A]A motif suggests it may be recognized by a Pro-directed kinase and more specifically a mitogen-activated protein (MAP) kinase. Pro-directed kinases have a stringent requirement for a proline in the n+1 position N terminal to the phosphorylation site and MAP kinases have a preference for an additional proline within the recognition motif (Pinna and Ruzzene, 1996). MAP kinases also require a D domain N terminal to the phosphorylation site and the domain should contain a LxL motif located 3–5 amino acids downstream from a region containing several basic residues (Sharrocks et al., 2000). The CSLF6 protein carries a putative D domain (KGKHGFLPL) 18 amino acids N terminal to the SHPSPYAA motif. Both the SHPSPYAA and LXL motifs are well conserved among CSLF6 variants of wheat (*Triticum aestivum*), rice (*Oryza sativa*), oats (*Avena sativa*) and Brachypodium supporting these elements have an important function. Whether differences in HvCSLF6 phosphorylation status caused by A590T variation is the underlying cause for 7H (1,3;1,4)-β-glucan QTL remains to be confirmed by further studies.

It is interesting to note that a recent association mapping study by Houston et al. (2014) did not find an association with *CslF6* gene. This may have been due to several factors including the structure of the population used in the study. The authors acknowledge that an attempt to associate a SNP in the third exon of *CslF6* may have been hampered by the selection against this allele in elite malting barley leading to an under representation of the TR251 allele in the association mapping panel studied. Another reason could be the effect of environment on the QTL associated with the *CslF6* marker. It was found in the second study (Chapter 4) and Li et al. (2008) that the 7H QTL is affected by environment. Since Houston et al. (2014) grew the population under glasshouse conditions it is likely that the effect caused by the allelic differences are not applicable in a low stress environment. The second study of 119 genotypes grown in greenhouse conditions (Chapter 4) using GLM, CSLF6_4105 marker was found to be associated grain (1,3;1,4)-β-glucan concentration. But when a panel of 94 barley genotypes were investigated, corrected for population structure and a higher FDR, no significant association could be detected between CSLF_4105 marker and (1,3;1,4)-β-glucan concentration. In an extreme example of environmental control, *OsCslF6* expression and (1,3;1,4)-β-glucan synthase activity is reduced in rice seedlings grown submerged compared to dry growing conditions

(Kimpara et al., 2008). It would be interesting to investigate the effect of drought stress on *HvCslF6* expression and (1,3;1,4)-β-glucan synthase activity in barley.

6.5 Association mapping to identify novel markers for (1,3;1,4)-β-glucan concentration in barley

In the association mapping study, ddRAD was used to genotype a population of 94 two-row spring genotypes to find novel markers associated with (1,3;1,4)-β-glucan concentration. The bioinformatics pipeline used created large differences in the number of markers detected, nevertheless the downstream results were similar. Both datasets could be used to distinguish substructure within this population which predicted two groups that could be roughly categorized as Canadian varieties and non-Canadian varieties. Statistical models integrating kinship matrix and / or substructure matrix generated through the Bayesian algorithm STRUCTURE v 2.3.4 or ordination analysis by PCoA lead to different markers associations (Table 5.1). Despite these differences a number of chromosomal regions were found to be associated regardless of the dataset or statistical model used. The region on 5H is of particular interest as it was identified by three different statistical models and both datasets (Chapter 5; Table 5.1) and has been previously identified in an association mapping study and a bi-parental genetic mapping population (Shu and Rasmussen, 2014; Islamovic et al., 2013). Islamovic et al. (2013) identified a syntenous region in rice near Os03g59340 (*OsCesA2*) whereas Shu and Rusmussen (2014) made no predictions of putative candidate genes in this specific genomic region. It is unlikely that *HvCesA2* is the gene responsible for this QTL as it is located on the opposite telomere on 5H in barley according to current physical map (Mayer et al., 2012) and genetic map of Burton et al. (2004). From the EnsemblPlant annotated database (version 23), one putative uncharacterized gene (MLOC_21074) was found in this region of the 5H telomere. A Pfam search showed that MLOC_21074 contains a zf-RVT domain which is a zinc binding domain commonly found in reverse transcriptases. This most likely suggests that this gene as a transposable element but the significance is entirely unknown.

Christensen and Scheller (2012) suggested that there is a sensing mechanism at the cell wall which monitors the amount of (1,3;1,4)-β-glucan and influences *HvCslF6* expression. Two transcription factors were identified as candidate genes for (1,3;1,4)-β-glucan in the AM study (Chapter 4). MLOC_21017 is an aparataxin related basic helix loop helix protein gene located at

QTL on chromosome 4H and MLOC_37432 is a C2H2 zinc finger domain protein gene positioned at QTL on chromosome 1H, both of which are likely involved in transcription repression. Christensen and Scheller (2012) have shown transcriptional repression of *CslF6* by the *lys3a* mutation decreases the concentration of (1,3;1,4)-β-glucan in barley grains and postulated that signals from the cell wall can influence *CslF6* expression. Recently it has been suggested that *CslF6* transcription levels do differ between TR251 and CDC Bold but no causative agent could be found within the gene or promoter (Professor Diane Mather, University of Adelaide, Australia, personal conversation). Transcriptional regulation of *CslF6* is likely to play a role in (1,3;1,4)-β-glucan synthesis but it would probably be controlled by a *trans* acting element such as a transcription factor or by epigenetic factors.

UDP-Glu is known to be the substrate required for cellulose, callose and (1,3;1,4)-β-glucan. MLOC_73315 is an uncharacterised gene at the QTL on chromosome 7H (Chapter 5) displaying a UDP-glucuronosyl / UDP-glucosyltransferase domain. This protein may have a role in production of UDP-Glc, which is the main building block for cellulose, callose and (1,3;1,4)-β-glucan biosynthesis. The inclusion of a UDP-glucosyltransferase as a subunit of (1,3;1,4)-β-glucan synthase complex has been proposed by Urbanowicz et al. (2004).

As mentioned earlier, phosphorylation is emerging as a major control mechanism in cell wall biosynthesis. From the AM study, two putative phosphorylation signalling genes were identified that could encode proteins with signalling functions. One candidate gene of interest is MLOC_3791, a predicted acid phosphatase gene. Transgenic overexpression of *purple acid phosphatase 12 (NtPap12)* increase cellulose and callose synthesis in tobacco cells (Kaida et al., 2009). In this regard MLOC_3791 may play a similar role in (1,3;1,4)-β-glucan synthesis, though additional studies will be needed to confirm this speculation. The second putative signaling gene was located near MLOC_53722, annotated as a putative serine / threonine-protein kinase gene. Pfam blast search reveals several domains within this protein including a bulb-type lectin domain, ATP binding site and an S-receptor-like serine / threonine-protein kinase domain. This predicted 839 amino acid protein contains two possible transmembrane domains one near the N-terminal from amino acid 2 - 24 and a second at amino acids 451 - 486 (Cao et al., 2006). S-receptor-like serine / threonine-protein kinase proteins respond to external signals and have been shown be involved in wound response in some plants (Pastuglia et al., 1997; 2002).

The results of association mapping study in chapter 5 suggest an extra layer of control of (1,3;1,4)-β-glucan synthesis in the barley grain. Increasing evidence has shown that phosphorylation of enzymes participating starch biosynthesis (Genberger, 2011) and production and degradation of various glucan polymers of the cell wall (Chen et al., 2010; Taylor, 2007). One example of the role of phosphorylation of CESA was recently demonstrated for the elongating internode of maize (Zhang et al., 2014). Microarray data shows that there is a disparity between *CesA* and *Csl* activity and accumulation of cellulose and hemicellulose products (Zhang et al., 2014). When a global correlation analysis was applied to secondary cell wall *ZmCesA* genes it was noted that two genes encoding protein kinases (Q653F8 and Q75V63) are highly correlated with cell wall biosynthesis. While cell wall biosynthesis may be grossly controlled by gene expression, increasing evidence suggests that control of the biosynthetic machinery is occurring at the post translation level (Zhang et al., 2014; Chen et al., 2010; Taylor, 2007; Christensen and Scheller, 2012).

Despite common marker systems and diverse barley genotypes studied, no common allele explaining a majority of the variance in (1,3;1,4)-β-glucan concentration in barley have been found. In association mapping the statistical association among a set of loci will decay more or less quickly depending on the amount of recombination events that have occurred during meiosis (Dawson, 2000). By studying individuals from a wide variety of genetic backgrounds a higher number of ancient meiotic events can be assumed to have occurred, allowing for greater mapping resolution. The advantage of association mapping is that as genetic diversity increases in the population the amount of ancient meiotic events also increases leading to ever increasing mapping resolution. The disadvantage of association mapping is that as the amount of genetic diversity may increase the number of alleles affecting the trait in question could also potentially increase, leading to a reduction in the amount of variation explained by each allele. This is a common theme in human medicine of "common disease, many rare variants" (Ingvarsson and Street, 2010). Genetic variants causing diseases should be expected to be associated with negative selection pressure which, by its very nature, would remove any common variants from the population. In this sense grain (1,3;1,4)-β-glucan concentration may be looked at through the lens of human disease association mapping. Intensive selection pressure against grain (1,3;1,4)-β-glucan concentration in malting breeding programs may have had a similar effect. This may be
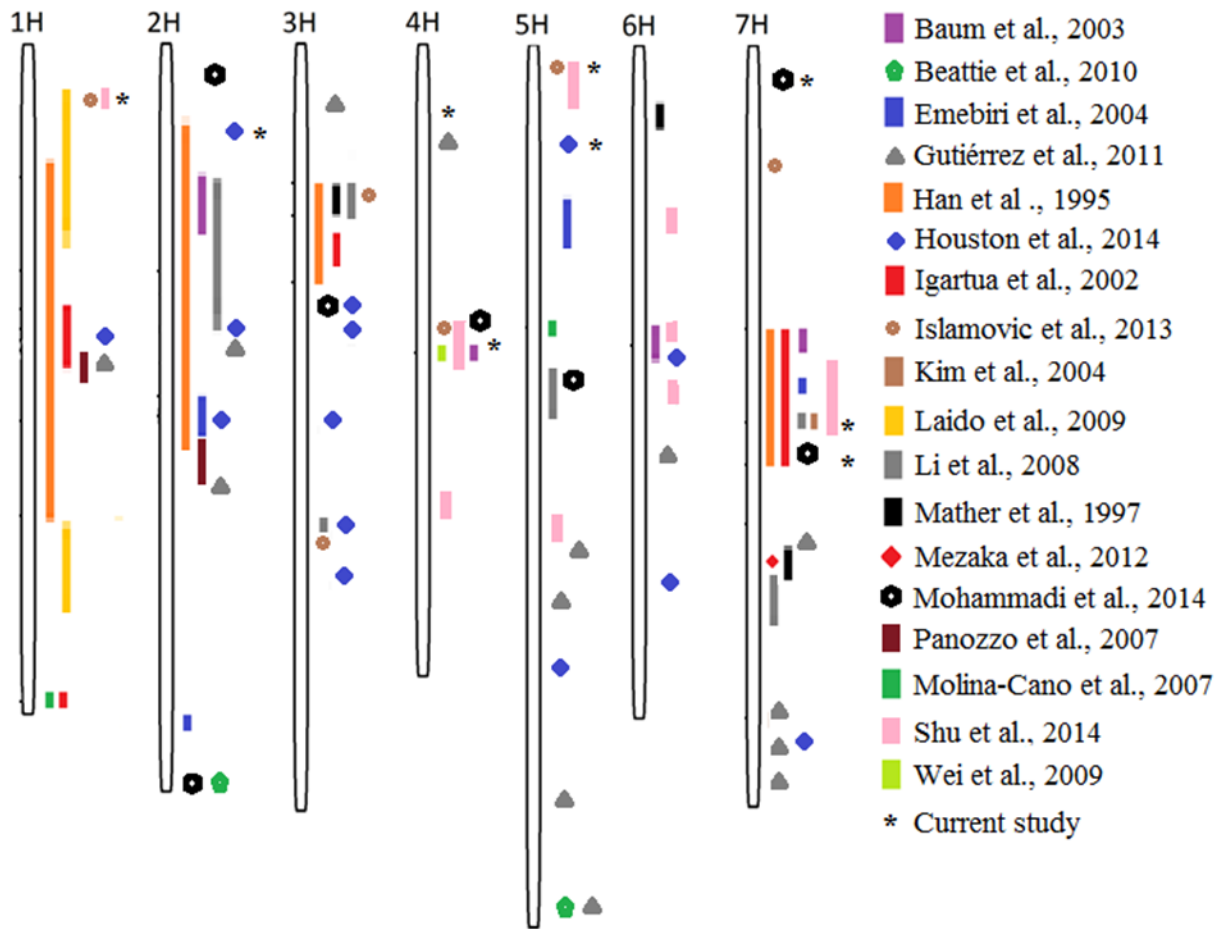
Figure 6.1. An updated summary of genomic regions associated with barley (1,3;1,4)-β-glucan grain or wort concentrations.

why there has been no "common variant" found in association mapping panels when mapping grain (1,3;1,4)-β-glucan concentration (Figure 6.1).

6.6 Future directions

Recently, Houston et al. (2014) have identified a marker associated with (1,3;1,4)-β-glucan that is predicted to be near *HvCslF9* but has not identified an underlying reason for the variation caused by this marker. Direct sequencing of *HvCslF9* from Morex and Steptoe may be able to identify a causative mutation and lead to a functional marker within this gene or rule it out as a possible candidate.

Recent studies have shed new light on the possible location, orientation and action of CSLF6. (Kim et al., article in press). Transgenic expression of *CslF6* from the model grass species *Brachypodium distachyonin* was performed in two heterologous systems, tobacco epidermal cells and the yeast *Pichia pastoris.* YFP-tagged BdCSLF6 shows localization to the Golgi apparatus similar to localization of CSLH1 by Dolbin et al. (2009). Subsequent analysis using anti-(1,3;1,4)-β-glucan immunogold labeling identifies (1,3;1,4)-β-glucan in the cell wall and Golgi apparatus of transformed tobacco cells. Topological studies using proteinase treatments indicate that the catalytic site is exposed to the cytosol and the protein contains an even number of transmembrane domains similar to the model proposed by Urbanowicz et al. (2004). Unlike the Urbanowicz model, the authors argue that BdCSLF6 alone can produce a full length (1,3;1,4)-β-glucan since heterologous expression of BdCSLF6 in *Pichia* resulted in a the production of (1,3;1,4)-β-glucan. The authors also argue that the evolutionary distance between yeast and plants makes it unlikely that a yeast glucosyltranferase would work with a plant CSLF6 protein.

The strongest direct evidence of the involvement of *CslF6* in (1,3;1,4)-β-glucan synthesis in barley has come from mutations affecting the protein (Tonooka et al., 2009; Taketa et al., 2012; Hu et al., 2014). Point mutations near the conserved aspartic acid residues and in the transmembrane domain have shown a drastic effect on the amount of (1,3;1,4)-β-glucan accumulated in the cell. On the other hand, constitutive expression of *CslF* suggest that the proteins may not act alone as high (1,3;1,4)-β-glucan concentration is not always obtained. Heterologous expression of rice *CslF* genes in *Arabidopsis* results in the synthesis of (1,3;1,4)-β-glucan preferentially in epidermal cells (Burton et al., 2006) and overexpression of *CslF* in

barley results in non-uniform deposition of (1,3;1,4)-β-glucan (Burton et al., 2011). This might suggests that *CslF6* is one component of a larger protein complex responsible for the synthesis of (1,3;1,4)-β-glucan in grain. Synthesis of large glucose polysaccharides such as starch and cellulose involve highly controlled and large protein complexes (Chen et al., 2010; Tetlow and Emes, 2014). Analysis of the *CesA / Csl* superfamily show that the CLSF and CESA proteins are very closely related sharing a greater amount of homology compared to almost all other Csl genes. CESA proteins in plants have been shown to function as CESA homo-oligomers stabilised by the class specific regions (CSR) and plant specific regions (PSR) in the cytosolic regions of the protein (Sethaphong et al., 2014). Furthermore, the chimeric protein analysis shows that the C-terminal domains including the CSR and PSR confer specificity as to their specific protein partners (Wang et al., 2006). The CslF family of proteins may have a similar system of function coupling with other Csl proteins or other glucosyltransferases to form a functional (1,3;1,4)-β-glucan synthesis apparatus. Considering (1,3;1,4)-β-glucan can be produced in Arabidopsis and tobacco, plants which do not create (1,3;1,4)-β-glucan, the partnering protein may be either highly evolutionarily conserved throughout plants or the CSLF core synthase may not be highly selective and weakly active in the absence of partners.

Overexpression of *CslF6* in barley using a constitutive promoter did result in three to four fold higher (1,3;1,4)-β-glucan concentration accumulation in the leaves. However, despite higher transcription levels in the grain there was no appreciable increase in grain (1,3;1,4)-β-glucan concentration. Interestingly, constitutive expression of *CslF4*, which is naturally found at moderate levels in the leaves did not increase the leaf (1,3;1,4)-β-glucan concentration, but did have a drastic effect in the grain (Burton et al., 2006). These results do suggest some level of control of at the post-transcriptional level in these respective tissues. It would be interesting to investigate the role of phosphorylation and sugar signalling as possible feedback mechanisms on enzyme activity and gene expression. While QTL analysis does point towards the involvement of the CSLF proteins only this thesis has offered a possible explanation of how natural genetic variation in *CslF6* may influence (1,3;1,4)-β-glucan synthesis *in vivo*. A possible phosphorylation site influenced by a second messenger system seems likely but remains unconfirmed. Biochemical analysis looking at the differences in enzymatic activity between the CSLF 590T and CSLF6 590A protein could prove our hypothesis that this amino acid shift causes differences in (1,3;1,4)-β-glucan synthesis.

LITERATURE CITED

Aastrup S, Jørgersen K (1988). Quantification of high molecular-weight (1-3)(1-4)-beta-D-glucan using calcofluor complex-formation and flow-injection analysis. 2. Determination of total beta-glucan content of barley and malt. Carlsberg Research Communications. 53:287-296.

Acquaah G (2007). *Principles of Plant Genetics and Breeding*. Oxford, UK: Blackwell Publishing.

Amado R, Neukom H (1985). Minor constituents of wheat flour: The pentosans. *In*: Hill RD, Munck L (eds.). *New approaches to research on cereal carbohydrates* (pp. 241-251). Amsterdam, The Netherlands: Elsevier Science Publishers.

Åman P, Hesselman K, Tilly AC (1985). Variation in the chemical composition of Swedish barleys. Journal of Cereal Science. 3:73-77.

Åman P, Newman CW (1986). Chemical composition of some different types of barley grown in Montana, U.S.A..Journal of Cereal Science. 4: 133-141.

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

Ames N, Rhymer C, Rossnagel B, Therrien M, Ryland D, Dua S, Ross K (2006). Utilization of diverse hulless barley properties to maximize food product quality. Cereal Foods World. 51:23-38.

Andersen KE, Bjergegaard C, Møller P, Sørensen JC, Sørensen H (2005). Compositional variations for r-Galactosides in different species of leguminosae, brassicaceae, and barley: A chemotaxonomic study based on chemometrics and highperformance capillary electrophoresis. Journal of Agriculture and Food Chemistry. 53:5809-5817.

Anderson JW (1995). Dietary fibre, complex carbohydrate and coronary artery disease. Canadian Journal of Cardiology. 11:55G–62G.

Anderson OD, Huo N, Gu YQ (2013). The gene space in wheat: the complete γ-gliadin gene family from the wheat cultivar Chinese Spring. Functional & Integrative Genomics. 13:261-273.

Andersson AAM, Kamal-Eldin A, Fras A, Boros D, Åman P (2008). Alkylresorcinols in wheat varieties in the HEALTHGRAIN diversity screen. Journal of Agricultural and Food Chemistry. 56:9722–9725.

Aniskov NI, Kalashnik NA, Kozlova GYA, Garis DV (2008). Inheritance of grain protein content upon crossing covered and hulless barley biological varieties. Russian Agricultural Sciences. 34:1-2.

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JD, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M (2010). Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature. 465: 627-631.

Ayele AG (2011). Heritability and genetic advance in recombinant inbred lines for drought tolerance and other related traits in Sorghum (Sorghum bicolor L.). Continental Journal of Agricultural Science. 5:1-9.

Azhaguvel P, Vidya-Saraswathi D, Komatsuda T (2006). High-resolution linkage mapping for the non brittle rachis locus *btr1* in cultivated × wild barley (*Hordeum vulgare*). Plant Science. 170:1087–1094.

Bacic A, Stone B (1981). Isolation and ultrastructure of aleurone cell-walls from wheat and barley. Australian Journal of Plant Physiology. 8:453-474.

Baik BK, Ullrich SE (2008). Barley for food: Characteristics, improvement, and renewed interest. Journal of Cereal Science. 48:233-242.

Balding D (2006). A tutorial on statistical methods for population association studies. Nature Reviews Genetics. 7:781–791.

Batra VIP, Bansal HC, Mehta SL (1982). Carbohydrate composition of developing grains of the high-lysine barley mutant (notch-2) and its parent (NP 113). Journal of Science Food and Agriculture. 33: 30-34.

Baum M, Grando S, Backes G, Jahoor A, Sabbagh A, Ceccarelli S (2003). QTLs for agronomic traits in the mediterranean environment identified in recombinant inbred lines of the cross 'arta' x H-spontaneum 41-1. Theoretical and Applied Genetics. 107:1215-1225.

Beattie AD, Edney MJ, Scoles GJ, Rossnagel BG (2010). Association mapping of malting quality data from Western Canadian two-row barley cooperative trials. Crop Science. 50:1649-1663.

Becker M, Vincent C, Reid JS (1995). Biosynthesis of (1,3)(1,4)-beta-glucan and (1,3)-beta-glucan in barley (Hordeum vulgare L.). Properties of the membrane-bound glucan synthases. Planta. 195:331-338.

Beecher G (2004). Proanthocyanidins: biological activities associated with human health. Pharmaceutical Biology. 42:2-20.

BeMiller JN, Whistler RL (2009). *Starch: Chemistry and Technology.* 3rd ed. Orlando, FL: Academic Press.

Bengtsson S, Åman P, Andersson A (1992). Structural studies on water-soluble arabinoxylans in rye grain using enzymatic hydrolysis. Carbohydrate Polymers. 17:277-284.

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. 57: 289–300.

Berry PM, Sterling M and Mooney SJ (2006). Development of a model of lodging for barley. Journal of Agronomy and Crop Science. 192:151-158.

Bhatty RS (1982). Distribution of lipids in embryo and bran - endosperm fractions of Risø 1508 and Hiproly barley grains. Cereal Chemistry. 59: 154–156.

Bhatty RS (1999). The potential of hull-less barley. Cereal Chemistry. 76:589-599.

Bhatty RS, Rossnagel BG (1997). Zero amylose lines of hull-less barley. Cereal Chemistry. 74:190-191.

Bhatty RS, Rossnagel BG (1998). Comparison of pearled and unpearled Canadian and Japanese barleys. Cereal Chemistry. 75:15-21.

Blom N, Gammeltoft S, Brunak S (1999). Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. Journal of Molecular Biology. 294:1351-1362.

Bothmer R, Jacobsen N, Baden C, Jorgensen R, Linde-Laursen I (1995). *An Ecogeographical Study of the Genus Hordeum.* Rome:IBPGR.

Bourdon I, Yokoyama W, Davis P, Hudson C, Backus R, Richter D, Knuckles B, Schneeman BO. (1999). Postprandial lipid, glucose, insulin, and cholecystokinin responses in men fed barley pasta enriched with β-glucan. American Journal of Clinical Nutrition. 69:55-63.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007). TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics. 23:2633-2635.

Bravi E, Marconi O, Perretti G, Fantozzi P (2012). Influence of barley variety and malting process on lipid content of malt. Food Chemistry. 135:1112-7.

Brennan CS, Cleary LJ (2005). The potential use of cereal (1-3,1-4)-beta-D-glucans as functional food ingredients. Journal of Cereal Science, 42:1‑13.

Buckeridge M, Rayon C, Urbanowicz B, Tine M, Carpita N (2004). Mixed linkage (1→3),(1→4)-β-D-glucans of grasses. Cereal Chemistry. 81:115-127.

Buckeridge M, Vergara C, Carpita N (1999). The mechanism of synthesis of a mixed-linkage (1→3),(1→4)β-D-glucan in maize. Evidence for multiple sites of glucosyl transfer in the synthase complex. Plant Physiology. 120:1105-1116.

Buckler ES, Thornsberry JM (2002). Plant molecular diversity and applications to genomics. Current Opinion in Plant Biology. 5:107-111.

Burton RA, Gidley MJ, Fincher GB (2010). Heterogeneity in the chemistry, structure and function of plant cell walls. Nature Chemical Biology. 6:724-732.

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

Burton RA, Collins HM, Kibble NAJ, Smith JA, Shirley NJ, Jobling SA, Henderson M, Singh RR, Pettolino F, Wilson SM, Bird AR, Topping DL, Bacic A, Fincher GB (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.

Burton RA, Fincher GB (2009). (1,3;1,4)-beta-D-glucans in cell walls of the poaceae, lower plants, and fungi: A tale of two linkages. Molecular Plant. 2:873-882.

Burton RA, Jobling SA, Harvey AJ, Shirley NJ, Mather DE, Bacic A, Fincher GB (2008). The genetics and transcriptional profiles of the cellulose synthase-like *HvCslF* gene family in barley. Plant Physiology. 146:1821-1833.

Burton RA, Shirley NJ, King BJ, Harvey AJ, Fincher GB (2004). The CesA gene family of barley. Quantitative analysis of transcripts reveals two groups of co-expressed genes. Plant Physiology. 134:224-236.

Cao B, Porollo A, Adamczak R, Jarrell M, Meller J (2006). Enhanced Recognition of Protein Transmembrane Domains with Prediction-based Structural Profiles, Bioinformatics. 22:303-9.

Carbonero P, Garcia-Olmedo F (1999). A multigene family of trypsin-amylase inhibitors from cereals. *In*: Shewry PR, Casey R (eds.). *Seed Proteins* (pp. 617-633). Dordrecht: Kluwer Academic Publishers.

Cattell RB, Vogelmann S (1977). A comprehensive trial of scree and KG criteria for determining the number of factors. Multivariate Behavioral Research. 12:289-325.

Cavallero A, Empilli S, Brighenti F, Stanca AM (2002). High (1-3,1-4)-β-glucan barley fractions in bread making and their effects on human glycemic response. Journal of Cereal Science. 36: 59-66.

Chen S, Ehrhardt DW, Somerville CR. (2010). Mutations of cellulose synthase (CESA1) phosphorylation sites modulate anisotropic cell expansion and bidirectional mobility of cellulose synthase. Proceedings of the National Academy Science of the United States of America. 107: 17188-17193.

Chibbar RN, Ganeshan S, Båga M, Khandelwal RN (2004). Carbohydrtae metabolism. *In*: Wringley C, Corke H, Walker C (eds). *Encyclopedia of Seed Science* (pp. 168-179). Oxford: Elsevier Academic Press.

Christensen U, Scheller HV (2012). Regulation of (1,3;1,4)-β-D-glucan synthesis in developing endosperm of barley *lys* mutants. Journal of Cereal Science. 55:69-76.

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica. 142:169-196.

Collins HM, RA, Topping DL, Liao M-L, Bacic A, Fincher GB (2010). Variability in fine structures of noncellulosic cell wall polysaccharides from cereal grains: potential importance in human health and nutrition. Cereal Chemistry. 87:272-282.

Comadran J, Thomas WTB, van Eeuwijk FA, Ceccarelli S, Grando S, Stanca AM, Pecchioni N, Akar T, Al-Yassin A, Benbelkacem A, Ouabbou H, Bort, J, Romagosa I, Hackett CA, Russell JR (2009). Patterns of genetic diversity and linkage disequilibrium in a highly structured Hordeum vulgare association-mapping population for the Mediterranean basin. Theoretical and Applied Genetics. 119:175-187.

Corrander J, Marttinen P, Mäntyniemi S (2006). *A Bayesian method for identification of stock mixtures from molecular marker data.* Fishery Bulletin, 104:550-558.

Cory AT, Båga M, Anyia A, Rossnagel BG, Chibbar RN (2012). Genetic markers for *CslF6* gene associated with (1,3;1,4)-beta-glucan concentration in barley grain. Journal of Cereal Science. 56:332-339.

Coutinho PM, Deleury E, Davies GJ, Henrissat B (2003). An evolving hierarchical family classification for glycosyltransferases. Journal of Molecular Biology. 328:307-317.

Cui W, Wood P, Blackwell B, Nikiforuk J (2000). Physicochemical properties and structural characterization by two-dimensional NMR spectroscopy of wheat beta-D-glucan - comparison with other cereal beta-D-glucans. Carbohydrate Polymers. 41:249-258.

Cui W, Wood PJ (2000). Relationships between structural features, molecular weight and rheological properties of cereal β-D-glucans. *In*: Nishinari K (ed.). *Hydrocolloids, Part 1: Physical Chemistry and Industrial Applications of Gels, Polysaccharides and Proteins* (pp. 159-168). Amsterdam, The Netherlands: Elsevier.

Dawson KJ (2000). The decay of linkage disequilibrium under random union of gametes: How to calculate Bennett's principal components. Theoritical Population Biology. 58:1-20.

Dawson KJ, Belkhir K (2001). A Bayesian approach to the identification of panmictic populations and the assignment of individuals. Genetics Research. 78:59-77.

Delcour J, Hoseney RC (2010). Starch. *In*: Delcour J, Hoseney RC (eds). *Principles in Cereal Science and Technology* (pp.23-51). 3rd ed. St. Paul, MN: American Association of Cereal Chemists.

Delseny M, Han B, Hsing Y (2010). High throughput DNA sequencing: The new sequencing revolution. Plant Science. 179:407-422.

Devos KM (2005). Updating the 'crop circle'. Current Opinion in Plant Biology. 8:155-162.

DeVries (2001). The definition of dietary fiber. AACC Report.

Doblin MS, Pettolino FA, Wilson SM, Campbell R, Burton RA, Fincher GB, 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.

Doehlert DC, Simsek S (2012). Variation in beta-glucan fine structure, extractability, and flour slurry viscosity in oats due to genotype and environment. Cereal Chemistry. 89:242-246.

Doerge R (2002). Mapping and analysis of quantitative trait loci in experimental populations. Nature. 3:43-52.

Doveri S, Lee D, Maheswaran M, Powell W (2008). Molecular markers: History, features and applications. *In*: Kole C, Abbot AG (eds). *Principles and Practices of Plant Genomics, Volume 1: Genome Mapping* (pp. 23-68). Enfield, USA: Science Publishers.

Duffus CM, Cochrane MP (1993). Formation of the barley grain: morphology, physiology and biochemistry. *In*: MacGregor AW, Bhatty RS (eds). *Barley: Chemistry and Technology* (pp. 31-72). St. Paul, MN: American Association of Cereal Chemists.

Duffus CM, Rosie R (1976). Changes in trace element composition of developing barley grain. Journal of Agricultural Science (Cambridge). 87:75-79.

Duran C, Appleby N, Edwards D, Batley J (2009). Molecular genetic markers: discovery, applications, data storage and visualisation. Current Bioinformatics. 61:16-27.

Durek P, Schmidt R, Heazlewood JL, Jones A, Maclean D, Nagel A, Kersten B, Schulze WX (2009). PhosPhAt: the Arabidopsis thaliana phosphorylation site database: an update. Nucleic Acids Research. 38:D828-D834.

Earl DA, vonHoldt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources. 4:359-361.

Eimert K, Villand P, Kilian A, Kleczkowski LA (1996). Cloning and characterization of several cDNAs for UDP-Glc pyrophosphorylase from barley (Hordeum vulgare) tissues. Gene. 170:227-232.

El Khoury D, Cuda C, Luhovyy BL, Anderson GH (2012). Beta glucan: health benefits in obesity and metabolic syndrome. Journal of Nutrition and Metabolism. 2012:851362.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011). A robust, simple genotype-by-sequencing (GBS) approach for high diversity species. PLos ONE. 6:e19379.

Emebiri L, Moody D, Panozzo J, Chalmers K, Kretschmer J, Ablett G (2003). Identification of QTLs associated with variations in grain protein concentration in two-row barley. Australian Journal of Agricultural Research. 54:1211-1221.

Erkkila AT, Herrington DM, Mozaffarian D, Lichtenstein AH (2005). Cereal fiber and whole-grain intake are associated with reduced progression of coronary-artery atherosclerosis in postmenopausal women with coronary artery disease. American Heart Journal. 150(1):94-101.

FAO (2014). Food and Agriculture Organization of the United Nations, Land Resources. http://www.fao.org/nr/land/databases information-systems/en/ (accessed Nov 27, 2014).

Farrokhi N, Burton, R, Brownfield L, Hrmova M, Wilson SM, Bacic A, Fincher GB (2006). Plant cell wall biosynthesis: genetic, biochemical and functional genomics approaches to the identification of key genes. Plant Biotechnology Journal. 4:145-167.

FDA news release (2005). http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2005/ucm108543.htm

Feingold DS, Neufeld EF, Hassid WZ (1958). Synthesis of a β-1,3-linked glucan by extracts of Phaseolus aureus seedlings. Journal of Biological Chemistry. 233:783-788.

Fincher GB (1975). Morphology and chemical composition of barley endosperm cell walls. Journal of the Institute of Brewing. 81, 116-122.

Fincher GB, Stone BA (1974). A water-soluble arabinogalactan-peptide from wheat endosperm. Australian Journal of Biological Sciences. 27:117-132.

Fincher GB, Stone BA (1986). Cell walls and their components in cereal grain technology. *In*: Pomeranz Y (ed.). *Advances in Cereal Science and Technology, Vol 8* (pp. 207–295). St. Paul, MI: American Association of Cereal Chemists.

Fincher GB. (2009). Revolutionary times in our understanding of cell wall biosynthesis and remodeling in the grasses. Plant Physiology. 149:27-37.

Flint-Garcia S, Thornsberry JM, Buckler ESIV (2003). Structure of linkage disequilibrium in plants. Annual Reviews in Plant Biology. 54:357–374.

Foster E, Prentice N (1987). Barley. *In*: Olson RA, Frey KJ (eds.). *Nutritional Quality of Cereal Grains: Genetic and Agronomic Improvement* (pp.303-309). Madison, WI: American Society of Agronomy.

Fujita M, Domon E, Doi Y (1999). Grain and starch characteristics of the double recessive lines for amylose-free and high amylose gene in barley. Breeding Science. 49:217-219.

Ganeshan S, Drinkwater JM, Repellin A, Chibbar RN (2010). Selected carbohydrate metabolism genes show coincident expression peaks in grains of in vitro-cultured immature spikes of wheat (Triticum aestivum L.). Journal of Agricultural and Food Chemistry. 58:4193-4201.

Gaut B (2002). Evolutionary dynamics of grass genomes. New Phytologist. 154:15-28.

Genberger P (2011). Regulation of starch biosynthesis in response to a fluctuating environment. Plant Physiology. 155:1566-1577.

Gibeaut D, Carpita N (1991). Tracing cell-wall biogenesis in intact-cells and plants – selective turnover and alteration of soluble and cell-wall polysaccharides in grasses. Plant Physiology. 97:551-561.

Gibeaut D, Carpita N (1993). Synthesis of $(1\rightarrow3)$, $(1\rightarrow4)$-β-D-glucan in the Golgi apparatus of maize coleoptiles. Proceedings of the National Academy of Sciences of the United States of America. 90:3850-3854.

Gibeaut D, Carpita N (1994). Biosynthesis of plant-cell wall polysaccharides. FASEB Journal. 8:904-915.

Gohl B, Alden S, Elwinger K, Thomke S (1978). Influence of β-glucanase on the feed value of barley for poultry and moisture content of excreta. British Poultry Science. 19:41-47.

Gore MA, Chia JM, Elshire RJ, Sun Q, Ersoz ES, Hurwitz BL, Peiffer JA, McMullen MD, Grills GS, Ross-Ibarra J, Ware DH, Buckler ES (2009). A first-generation haplotype map of maize. Science. 326:1115-1117.

Greber B, Waite D, Fahy B, Hylton C, Parker M, Laurie D, Smith AM, Denyer K (2000). Use of barley mutants to understand starch synthesis. Proceedings of Eighth International Barley Genetics Symposium. Adelaide, October 2000, vol. 1, pp. 196-198.

Gruppen H, Kormelink FJM, Voragen AGJ (1993).Water-unextractable cell wall material from wheat flour. 3. A structural model for arabinoxylans. Journal of Cereal Science. 18:111-128.

Gubatz S, Shewry PR (2011). The development, structure and composition of the 134 barley grain. *In*: Ullrich SE (ed). *Barley: Production, improvement and uses* (pp. 391-438). West Sussex, UK: John Wiley and Sons.

Güler M (2003). Barley grain β-glucan content as affected by nitrogen and irrigation. Field Crop Research. 84:335-340.

Gutiérrez L, Cuesta-Marcos A, Castro AJ, von Zitzewitz J, Schmitt M, Hayes PM (2011). Association mapping of malting quality quantitative trait loci in winter barley: positive signals from small germplasm arrays. The Plant Genome Journal. 4:256-272.

Gymer PT (1978). The genetics of the six-row/two-row character. Barley Genetics Newsletter. 8:44-46.

Han F, Ullrich S, Chirat S, Menteur S, Jestin L, Sarrafi A, Hayes P, Jones B, Blake T, Wesenberg D, Kleinhofs A, Kilian A (1995). Mapping of beta-glucan content and beta-glucanase activity loci in barley-grain and malt. Theoretical and Applied Genetics. 91:921-927.

Hang A, Ohert D, Inez AG, Burton C (2007). Barley amylose and β-glucan: Their relationship to protein, agronomic traits, and environmental factors. Crop Science 47:1754-1760.

Hardy OJ, Vekemans X (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes. 2:618-620.

Harlan JR (1979). On the origin of barley. *In*: *Barley: Origin, Botany, Culture, Winter Hardiness, Genetics, Utilization, Pests. Agriculture Handbook No. 338* (pp. 10-36). Washington, DC: USDA.

Hashimoto S, Shogren MD, Bolte LC, Pomeranz Y (1987). Cereal pentosans: their estimation and significance. III. Pentosans in abraded grains and milling by-products. Cereal Chemistry. 64:39-41.

Hayes PM, Cerono J, Witsenhoer H, Kuiper M, Zabeau M, Sato K, Kleinhofs A, Kudrna D, Kilian A, Saghai-Maroof M, Hoffman D, and North American Barley Genome Mapping Project USA (1997). Characterizing and exploiting genetic diversity and quantitative traits in barley (hordeum vulgare) using AFLP markers. Journal of Agricultural Genomics. 3:1-18.

Heck GR, Chamberlain AK, Ho TH (1993). Barley embryo globulin 1 gene, Beg1 : characterisation of cDNA, chromosome mapping and regulation of expression. Molecular and General Genetics. 239: 209-218.

Hecker KD, Meier ML, Newman RK, Newman CW (1998). Barley β-glucan is effective as a hypocholesterolaemic ingredient in foods. Journal of Science Food and Agriculture. 77:179-183.

Henry RJ (1986). Genetic and environmental variation in the pentosan and β-glucan contents of barley, and their relation to malting quality. Journal of Cereal Science. 4:269-277.

Henry RJ (1987). Pentosan and (1→3),(1→4)-β-glucan concentrations in endosperm and whole grain of wheat, barley, oats and rye. Journal of Cereal Science. 6:253-258.

Henry RJ (1988). The carbohydrates of barley grains – a review. Journal of the Institute of Brewing. 94:71-78.

Hizukuri S (1985). Relationship between the distributions of the chain length of amylopectin and the crystalline structure of starch granules. Carbohydrate Research. 141:295-306.

Hoffmann RA, Roza M, Maat J, Kamerling JP, Vliegenthart JFG (1991). Structural characteristics of the cold-water-soluble arabinoxylans from the white flour of the soft wheat variety Kadet. Carbohydrate Polymers. 15:415-430.

Hogy P, Poll C, Marhan S, Kandeler E, Fangmeier A (2013). Impacts of temperature increase and change in precipitation pattern on crop yield and yield quality of barley. Food Chemistry. 136:1470-1477.

Holtekjølen AK, Baevre AB, Rødbotten M, Berg H, Knutsen SH (2008). Antioxidant properties and sensory profiles of breads containing barley flour. Food Chemistry. 110:414-421.

Holtekjølen AK, Uhlen AK, Bråthen E, Sahlstrøm S, Knutsen SH (2006). Contents of starch and non-starch polysaccharides in barley varieties of different origin. Food Chemistry. 94:348-358.

Houston K, Russell J, Schreiber M, Halpin C, Oakey H, Washington JM, Booth A, Shirley N, Burton RA, Fincher GB, Waugh R (2014). A genome wide association scan for (1,3;1,4)-β-glucan content in the grain of contemporary 2-row Spring and Winter barleys. BMC Genomics. 15:907. doi:10.1186/1471-2164-15-907

Hu G, Burton C, Hong Z, Jackson E (2014). A mutation of the cellulose-synthase-like (*CslF6*) gene in barley (Hordeum vulgare L.) partially affects the β-glucan content in grains. Journal of Cereal Science. 59:189-195.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009). Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources. 9:1322-1332.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009). Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources. 9:1322-1332.

Hung PV, Maeda T, Miskelly D, Tsumori R, Morita N (2008). Physicochemical characteristics and fine structure of high-amylose wheat starches isolated from Australian wheat cultivars. Carbohydrate Polymers. 71:656-663.

ICH harmonized tripartite guidelines (2005). Validation of analytical procedures: Text and methodology Q2(R1). <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/ Guidelines/Quality/Q2_R1/ Step4/Q2_R1__Guideline.pdf > pp. 1–17.

Igartua E, Hayes PM, Thomas WTB, Meyer R, Mather DE (2002). Genetic control of quantitative grain and malt quality traits in barley. Journal of Crop Production. 5:131-164.

Igartua E, Moralejo M, Maria-Casas A, Torres L, Molina-Cano J (2013). Whole-genome analysis with SNPs from BOPA1 shows clearly defined groupings of western Mediterranean, Ethiopian, and Fertile Crescent barleys. Genetic Resources and Crop Evolution. 60:251-264.

Inglett GE (1990). Hypocholesterolemic beta-glucan-amylodextrins from oats as a dietary fat-replacements. In: *Proceedings of Symposium on Beta-Glucans: Biotechnology and Nutrition, 199th National Meeting of the American Chemical Society,* Boston, MA.

Ingvarsson PK, Street NR (2011). Tansley review Association genetics of complex traits in plants. New Phytologist. 198: 909-922.

Irakli M, Biliaderis CG, Izydorczyk MS, Papadoyannis IN (2004). Isolation, structural features and rheological properties of water-extractable beta-glucans from different Greek barley cultivars. Journal of the Science of Food and Agriculture. 84:1170-1178.

Islamovic E, Obert DE, Oliver RE, Harrison SA, Ibrahim A, Marshall JM, Miclaus KJ, Hu G, Jackson EW (2013). Genetic dissection of grain beta-glucan and amylose content in barley (hordeum vulgare L.). Molecular Breeding. 31:15-25.

Izydorczyk MS, Storsley J, Labossiere D, MacGregor AW, Rossnagel BG (2000). Variation in total and soluble β-glucan content in hulless barley: Effects of thermal, physical, and enzymic treatments. Journal of Agricultural and Food Chemistry. 48:982-989.

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

Jadhav SJ, Lutz SE, Ghorpade VM, Salunkhe DK (1998). Barley: chemistry and value-added processing. Critical Reviews in Food Science and Nutrition. 38(2):123–171.

Jansen R, Stam P (1994). High resolution of quantitative traits into multiple loci via interval mapping. Genetics. 136:1447-1455.

Janthakahalli NV (2004). Regulation of fructan metabolism in barley leaves. PhD Thesis, Faculty of Science. University of Basel.

Jarves JK, Barkworth ME (1992). Morphological variations and genome constitution in some perennial Triticeae. Botanical Journal of the Linnean Society. 103:167-180.

Jende-Strid B (1985). Phenolic acids in grains of wild-type barley and proanthocyanidin-free mutants . Carlsberg Research Communications. 50:1-14.

Jiang G, Vasanthan T (2000). Effect of extrusion cooking on the primary structure and water solubility of beta-glucans from regular and waxy barley. Cereal Chemistry. 77:396-400.

Jui PY, Choo TM, Ho KM, Konishi T, Martin RA (1997).Genetic basis of a two-row by six-row cross of barley using doubled-haploid lines. Theoretical and Applied Genetics 94:549-556.

Kaczmarczyk A, Bowra S, Elek Z, Vincze E (2012). Quantitative RT–PCR based platform for rapid quantification of the transcripts of highly homologous multigene families and their members during grain development. BMC Plant Biology. 12:184.

Kaida R, Satoh Y, Bulone V, Yamada Y, Kaku T, Hayashi T, Kaneko TS (2009). Activation of β-glucan synthases by wall-bound Purple Acid Phosphatase in Tobacco Cells. Plant Physiology. 150:1822-1830.

Kakeda K, Ishihara N, Izumi Y, Sato K, Taketa S (2011). Expression and functional analysis of the barley Nud gene using transgenic rice. Breeding Science 61:35-42.

Kan YC, Kan YF, Beaudoin F, Leader DJ, Edwards K, Poole R, Wang DW, Mitchell RAC, Shewry PR (2006). Transcriptome analysis reveals differentially expressed storage protein transcripts in seeds of Aegilops and wheat. Journal of Cereal Science. 44:75-85.

Kandemir N, Kudrna DA, Ullrich SE, Kleinhofs A (2000). Molecular marker assisted genetic analysis of head shattering in six-rowed barley. Theoretical and Applied Genetics. 101(1-2):203-210.

Kato T, Sasaki A, Takeda G (1995). Genetic variation of β-glucan contents and β-glucanase activities in barley, and their relationship to malting quality. Breeding Science 45:471-477.

Kauppinen S (1992). Structure and expression of the Kas12 gene encoding a beta-ketoacyl-acyl carrier protein synthase-i isozyme from barley. Journal of Biological Chemistry. 267:23999-24006.

Kellogg EA (2001). Evolutionary history of grasses. Plant Physiology. 125:1198-1205.

Kersey PJ, Allen JE, Christensen M, Davis P, Falin LJ, Grabmueller C, Hughes DST, Humphrey J, Kerhornou A, Khobova J, Langridge N, McDowall MD, Maheswari U, Maslen G, Nuhn M, Kee Ong C, Paulini M, Pedro H, Toneva I, Tuli M, Walts B, Williams G, Wilson D,

Youens-Clark K, Monaco MK, Stein J, Wei X, Ware D, Bolser DM, Howe KL, Kulesha E, Lawson D, Staines DM (2014). Ensemble Genomes 2013: scaling up access to genome-wide data. Nucleic Acids Research. 42: D546-D552.

Kim H, Park K, Baek S, Kim J (2011). Inheritance of (1-3)(1-4)-beta-D-glucan content in barley (*Hordeum vulgare* L.). Journal of Crop Science and Biotechnology. 14:239-245.

Kim SJ, Zemelis S, Keegstra K, Brandizzi F (article in press) The catalytic site of CSFL6 supports a channeling model for the biosynthesis of mixed-linkage glucan. The Plant Journal. doi: 10.1111/tpj.12748

Kimpara T, Aohara T, Soga K, Wakabayashi K, Hoson T, Tsumuraya Y, Kotake T (2008). β-(1,3;1,4)-glucan synthase activity in rice seedlings under water. Annals of Botany. 102:221-226.

Kirby EJM, Riggs TJ (1978). Developmental consequences of two-row and six-row ear type in spring barley: 2. Shoot apex, leaf and tiller development. Journal of Agricultural Science (Cambridge). 91:207-216.

Kirkman MA, Shewry PR, Miflin BJ (1982). The effect of nitrogen nutrition on the lysine content and protein composition of barley seeds. Journal of the Science of Food and Agriculture. 33:115-127.

Kleinhofs A (1997). Integrating barley RLFP and classical marker maps. Barley Genetics Newsletter. 27:105-112.

Klopfenstein C, Hoseney R (1987). Cholesterol-lowering effect of beta-glucan-enriched bread. Nutrition Reports International. 36:1091-1098.

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A, Lundqvist U, Fujimura T, Matsuoka M, Matsumoto T, Yano M (2007). Six-rowed barley originated from a mutation in a homeodomain leucine zipper I-class homeobox gene. Proceedings of the National Academy of Sciences of the United States of America. 104:1424-1429.

Koppolu R, Anwar N, Sakuma S, Tagiri A, Lundqvist U, Pourkheirandish M, Rutten T, Seiler C, Himmelbach A, Ariyadasa R, Youssef HM, Stein N, Sreenivasulu N, Komatsuda T, Schnurbusch T (2013). Six-rowed spike4 (Vrs4) controls spikelet determinacy and row-type in barley. Proceedings of the National Academy of Sciences of the United States of America. 110:13198–203.

Kramer HH, Blander BA (1961). Orientating linkage maps on the chromosomes of barley. Crop Science. 1:339-342.

Kurek I, Kawagoe Y, Jacob-Wilk D, Doblin M, Delmer D (2002). Dimerization of cotton fiber cellulose synthase catalytic subunits occurs via oxidation of the zinc-binding domains.

Proceedings of the National Academy of Sciences of the United States of America. 99:11109-11114.

LaFrance JT, Watts MJ (1986). The value of protein in feed barley for beef, dairy, and swine feeding. Western Journal of Agricultural Economics. 11:76-81.

Laido G, Barabaschi D, Tondelli A, Gianinetti A, Stanca AM, Nicosia OLD, Di Fonzo N, Francia E, Pecchioni N (2009). QTL alleles from a winter feed type can improve malting quality in barley. Plant Breeding. 128:598-605.

Lander ES, Botstein D (1989). Mapping Medelian factors under- lying quantitative traits using RFLP linkage maps. Genetics. 121:185-199.

Langmead B, Salzberg S (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods. 9:357-359.

Lazaridou A, Biliaderis C, Izydorczyk M (2003). Molecular size effects on rheological properties of oat beta-glucans in solution and gels. Food Hydrocolloids. 17:693-712.

Lazaridou A, Biliaderis C, Micha-Screttas M, Steele B (2004). A comparative study on structure-function relations of mixed-linkage (1-3),(1-4) linear beta-D-glucans. Food Hydrocolloids. 18:837-855.

Lazaridou A, Biliaderis CG (2007). Molecular aspects of cereal β-glucan functionality: Physical properties, technological applications and physiological effects. Journal of Cereal Science. 46:101-118.

Levinson G, Gutman G (1987). Slipped-strand mispairing - a major mechanism for DNA-sequence evolution. Molecular Biology and Evolution. 4:203-221.

Li J, Båga M, Rossnagel BG, Legge WG, Chibbar RN (2008). Identification of quantitative trait loci for beta-glucan concentration in barley grain. Journal of Cereal Science. 48:647-655.

Li Z, Li D, Du X, Wang H, Larroque O, Jenkins CLD, Jobling SA, Morell MK (2011). The barley amo1 locus is tightly linked to the starch synthase IIIa gene and negatively regulates expression of granule-bound starch synthetic genes. Journal of Experimental Botany. 62:5217-5231.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25: 2078-2079.

Litt M, Luty JA (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetics. 44:397.

Liu B (1998). *Statistical genomics: Linkage, mapping and QTL analysis*. Boca Raton, Florida: CRC Press.

Liu DJ, Pomeranz Y, Robbins GS (1975). Mineral content of developing and malted barley. Cereal Chemistry. 52:678-686.)

Liu KS, Peterson KL, Raboy V (2007). Comparison of the phosphorus and mineral concentrations in bran and abraded kernel fractions of a normal barley (*Hordeum vulgare*) cultivar versus four low phytic acid isolines . Journal of Agricultural and Food Chemistry. 55:4453-4460 .

Liu RH (2004). New finding may be key to ending confusion over link between fiber, colon cancer. American Institute for Cancer Research Press Release. November 3, 2004.

Lott JNA, Ockenden I, Raboy V, Batten GD (2000). Phytic acid and phosphorus in crop seeds and fruits: a global estimate. Seed Science Research. 10:11-33.

Lundqvist U, Franckowiak JD, Konishi T (1997). New and revised descriptions of barley genes. Barley Genetics Newsletter. 26:22-516.

Lundqvist U, Lundqvist A (1987). An intermedium gene present in a commercial six-row variety of barley. Hereditas. 107:131-135.

MacGregor AW, Fincher GF (1993). Carbohydrates of the barley grain. *In*: MacGregor AW, Bhatty RS (eds). *Barley: Chemistry and Technology* (pp. 73-130). St. Paul, MN: American Association of Cereal Chemists.

MacLeod AM (1953). Studies on the free sugars of the barley grain. IV. Low-molecular fructosans. Journal of the Institute of Brewing. 59:462-469.

Mammadov J, Aggarwal R, Buyyarapu R, Kumpatla S (2012). SNP markers and their impact on plant breeding. International Journal of Plant Genomics. Doi:10.1155/2012/72839.

Marconi E, Graziano M, Cubadda R (2000). Composition and utilization of barley pearling by-products for making functional pastas rich in dietary fiber and β-glucans. Cereal Chemistry. 77:133-139.

Marquez-Cedillo LA, Hayes PM, Kleinhofs A, Legg, WG, Rossnagel BG, Sato K, Ullrich SE, Wesenberg DM (2001). QTL analysis of agronomic traits in barley based on the doubled haploid progeny of two elite North American varieties representing different germplasm groups. Theoretical and Applied Genetics. 103:625–637.

Mascher M, Wu S, St. Amand P, Stein N, Poland J (2013). Application of genotyping-by-sequencing on semiconductor sequencing platforms: A comparison of genetic and reference-based marker ordering in barley. PlosOne. 8:e76925.

Mather D, Tinker N, LaBerge D, Edney M, Jones B, Rossnagel B, Legge W, Briggs K, Irvine R, Falk D, Kasha K (1997). Regions of the genome that affect grain and malt quality in a North American two-row barley cross. Crop Science. 37:544-554.

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

McCleary B, Codd R (1991). Measurement of (1→3),(1→4)-beta-D-glucan in barley and oats - a streamlined enzymatic procedure. Journal of the Science of Food and Agriculture. 55:303-312.

McGuire CF, Hockett EA (1981). Effect of awn length and naked on malting quality of Betzes barley. Crop Science. 21:18-21.

McIntosh G, Leleu R, Kerry A, Goldring M (1993). Barley-grain for human food use. Food Australia. 45:392-394.

Meikle PJ, Hoogenraad NJ, Bonig I, Clarke AE, Stone BA (1994). A (1→3,1→4)-β-glucan-specific monoclonal antibody and its use in the quantitation and immunocytochemical location of (1→3,1→4)-β-glucans. The Plant Journal. 5:1-9.

Meikle PJ, Ng KF, Johnson E, Hoogenraad NJ, Stone BA (1991). The β-glucan synthase from Lolium multiflorum: Detergent solubilization, purification using monoclonal antibodies and pho- toaffinity labeling with a novel photoreactive pyrimidine analog of uridine 5′-diphosphoglucose. Journal of Biological Chemistry. 266:22569-22581.

Mezaka I, Bleidere M, Legzdina L, Rostoks N (2011). Whole genome association mapping identifies naked grain locus NUD as determinant of beta-glucan content in barley. Zemdirbyste-Agriculture. 98:283-292.

Mikkelsen MS, Jespersen BM, Larsen FH, Blennow A, Engelsen SB (2013). Molecular structure of large-scale extracted beta-glucan from barley and oat: Identification of a significantly changed block structure in a high beta-glucan barley mutant. Food Chemistry. 136:130-138.

Mohammadi M, Endelman JB, Nair S, Chao S, Jones SS, Muehlbauer GJ, Ullrich SE, Baik B, Wise ML, Smith KP (2014). Association mapping of grain hardness, polyphenol oxidase, total phenolics, amylose content, and β-glucan in US barley breeding germplasm. Molecular Breeding. 34:1229-1243.

Molina-Cano JL, Ramo T, Ellis R, Swanston J, Bain H, Uribeechertia T, Perex-Vendrell A (1995). Effect of grain composition on water-uptake by malting barley – a genetic and environmental-study. Journal of the Institute of Brewing. 101:79-83.

Molina-Cano JL, Moralejo M, Elia M, Munoz P, Russell JR, Perez-Vendrell AM, Ciudad F, Swanston JS (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.

Moralejo M, Swanston J, Munoz P, Prada D, Elia M, Russell J, Ramsay L, Cistue L, Codesal P, Casas A, Romagosa I, Powell W, Molina-Cano J (2004). Use of new EST markers to elucidate the genetic differences in grain protein content between European and North American two-rowed malting barleys. Theoretical and Applied Genetics. 110:116-125.

Moreau RA, Flores RA, Hicks KB (2007). Composition of functional lipids in hulled and hulless barley infractions obtained by scarification and in barley oil. Cereal Chemistry. 84:1-5.

Morgan A, Riggs T (1981). Effects of drought on yield and on grain and malt characters in spring barley. Journal of the Science of Food and Agriculture. 32:339-346.

Morrison WR (1993). Barley lipids. *In*: MacGregor AW, Bhatty RS (eds). *Barley: Chemistry and Technology* (pp.199-246). St. Paul, MN: American Association of Cereal Chemists.

Munck L, Møller B, Jacobsen S, Søndergaard I (2004). Near infrared spectra indicate specific mutant endosperm genes and reveal a new mechanism for substituting starch with $(1\rightarrow3,1\rightarrow4)$-β-glucan in barley. Journal of Cereal Science. 40:213-222.

Nagashima H, Ishikawa H (1995). Genotypic variation of hordein in Japanese barley cultivars. Breeding Science. 45:87-90.

Nemeth C, Freeman J, Jones HD, Sparks C, Pellny TK, Wilkinson MD, Dunwell J, Andersson AAM, Åman P, Guillon F, Saulnier L, Mitchell RAC, Shewry PR (2010). Down-regulation of the *CslF6* gene results in decreased (1,3;1,4)-beta-D-glucan in endosperm of wheat. Plant Physiology. 152:1209-1218.

Neukom H, Markwalder H (1975). Isolation and characterization of an arabinogalactan from wheat flour. Carbohydrate Research. 39:387-389.

Nevo E (1992). Origin, evolution, population genetics and resources for breeding of wild barley, *Hordeum spontaneum,* in the Fertile Crescent. *In*: Shewry PR (ed.). *Barley: Genetics, Molecular Biology and Biotechnology* (pp. 19-43). Wallingford, UK: CAB International.

Newman CW, McGuire CF (1985). Nutritional quality of barley. *In*: Rasmusson CD (ed.). *Barley* (pp. 403-456). Madison, Wisconsin: American Society of Agronomy.

Newman CW, Newman RK (1992). Nutritional aspects of barley seed structure and composition. *In*: Shewry PR (ed.). *Barley: Genetics, Molecular Biology and Biotechnology* (pp. 351-368). Wallingford, UK: CAB International.

Newman CW, Newman RK (2006). A brief history of barley foods. Cereal Foods World. 51:4-71.

Newman CW, Newman RK (2008). *Barley for food and health: Science, technology and products*. Hoboken, NJ: Wiley.

Nordborg M, Weigel D (2008). Next-generation genetics in plants. Nature. 456:720-723.

Nühse TS, Stensballe A, Jensen ON, Peck SC (2004). Phosphoproteomics of the Arabidopsis plasma membrane and a new phosphorylation site database. Plant Cell. 16:2394-2405.

Ockenden I, Dorsch JA, Reid MM, Lin L, Grant LK, Raboy V, Lott JNA (2004). Characterization of the storage of phosphorus, inositol phosphate and cations in grain tissues of four barley (*Hordeum vulgare* L.) low phytic acid genotypes. Plant Science. 167:1131-1142.

Oikawa A, Nagai K, Kato K, Kidou S (2009). Gene silencing of barley P23k involved in secondary wall formation causes abnormal tiller formation and intercalary elongation. Breeding Science. 59:664.

Oikawa A, Rahman A, Yamashita T, Taira H, Kidou S (2007).Virus-induced gene silencing of P23kin barley leaf reveals morphological changes involved in secondary wall formation. Journal of Experimental Botany. 58:2617-2625.

Oraguzie NC, Rikkerink EHA, Gardiner SE, de Silva HN (2007). *Association Mapping in Plants*. New York, NY: Springer.

Oziel A, Hayes P, Chen F, Jones B (1996). Application of quantitative trait locus mapping to the development of winter-habit malting barley. Plant Breeding. 115:43-51.

Paliyath G, Poovaiah BW (1988). Promotion of β-glucan synthesis activity in corn microsomal membranes by calcium and protein phosphorylation. Plant Cell Physiology. 29:67-73.

Panozzo JF, Eckermann PJ, Mather DE, Moody DB, Black CK, Collins HM, Barr AR, Lim P, Cullis BR (2007). QTL analysis of malting quality traits in two barley populations. Australian Journal of Agricultural Research. 58:858-866.

Pastuglia M, Roby D, Dumas C, Cock JM (1997). Rapid induction by wounding and bacterial infection of an S gene family receptor-like kinase gene in Brassica oleracea. The Plant Cell. 9:49-60.

Pastuglia M, Swarup R, Rocher A, Saindrenan P, Roby D, Dumas C, Cock JM (2002). Comparison of the expression patterns of two small gene families of S gene family receptor kinase genes during the defence response inBrassica oleracea and Arabidopsis thaliana. Gene. 282: 215-225.

Paterson AH (2002). What has QTL mapping taught us about plant domestication? New Phytology. 154:591-608.

Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. 28: 2537-2539.

Pereira M, O'Reilly E, Augustsson K, Fraser G, Goldbourt U, Heitmann B, Hallmans G, KnektP, Liu S, Pietinen P, Spiegelman D, Stevens J, Virtamo J, Willett W, Ascherio A (2004). Dietary fiber and risk of coronary heart disease − A pooled analysis of cohort studies. Archives of Internal Medicine. 164:370-376.

Perez-Vendrell A, Brufau J, Molina-Cano J, Francesch M, Guasch J (1996). Effects of cultivar and environment on beta-(1,3)-(1,4)-D-glucan content and acid extract viscosity of Spanish barleys. Journal of Cereal Science. 23:285-292.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012). Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PloS ONE. 7:e37135.

Peterson DM (1994). Barley tocols: effects of milling, malting and mashing. Cereal Chemistry. 71:42-44.

Pilling E, Hofte H (2003). Feedback from the wall. Current Opinion in Plant Biology. 6:611-616.

Poland JA, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, Dreisigacker S, Crossa J, Sanchez-Villeda H, Sorrells M, Jannink J (2012b). Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genenetics. 5:103-113.

Poland JA, Brown PJ, Sorrells ME, Jannink J (2012a). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLos ONE 7:e32253.

Poppitt SD, van Drunen JDE, McGill AT, Mulvey TB, Leahy FE (2007). Supplementation of a high-carbohydrate breakfast with barley β-glucan improves postprandial glycaemic response for meals but not beverages. Asia Pacific Journal of Clinical Nutrition 16, 16-24.

Powell W, Caligari P, Swanston J, Jinks J (1985). Genetic investigations into beta-glucan content in barley. Theoretical and Applied Genetics. 71:461-466.

Powell W, Ellis RP, Thomas WTB (1990). The effects of major genes on qualitatively varying characters in barley. III. The two row/six row locus (V-v). Heredity 65: 259-264.

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

Rallabhandi P (2012). Gluten and celiac disease – an immunological perspective. Journal of AOAC International. 95:349-355.

Ramsay L, Comadran J, Druka A, Marshall DF, Thomas WT, Macaulay M, MacKenzie K, Simpson C, Fuller J, Bonar N, Hayes PM, Lundqvist U, Frankowiak JD, Close TJ, Muehlbauer GJ, Waugh R (2011). INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1. Nature Genetics. 43:169-172.

Rastogi V, Oaks A (1986). Hydrolysis of storage proteins in barley endosperm. Plant Physiology. 81: 901-906.

Ray PM (1973). Regulation of β-glucan synthetase activity by auxin in pea stem tissue. II. Metabolic requirements. Plant Physiology. 51:609-614.

Richard I, Beckman JS (1995). How neutral are synonymous codon mutations: Nature Genetics. 10:259.

Richmond TA, Somerville CR (2000). The cellulose synthase superfamily. Plant Physiology. 124:495-498.

Risch N, Merikangas K (1996). The future of genetic studies of complex human diseases. Science. 273:1516-1517.

Ritland K (1996). Estimators for pairwise relatedness and individual inbreeding coefficients. Genetical Research. 67:175-185.

Robinson HF, Comstock RE, Harvey VH (1949). Estimates of heritability and degree of dominance in corn. Agronomy Journal. 41:353-359.

Rohde W, Becker D, Salamini F (1988). Structural analysis of the waxy locus from Hordeum vulgare. Nucleic Acids Research. 16:1785-1786.

Ross AB, Shepherd MJ, Schüpphaus M, Sinclair V, Alfaro B, Kamal-Eldin A, Åman P (2003). Alkylresorcinols in cereals and cereal products. Journal of Agricultural and Food Chemistry. 51:4111-4118.

Rozen S, Skaletsky HJ (2000). Primer3 on the WWW for general users and for biologist programmers. *In*: Misener S, Krawetz S (eds.). *Methods in Molecular Biology. Bioinformatics Methods and Protocols, vol. 132* (pp. 365-386). Totowa, NJ, USA: Humana Press Inc.

Schlotterer C (2000). Evolutionary dynamics of microsatellite DNA. Chromosoma. 109:365-371.

Schreiber M, Wright F, MacKenzie K, Hedley PE, Schwerdt JG, Little A, Burton R, Fincher GB, Marshall D, Waugh R, Halpin C (2014). The barley genome sequence assembly reveals

three additional members of the *CslF* (1,3;1,4)-β-glucan synthase gene family. PloS one. 9:e90888.

Sethaphong L, Haigler CH, Kubicki JD, Zimmer J, Bonetta D, DeBolt S, Yingling YG (2013). Tertiary model of a plant cellulose synthase. Proc Natl Acad Sci USA.110: 7512-7517.

Shewry PR (1993). Barley seeds proteins. *In*: MacGregor AW, Bhatty RS (eds). *Barley: Chemistry and Technology* (pp. 1-5). St. Paul, MN: American Association of Cereal Chemists.

Shu X, Rasmussen SK (2014). Quantification of amylose, amylopectin, and β-glucan in search for genes controlling the three major quality traits in barley by genome-wide association studies. Frontiers in Plant Science. 5:197. doi: 10.3389/fpls.2014.00197

Simmonds DH (1978). Structure, composition and biochemistry of cereal grains. *In*: Pomeranz Y (ed.), *Cereals '78: Better Nutrition for the World's Millions* (p.105). St. Paul, MN: American Association of Cereal Chemists.

Singh M, Ceccarelli S, Hamblin J (1993). Estimation of heritability from varietal trials data. Theoretical and Applied Genetics. 86:437-441.

Sneller CH, Mather DE, Crepieux S (2009). Analytical approaches and population types for finding and utilizing QTL in complex plant populations. Crop Science. 49:363-380.

Sollid LM (2000). Molecular basis of celiac disease. Annual Review of Immunology. 18:53-81.

Sopanen T, Lauriere C (1989). Release and activity of bound, β-amylase in a germinating barley grain. Plant Physiology. 89:244-249.

Sørensen I, Pettolino FA, Wilson SM, Doblin MS, Johansen B, Bacic A, Willats WGT (2008). Mixed-linkage (1→3),(1→4)-β-D-glucan is not unique to the poales and is an abundant component of equisetum arvense cell walls. Plant Journal. 54:510-521.

Sreenivasulu N, Usadel B, Winter A, Radchuk V, Scholz U, Stein N, Weschke W, Strickert M, Close TJ, Stitt M, GranerA, Wobus U (2008). Barley grain maturation and germination: Metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. Plant Physiology. 146:1738-1758.

Statistics Canada (2014), http://www.statcan.gc.ca/pub/22-002-x/2012006/t059-eng.htm

Storsley J, Izydorczyk M, You S, Biliaderis C, Rossnagel B (2003). Structure and physicochemical properties of beta-glucans and arabinoxylans isolated from hull-less barley. Food Hydrocolloids. 17:831-844.

Sunyaev S, Hanke J, Aydin A, Wirkner U, Zastrow I, Reich J, Bork P (1999). Prediction of nonsynonymous single nucleotide polymorphisms in human disease associated genes. Journal of Molecular Medicine. 77:754-760.

Swanston J, Ellis R, Stark J (1995). Effects on grain and malting quality of genes altering barley starch composition. Journal of Cereal Science. 22:265-273.

Syvanen AC (2001). Accessing genetic variation: Genotyping single nucleotide polymorphisms. Nature Reviews Genetics. 2:930-942.

Tabata M (1961). Studies of a gametophyte factor in barley. Japan Journal of Genetics 36:157-167.

Takeda S, Mano S, Ohto M, Nakamura K (1994). Inhibitors of protein phosphatases I and 2A block the sugar inducible gene expression in plants. Plant Physiology. 106:567-574.

Takeda Y, Shitaozono T, Hizukuri S (1990). Structures of subfractions of corn amylose. Carbohydrate Research. 199:207-214.

Taketa S, Amano S, Tsujino Y, Sato T, Saisho D, Kakeda K, Nomura M, Suzuki T, Matsumoto T, Sato K, Kanamori H, Kawasaki S, Takeda K (2008). Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. Proceedings of the National Academy of Sciences of the United States of America. 105:4062-4067.

Taketa S, Kikuchi S, Awayama T, Yamamoto S, Ichii M, Kawasaki S (2004). Monophyletic origin of naked barley inferred from molecular analyses of a marker closely linked to the naked caryopsis gene (nud). Theoretical and Applied Genetics. 108:1236-1242.

Taketa S, Yuo T, Tonooka T, Tsumuraya Y, Inagaki Y, Haruyama N, Larroque O, Jobling SA (2012). Functional characterization of barley betaglucanless mutants demonstrates a unique role for *CslF6* in (1,3;1,4)-beta-D-glucan biosynthesis. Journal of Experimental Botany. 63:381-392.

Tanksley SD (1993). Mapping polygenes. Annual Review of Genetics. 27:205-233.

Tanno K, Takeda K (2004). On the origin of six-rowed barley with brittle rachis, agriocrithon (*Hordeum vulgare* ssp. f. *agriocrithon* (Åberg) Bowd.), based on a DNA marker closely linked to the *vrs1* (six-row gene) locus. Theoretical and Applied Genetics. 110:145-150.

Tanno K, Willcox G (2006). How fast was wild wheat domesticated? Science 311:1886.

Taylor NG (2007). Identification of cellulose synthase AtCesA7 (IRX3) in vivo phosphorylation sites. A potential role in regulating protein degradation. Plant Molecular Biology. 64:161–171.

Tester RF, Morrison WR (1992). Swelling and gelatinization of cereal starches. 3. Some properties of waxy and normal nonwaxy barley starches. Cereal Chemistry. 69:654-658.

Tetlow IJ, EmesMJ (2014). A review of starch-branching enzymes and their role in amylopectin biosynthesis. IUBMB Life. 66:546-558.

Thomas, W (2003). Prospects for molecular breeding of barley. Annals of Applied Biology. 142:1-12.

Timmermans JW, van Leeuwen MB, Tournois H, de Wit D, Vliegenthart JFG (1994). Quantification analysis of the molecular weight distribution of inulin by means of anion exchange HPLC with pulsed amperiometric detection. Journal of Carbohydrate Chemistry. 13:881-888.

Tohno-Oka T, Nawada N, Yoshioka T (2004). Relationship between grain hardness and endosperm cell wall polysaccharides in barley. *In*: *Proceedings of the Ninth International Barley Genetic Symposium*, Brno, Czech Republic, pp. 595-600.

Tondelli A, Xu X, Moragues M, Schnaithmann F, Sharma R, Ingvardsen C, Manninen O, Comadran J, Russel J, Waugh R, Schulman AH, Pillen K, Rasmussen SK, Kilian B, Cattivelli L, Thomas WTB, Flavel A (2013). Structural and temporal variation in the genetic diversity of a European collection of spring 2-row barley cultivars and utility for association mapping of quantitative traits. Plant Genome. 6:1-14.

Tonooka T, Aoki E, Yoshioka T, Taketa S (2009). A novel mutant gene for (1-3,1-4)-beta-D-glucanless grain on barley (*Hordeum vulgare L.*) chromosome 7H. Breeding Science. 59:47-54.

Trethewey JAK, Campbell LM, Harris PJ (2005). (1,3;1,4)-β-glucans in the cell walls of the Poales (sensu lato): an immunogold labeling study using a monoclonal antibody. American Journal of Botany. 92:1660-1674.

Tsai CJ, Leitzmann MF, Willett WC, Giovannucci EL (2004). Long-term intake of dietary fiber and decreased risk of cholecystectomy in women. American Journal of Gastroenterology. 99:1364-70.

Tsuchiya K, Urahara T, Konishi T, Kotake T, Tohno-oka T, Komae K, Kawada N, Tsumuraya Y (2005). Biosynthesis of (1→3),(1→4)-β-glucan in developing endosperms of barley (*Hordeum vulgare*). Physiologica Plantarum. 125:181-191.

Ullrich SE (2002). Genetics and breeding of barley feed quality. *In*: Slafer, GA, Molina-Cano, JL, Araus JL, Romagosa I (eds.). *Barley Science: Recent advances from molecular biology to agronomy of yield and quality* (pp. 115-142). Binghamton, NY: Haworth Press.

Ullrich SE, Clancy JA, Eslick RF, Lance RCM (1986). Beta-glucan content and viscosity of extracts from waxy barley. Journal of Cereal Science. 4:279-285.

Urbanowicz B, Rayon C, Buckeridge M, Carpita N (2004). Topology of the Maize Mixed Linkage (1→3),(1→4)-β-D-Glucan Synthase at the Golgi Membrane. Plant Physiology. 134:758-768.

van Gool D, Vernon L (2006). Potential impacts of climate change on agricultural land use suitability: barley. Report No. 302. Department of Agriculture, Government of Western Australia, Perth.

van Ooijen JW (2004). MapQTL 5: Software for the Mapping of Quantitative Trait Loci in Experimental Populations. Kyazma BV, Wageningen, The Netherlands.

van Ooijen JW, Voorrips RE (2001). JoinMap3.0: Software for the Calculation of Genetic Linkage Maps. Plant Research International, Wageningen, The Netherlands.

Vergara CE, Carpita NC (2001). Beta-D-glycan synthases and the CesA gene family: Lessons to be learned from the mixed-linkage (1→3),(1→4)beta-D-glucan synthase. Plant Molecular Biology. 47:145-160.

Vinkx CJA, Delcour JA (1996). Rye (*Secale cereale* L.) arabinoxylans: a critical review. Journal of Cereal Science. 24:1-14.

von Bothmer R, Komatsuda T (2011). Barley Origin and Related Species. *In*: Ullrich SE (ed.). *Barley: Production, Improvement, and Uses* (pp. 14-62). West Sussex, UK: Wiley-Blackwell.

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

Wang J, Zhang G, Chen J, Wu F (2004). The changes of beta-glucan content and beta-glucanase activity in barley before and after malting and their relationships to malt qualities. Food Chemistry. 86:223-228.

Wang J, Howles PA, Cork AH, Birch RJ, Williamson RE (2006). Chimeric proteins suggest that the catalytic and/or C-terminal domains give CesA1 and CesA3 access to their specific sites in the cellulose synthase of primary walls. Plant Physiology. 142: 685-695.

Washington JM, Box A, Barr AR (2000). Developing waxy barley cultivars for food, feed and malt. *In*: Logue S (ed.). *Barley Genetics VIII* (pp. 303-306). 8th International Barley Genetics Symposium, Adelaide, Australia. Univ. of Adelaide, SA, Australia.

Wei K, Xue D, Huang Y, Jin X, Wu F, Zhang G (2009). Genetic mapping of quantitative trait loci associated with beta-amylase and limit dextrinase activities and beta-glucan and protein fraction contents in barley. Journal of Zhejiang University – Science B. 10:839-846.

Weiss, KM, and Clark, AG. (2002). Linkage disequilibrium and the mapping of complex human traits. TRENDS in Genetics, 18: 19-24.

Wenzl, P, Li H, Carling J, Zhou M, Raman H, Paul E, Hearnden P, Maier C, Xia L, Caig V, Ovesna J, Cakir M, Poulsen D, Wang J, Raman R, Smith KP, Muehlbauer GJ, Chalmers KJ, Kleinhofs A, Huttner E, Kilian A (2006). A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BioMed Central Genomics. 7:206-228.

Wilson SM, Burton RA, Collins HM, Doblin MS, Pettolino FA, Shirley N, Fincher GB, Bacic A (2012). Pattern of deposition of cell wall polysaccharides and transcript abundance of related cell wall synthesis genes during differentiation in barley endosperm. Plant Physiology. 159:655-670.

Wilson SM, Burton RA, Doblin MS, Stone BA, Newbigin EJ, Fincher GB, Bacic A (2006). Temporal and spatial appearance of wall polysaccharides during cellularization of barley (Hordeum vulgare) endosperm. Planta. 224:655-667.

Wolfe SA, Nekludova L, Pabo CO (2000). DNA recognition by Cys2His2 zinc finger proteins. Annual Review of Biophysics and Biomolecular Structure. 29:183-212.

Wood P, Weisz J, Beer M, Newman C, Newman R (2003). Structure of $(1{\rightarrow}3)(1{\rightarrow}4)$-β-D-glucan in waxy and nonwaxy barley. Cereal Chemistry. 80:329-332.

Wood P, Weisz J, Blackwell BA (1994). Structural studies of (1-3),(1-4)-β-d-glucans by 13 C-nuclear magnetic resonance spectroscopy and by rapid analysis of cellulose-like regions using high-performance anion-exchange chromatography of oligosaccharides released by lichenase. Cereal Chemistry. 71:301-307.

Wood PJ (1986). Oat b-glucan: structure, properties and health claims. *In*: Webster FH (ed.). *Oats: Chemistry and Technology* (pp. 121-152). St. Paul, MN: American Association of Cereal Chemists.

Wood PJ (2007). Cereal β-glucans in diet and health. Journal of Cereal Science. 46:230-238.

Wood PJ (2010). Oat and rye beta-glucan: Properties and function. Cereal Chemistry. 87:315-330.

Wood PJ, Beer MU (1998). Functional oat products. In: Mazza G (ed.). Functional Foods. Biochemical & Processing Aspects (pp. 1-38). Lancaster: Technomic Publishing Co., Inc.

Wood PJ, Fulcher RG (1983). Dye interactions. A basis for specific detection and histochemistry of polysaccharides. Journal of Histochemistry and Cytochemistry. 31: 823-826.

Xue Q, Wang L, Newman RK, Newman CW, 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.

Yong W, Link B, O'Malley R, Tewari J, Hunter C, Lu C, Li X, Bleecker A, Koch K, McCann M, McCarty D, Patterson S, Reiter W, Staiger C, Thomas S, Vermerris W, Carpita N (2005). Genomics of plant cell wall biogenesis. Planta. 221:747-751.

Young ND (2001). Constructing a plant genetic linkage map with DNA Markers. *In*: Phillips RL, Vasil IK (eds.). *DNA-Based Markers in Plants* (pp. 31-47). Dordrecht: Kluwer Academic Publishers.

Zeng ZB (1994). Precision mapping of quantitative trait loci. Genetics. 136:1457-1468.

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marloram P, Nordborg M (2007). An Arabidopsis example of association mapping in structured samples. PLoS Genet. 3:1-4.

Zohary D, Weiss E, Hopf M (2012). *Domestication of Plants in the Old World*. 4th edition. Oxford, UK: Oxford University Press.

Zuur AF, Leno EN, Smith GM (2007). *Statistics for Biology and Health - Analysing Ecological Data*. New York, NY: Springer.

Figure S1. Alignment of nucleotide sequence for *HvCslF6* from barley CDC Bold and TR251. Start codon is highlighted in red.

```
CDCBold   NNCGTAAAGGAGAGTGAGTGCGTGCATTGAGGACGACGGCCATGGCGCCAGCGGTGGCCG 60
TR251     ************************************************************ 60

CDCBold   GAGGGGGCCGCGTGCGGAGCAATGAGCCGGTTGCTGCTGCTGCCGCCGCGCCGGCGGCCA 120
TR251     ************************************************************ 120

CDCBold   GCGGCAAGCCCTGCGTGTGCGGCTTCCAGGTTTGCGCCTGCACGGGGTCGGCCGCGGTGG 180
TR251     ************************************************************ 180

CDCBold   CCTCCGCCGCCTCGTCGCTGGACATGGACATCGTGGCCATGGGGCAGATCGGCGCCGTCA 240
TR251     **********************************A************************* 240

CDCBold   ACGACGAGAGCTGGGTGGGCGTGGAGCTCGGCGAAGATGGCGAGACCGACGAAAGCGGTG 300
TR251     ************************************************************ 300

CDCBold   CCGCCGTTGACGACCGCCCCGTATTCCGCACCGAGAAGATCAAGGGTGTCCTCCTCCACC 360
TR251     ************************************************************ 360

CDCBold   CCTACCGGTACGTCCTGCTCCCACAACTAAACAGAAACTCCCTATATCTGCGTCACACTC 420
TR251     ************************************************************ 420

CDCBold   AACAATTAATCCAACTAAGTCTCTCTACTACTCTAGTATTTATTTTTACTCTCTATCTGC 480
TR251     *********************************A************************** 480

CDCBold   ACAACAAGCGCTACTACAATTAACCCAACAAGCACCACGCCAGGTTGACAGTCAGGATAA 540
TR251     ************************************************************ 540

CDCBold   TTTGATCTTGACCGGAGTAAGTACTAGTACTAGGTCGGTGTTAATCAGAGTAATTATTGC 600
TR251     ************************************************************ 600

CDCBold   ACTAGTTAATTAAAATTTGAGTAATCCGAGACAGGTGCACGTTAGGGCCGGGCCAATGAT 660
TR251     ************************************************************ 660

CDCBold   GGCTCGAATCCACCCAAAATAGCGCGTCCCGGTGTGGGCTGTCGGCTCGGTGCTTCTTCC 720
TR251     C*********************************************************** 720

CDCBold   TTCCATTTTACTAGTCGCAGTCACTGCAGCTTGGGCCCACGGGAGGGGACGTTAGCCGTT 780
TR251     ************************************************************ 780

CDCBold   GGGCCTGCCTGGCAGGTGGGCCCCGGTGGCCACCCTGGCGGCTCATAAATCCTTGCTACT 840
TR251     ************************************************************ 840

CDCBold   TTGGAGCTGTAATGGACGCTCTGCAATAGCAATAGGAATCCGAGGTGAAACGACGACAGT 900
TR251     ************************************************************ 900

CDCBold   GGGCATGGCATGGCTTGCATGTGAATCCAAGCCACATCATTAAAAGCATCCTCCCTGGGC 960
TR251     ************************************************************ 960

CDCBold   ACGTCGCGGTGAGAAAGTTGGATAAACTTTTGGGGGTTCGGACAAGATGAGAAAAAGCAA 1020
TR251     ************************************************************ 1020

CDCBold   GTAACATGTCCCTTTTTTGGCACCGAAGAAATCCTATGTACGGCGAGTTTTTTCTGCATC 1080
TR251     ********G*************************************************** 1080

CDCBold   TAGTTATGGGTAGATGTACGTTAGTTTTTGTTGAGCGTTTTATGTGCTACATATATGGAG 1140
TR251     ************************************************************ 1140
```

```
CDCBold    AAAAAGAGAAAAATATTATCATGTCATGTCATGCCATGCCATGAGAAGGAGGAGAAGAAA 1200
TR251      ****************************************--------------******** 1186

CDCBold    AAGAAGTGTTGCCACCGCTCGAATGCCTTTTTTTTTCTTTCGGAAGGATGCGTGAGTCATG 1260
TR251      *******************************C**************************** 1246

CDCBold    TTGGCACCGAGAAAAGCCATATTAAAGTGGCAGAGTTACAACTCAGAATAAATGCGGGTG 1320
TR251      ************************************************************ 1306

CDCBold    TTACAAAAACACTAGGATATGTGAAGGGCACTCGGCACAACACTTTAAGACTACACAATT 1380
TR251      ************************************************************ 1366

CDCBold    GAAAAAGTGTACTCTCTGTATCTAAATAATTATAATTGAAAAGAATTAATCTATATATGA 1440
TR251      ************************************************************ 1426

CDCBold    AAGTAGTAATGATTAGCGGTAGAAGGTTCCAACGACTTTTTTGGCGCCAATAGCAAGAAG 1500
TR251      ************************************************************ 1486

CDCBold    AAAGAAAAAGAAAAAAATCTTTACTGTACAGTATAACGAGAAAAGAGGCCCATTAATAGA 1560
TR251      ******************************--*************************** 1544

CDCBold    GCAACGAATCGAGCGGCACCACCTCTGGCGGTCACGTCCATGCCCTCGCACGACGGATGA 1620
TR251      ***********C************************************************ 1604

CDCBold    GGCCCGGGGGGTCCTACTGACAGCCGAAGCATGTCGGTGCTCAAACACGGCGCCGTTTGC 1680
TR251      ************************************************************ 1664

CDCBold    TGCCAAGTGTGCCAGCTCGCACTCATTGACTTGCCAGCTCTCTCCTTGGTTGTCAATGAG 1740
TR251      ************************************************************ 1724

CDCBold    AACATGATGCCTTTTGGCATTTGCAAACTTATTAAAACTAGCTGTCGTCCGATAGGGAAA 1800
TR251      ************************************************************ 1784

CDCBold    AGAAAAGAAAAGAAAAGAATAAGAAAAAAAAGGACAAAGAGAAAAGATGAACATGGCGCAT 1860
TR251      ************************************************************ 1844

CDCBold    GTTCCCTCCAATAATTGCAGGCACCAACACTGGGTCGATTAATCCAACAACAATATTTTA 1920
TR251      ************************************************************ 1904

CDCBold    CTATACCAGACGAGAGTACAGTAGTCGGGTGATGATGGACTGTAACTGACTGAGTATGAA 1980
TR251      ************************************************************ 1964

CDCBold    TGACTGTAATGCAGGGTGCTGATTTTCGTTCGTCTGATCGCCTTCACGCTGTTCGTGATC 2040
TR251      ************************************************************ 2024

CDCBold    TGGCGTATCTCCCACAAGAACCCAGACGCGATGTGGCTGTGGGTGACATCCATCTGCGGC 2100
TR251      ************************************************************ 2084

CDCBold    GAGTTCTGGTTCGGTTTCTCGTGGCTGCTGGATCAGCTGCCCAAGCTGAACCCCATCAAC 2160
TR251      ************************************************************ 2144

CDCBold    CGCGTGCCGGACCTGGCGGTGCTGCGGCAGCGCTTCGACCGCCCCGACGGCACCTCCACG 2220
TR251      ************************************************************ 2204

CDCBold    CTCCCGGGGCTGGACATCTTCGTCACCACGGCCGACCCCATCAAGGAGCCCATCCTCTCC 2280
TR251      ************************************************************ 2264

CDCBold    ACCGCCAACTCGGTGCTCTCCATCCTGGCCGCCGACTACCCCGTGGACCGCAACACATGC 2340
TR251      *********************************************************** 2324

CDCBold    TACGTCTCCGACGACAGTGGCATGCTGCTCACCTACGAGGCCCTGGCAGAGTCCTCCAAG 2400
TR251      *********************************************************** 2384
```

160

```
CDCBold    TTCGCCACGCTCTGGGTGCCCTTCTGCCGCAAGCACGGGATCGAGCCCAGGGGTCCGGAG 2460
TR251      ************************************************************ 2444

CDCBold    AGCTACTTCGAGCTCAAGTCACACCCTTACATGGGGAGAGCCCAGGACGAGTTCGTCAAC 2520
TR251      ************************************************************ 2504

CDCBold    GACCGCCGCCGCGTTCGCAAGGAGTACGACGAGTTCAAGGCCAGGATCAACAGCCTGGAG 2580
TR251      ************************************************************ 2564

CDCBold    CATGACATCAAGCAGCGCAACGACGGGTACAACGCCGCCATTGCCCACAGCCAAGGCGTG 2640
TR251      ************************************************************ 2624

CDCBold    CCCCGGCCCACCTGGATGGCGGACGGCACCCAGTGGGAGGGCACATGGGTCGACGCCTCC 2700
TR251      ************************************************************ 2684

CDCBold    GAGAACCACCGCAGGGGCGACCACGCCGGCATCGTACTGGTCAGTATCCATCCATCTTTC 2760
TR251      ************************************************************ 2744

CDCBold    TGCTGCTTATATTACTCTTAGGTTACTCTTATCGTCTCTTTCCTATACTGTACATGCATG 2820
TR251      *********************************************C*********** 2804

CDCBold    CATGCTGCTATTCTTGGAATCGTGGTTGGTTACTACTCCACCATGCAAAAATAACAAGAA 2880
TR251      ************************************************************ 2864

CDCBold    GAGGAATCTTGGTTAGTTAGGGCCTCGTTGTTATATTAGTGGCCATCTGATGTGATGCCT 2940
TR251      ************************************************************ 2924

CDCBold    GCCGGCTGTGCCCATCCATATCCATGGAAGATTTCGACAGAATCGACGTGGTGATAGTCG 3000
TR251      **************************************************T* 2984

CDCBold    AGAGTGCAACCACCACCCAGAGCCAGCCAAGCACATGCATGCTTCTCTTCTCGTCTCGTC 3060
TR251      ************************************************************ 3044

CDCBold    GTGTGGCCAGCAGCGCATTCATGCTATTGCTGTGACGAGGGAGGAATGGTAGTTGGGGTG 3120
TR251      *********************************************G********* 3104

CDCBold    GTCCTTTCCCCCCGACAGCACTACAGCCTCCACTTTATGACCCATTTAATTCACCGGCCC 3180
TR251      ************************************************T****** 3164

CDCBold    TGCTTTGTTGTAACCGCCTTCTCATCTCAATCAATCATTCATTCATTCATAAGTTTACTC 3240
TR251      *********************C************************************** 3224

CDCBold    ACTCTTTGTTACTACTCGAACCACTAATCAGGAAGGAGTAGGAGTAATGCAGATTTACTA 3300
TR251      ************************************************************ 3284

CDCBold    TTGACAGTTAAAGGAGTAAAAAGAAGGAAGCACAATTACAGAACCTTGTTTTTTTTTACT 3360
TR251      **A********************************************************* 3344

CDCBold    ACTGTACGTAAGGTGTAAGAATGGAGTGCTGACAGAGAATGGATGCAGGTGCTGCTGAAC 3420
TR251      ************************************************************ 3404

CDCBold    CACCCGAGCCACCGCCGGCAGACGGGCCCGCCGGCGAGCGCTGACAACCCACTGGACTTG 3480
TR251      ************************************************************ 3464

CDCBold    AGCGGCGTGGATGTGCGTCTCCCCATGCTGGTGTACGTGTCCCGTGAGAAGCGCCCCGGG 3540
TR251      ************************************************************ 3524

CDCBold    CACGACCACCAGAAGAAGGCCGGTGCCATGAACGCGCTTACCCGCGCCTCGGCGCTGCTC 3600
TR251      ************************************************************ 3584

CDCBold    TCCAACTCCCCCCTTCATCCTCAACCTCGACTGCGATCATTACATCAACAACTCCCAGGCC 3660
TR251      ************************************************************ 3644
```

```
CDCBold    CTTCGCGCCGGCATCTGCTTCATGGTGGGACGGGACAGCGACACGGTTGCCTTCGTCCAG 3720
TR251      ************************************************************ 3704

CDCBold    TTCCCGCAGCGCTTCGAGGGCGTCGACCCCACCGACCTCTACGCCAACCACAACCGCATC 3780
TR251      ************************************************************ 3764

CDCBold    TTCTTCGACGGCACCCTCCGTGCCCTGGACGGCATGCAGGGCCCCATCTACGTCGGCACT 3840
TR251      ************************************************************ 3824

CDCBold    GGGTGTCTCTTCCGCCGCATCACCGTCTACGGCTTCGACCCGCCGAGGATCAACGTCGGC 3900
TR251      ************************************************************ 3884

CDCBold    GGTCCCTGCTTCCCCAGGCTCGCCGGGCTCTTCGCCAAGACCAAGTACGAGAAGCCCGGG 3960
TR251      ************************************************************ 3944

CDCBold    CTCGAGATGACCACGGCCAAGGCCAAGGCCGCGCCCGTGCCCGCCAAGGGTAAGCACGGC 4020
TR251      ************************************************************ 4004

CDCBold    TTCTTGCCACTGCCCAAGAAGACGTACGGCAAGTCGGACGCCTTCGTGGACACCATCCCG 4080
TR251      ************************************************************ 4064

CDCBold    CGCGCGTCGCACCCGTCGCCCTACGCCGCGGCGGCTGAGGGGATCGTGGCCGACGAGGCG 4140
TR251      ****************************A******************************* 4124

CDCBold    ACCATCGTCGAGGCGGTGAACGTGACGGCCGCCGCGTTCGAGAAGAAGACCGGCTGGGGC 4200
TR251      ************************************************************ 4184

CDCBold    AAAGAGATCGGCTGGGTGTACGACACCGTCACGGAGGACGTGGTCACCGGCTACCGGATG 4260
TR251      ************************************************************ 4244

CDCBold    CATATCAAGGGGTGGCGGTCACGCTACTGCTCCATCTACCCACACGCCTTCATCGGCACC 4320
TR251      ************************************************************ 4304

CDCBold    GCCCCCATCAACCTCACGGAGAGGCTCTTCCAGGTGCTCCGCTGGTCCACGGGATCCCTC 4380
TR251      ************************************************************ 4364

CDCBold    GAGATCTTCTTCTCCAAGAACAACCCGCTCTTCGGCAGCACATACCTCCACCCGCTGCAG 4440
TR251      ************************************************************ 4424

CDCBold    CGCGTCGCCTACATCAACATCACCACTTACCCCTTCACCGCCATCTTCCTCATCTTCTAC 4500
TR251      ************************************************************ 4484

CDCBold    ACCACCGTGCCGGCGCTATCCTTCGTCACCGGCCACTTCATCGTGCAGCGCCCGACCACC 4560
TR251      ************************************************************ 4544

CDCBold    ATGTTCTACGTCTACCTGGGCATCGTGCTATCCACGCTGCTCGTCATCGCCGTGCTGGAG 4620
TR251      ************************************************************ 4604

CDCBold    GTCAAGTGGGCCGGGGTCACAGTCTTCGAGTGGTTCAGGAACGGCCAGTTCTGGATGACA 4680
TR251      ************************************************************ 4664

CDCBold    GCAAGTTGCTCCGCCTACCTCGCCGCCGTCTGCCAGGTGCTGACCAAGGTGATATTCCGG 4740
TR251      ************************************************************ 4724

CDCBold    CGGGACATCTCCTTCAAGCTCACATCCAAGCTACCCTCGGGAGACGAGAAGAAGGACCCC 4800
TR251      ************************************************************ 4784

CDCBold    TACGCCGACCTCTACGTGGTGCGCTGGACGCCGCTCATGATTACACCCATCATCATCATC 4860
TR251      *******************************************C**************** 4844

CDCBold    TTCGTCAACATCATCGGATCCGCCGTGGCCTTCGCCAAGGTTCTCGACGGCGAGTGGACG 4920
TR251      ************************************************************ 4904
```

162

```
CDCBold    CACTGGCTCAAGGTCGCCGGCGGCGTCTTCTTCAACTTCTGGGTGCTCTTCCACCTCTAC 4980
TR251      ************************************************************ 4964


CDCBold    CCCTTCGCCAAGGGCATCCTGGGGAAGCACGGAAAGACGCCAGTCGTGGTGCTCGTCTGG 5040
TR251      ************************************************************ 5024


CDCBold    TGGGCATTCACCTTCGTCATCACCGCCGTGCTCTACATCAACATCCCCCACATGCATACC 5100
TR251      ************************************************************ 5084


CDCBold    TCGGGAGGCAAGCACACAACGGTGCATGGTCACCATGGCAAGAAGTTGGTCGACACAGGG 5160
TR251      ************************************************************ 5144


CDCBold    CTCTATGGCTGGCTCCATTGATGACTTTGCCCGGACAAGACGACCTGAGACAAGAAACAA 5220
TR251      ************************************************************ 5204


CDCBold    CTCATCCACTCAACAGTCAGTGCATGCATCCATCTCATCGAGAAGCAGAGCCCGCCAAAG 5280
TR251      ************************************************************ 5264


CDCBold    TTTGAATTTTTTAATTTTTTTTCTTCACTTTTTTGCCCGTTTCTTTTTAGTTTTGTCCAG 5340
TR251      ************************************************************ 5324


CDCBold    AAAAAAGATGGTGTTGATTTGATTTAGTTTATAATTACCTGTGGTAATTAATTATGTACT 5400
TR251      ************************************************************ 5384


CDCBold    TAATTATACATTCCGCGAACAACAAGGGAGACAGACGACTTACGGGGTACTGGCTCGGGT 5460
TR251      ************************************************************ 5444


CDCBold    GGTAAGAGCTTGCACTGTACTGTACATGCTCGACGATGTATAGAGATGCACAGAGGAGAG 5520
TR251      **************A********************************************* 5504


CDCBold    GATGGGAGTGCTGGGACCGTGGGGTGGACGGCGGTATTCTTTTAGTATTATATATGGAAA 5580
TR251      ************************************************************ 5564


CDCBold    CAATAAATTTAATTTCATTAATTCGTTCATGTGTTCACCTGTGGCCAATGACTTCATTTT 5640
TR251      ************************************************************ 5624


CDCBold    TTTTGTCAGGGAATCTACTTATAGTAGTACTACTTATTTGTTATTACTTCGTCCAAAAAT 5700
TR251      ************************************************************ 5684


CDCBold    AAGTATATTTACCCACTAATTGCTTGACAGTGTTTATTCCTTCGCCAAAATTAATGATTA 5760
TR251      ************************************************************ 5744


CDCBold    AAATAGCTTGACAGTAGTGCGGGACTAGTCAGTCATGACGATTGAGCCGAGCATGTTCTC 5820
TR251      ************************************************************ 5804


CDCBold    TTCATTATGTTTCCGGAACTTTATTCTGTCCCACTATGTATTACCGAATATTTGACTCGT 5880
TR251      ************************************************************ 5864


CDCBold    AAAAAGCTGCAAGCCATGTGATCGATCTGCAGTTGCACTGTAAATGTTCATGATCTACTT 5940
TR251      ************************************************************ 5924


CDCBold    TCATCGTGGACGTTGGAGTTGGACCGTATACTCCGCACCTCTTTTACAGAAAAACCACCA 6000
TR251      ************************************************************ 5984


CDCBold    TCTATCATGGTATAATTCTTGTCGTGTTCCTATGACTTTCTTCAACCAAGACCCTTTTAC 6060
TR251      TCTATCATGGTATAATTCTTGTCGTGTTCCTATGACTTTCTTCAACCAAGACCCTTTTAC 6044
           ***********************************************************


CDCBold    ATCAACTAAAAATGTAATACGGGATGAAACCACCATCTATCGGGGAACGGTCGGTGGAAC 6120
TR251      ATCAACTAAAAATGTAATACGGGATGAAACCACCATCTATCGGGGAACGGTCGGTGGAAC 6104
           ***********************************************************


CDCBold    TAGCATATCGAGGAGACTTGTCGTCTCATTGAAAAGATGAATCTACAGGAGGATGCGGAA 6180
```

163

```
TR251        C**********************************************************  6164


CDCBold      GATGATCTTGTGTGGGAAGAGGAATTAGACGTGGAGGAAATCAAGCCAAAGTGGCTGGTG  6240
TR251        **********************************************************    6224


CDCBold      ATCGGGCGTCTTTTGGCGCAGAAATCCTTCACTAACAGCACGCTGATTGCACACATGAAA  6300
TR251        **********************************************************    6284


CDCBold      GCTACTTGGAATCCAGCACGAACAATGGTGTGGATGAGGATCAACGCCAACCTATTCACC  6360
TR251        **********************************************************    6344


CDCBold      ATCGAATTCAATTGCCTTGGAGACTGGAACAAAGCAATGCATGAGGGCCCATGGGATTTT  6420
TR251        **********************************************************    6404


CDCBold      CGTGGTCTCGCGCTAATCCTGACACAATATGATGGATTCTCCGAACCTGAGAAAGTCAAA  6480
TR251        *********************------------------------------------    6426


CDCBold      CTCGATAGGTTAGAAACTTGGTGCCAAATTCATAGGCTCCCTGATGGGAGTACCAGCCTG  6540
TR251        ------------------------------------------------------------
```

164

Figure S2. Alignment of amino acid sequence for HvCSLF6 from barley CDC Bold and TR251

```
CDC Bold     MAPAVAGGGRVRSNEPVAAAAAAPAASGKPCVCGFQVCACTGSAAVASAASSLDMDIVAM 60
TR251        ************************************************************ 60

CDC Bold     GQIGAVNDESWVGVELGEDGETDESGAAVDDRPVFRTEKIKGVLLHPYRVLIFVRLIAFT 120
TR251        ************************************************************ 120

CDC Bold     LFVIWRISHKNPDAMWLWVTSICGEFWFGFSWLLDQLPKLNPINRVPDLAVLRQRFDRPD 180
TR251        ************************************************************ 180

CDC Bold     GTSTLPGLDIFVTTADPIKEPILSTANSVLSILAADYPVDRNTCYVSDDSGMLLTYEALA 240
TR251        ************************************************************ 240

CDC Bold     ESSKFATLWVPFCRKHGIEPRGPESYFELKSHPYMGRAQDEFVNDRRRVRKEYDEFKARI 300
TR251        ************************************************************ 300

CDC Bold     NSLEHDIKQRNDGYNAAIAHSQGVPRPTWMADGTQWEGTWVDASENHRRGDHAGIVLVLL 360
TR251        ************************************************************ 360

CDC Bold     NHPSHRRQTGPPASADNPLDLSGVDVRLPMLVYVSREKRPGHDHQKKAGAMNALTRASAL 420
TR251        ************************************************************ 420

CDC Bold     LSNSPFILNLDCDHYINNSQALRAGICFMVGRDSDTVAFVQFPQRFEGVDPTDLYANHNR 480
TR251        ************************************************************ 480

CDC Bold     IFFDGTLRALDGMQGPIYVGTGCLFRRITVYGFDPPRINVGGPCFPRLAGLFAKTKYEKP 540
TR251        ************************************************************ 540

CDC Bold     GLEMTTAKAKAAPVPAKGKHGFLPLPKKTYGKSDAFVDTIPRASHPSPYAAAAEGIVADE 600
TR251        ***********************************************T********** 600

CDC Bold     ATIVEAVNVTAAAFEKKTGWGKEIGWVYDTVTEDVVTGYRMHIKGWRSRYCSIYPHAFIG 660
TR251        ************************************************************ 660

CDC Bold     TAPINLTERLFQVLRWSTGSLEIFFSKNNPLFGSTYLHPLQRVAYINITTYPFTAIFLIF 720
TR251        ************************************************************ 720

CDC Bold     YTTVPALSFVTGHFIVQRPTTMFYVYLGIVLSTLLVIAVLEVKWAGVTVFEWFRNGQFWM 780
TR251        ************************************************************ 780

CDC Bold     TASCSAYLAAVCQVLTKVIFRRDISFKLTSKLPSGDEKKDPYADLYVVRWTPLMITPIII 840
TR251        ************************************************************ 840

CDC Bold     IFVNIIGSAVAFAKVLDGEWTHWLKVAGGVFFNFWVLFHLYPFAKGILGKHGKTPVVVLV 900
TR251        ************************************************************ 900

CDC Bold     WWAFTFVITAVLYINIPHMHTSGGKHTTVHGHHGKKLVDTGLYGWLH 947
TR251        *********************************************** 947
```

165

Supplementary Table 1. Genes within intervals Ensembl Genome Release 23

Putative Genes	Description
1H:4,296,096 - 4,496,069	
MLOC_19284	Defense response
MLOC_74537	Nuclear inhibitor of protein phosphatase 1
MLOC_44678	Oxidation-reduction process, 12-oxophytodienoate reductase 1
MLOC_37184	Actin nucleation, Formin-like protein 14
MLOC_44602	Vacuolar ATP synthase proteolipid subunit
MLOC_37432	Putative zinc finger protein
MLOC_76007	Methylation, 5-pentadecatrienyl resorcinol O-methyltransferase
MLOC_50847	Defence resonse, Defensin
MLOC_81200	Structural constituent of ribosome
MLOC_26623	Uncharacterized
MLOC_923	Uncharacterized
MLOC_64274	Protein binding, WD repeat-containing protein
2H-79055395 - 80248284	
MLOC_57325	Microtubule based movement, Kinesin-like protein
MLOC_56621	Histone H3-T11 phosphorylation, gene silencing, haspin
MLOC_56623	NPSN13, Plant SNARE 13, involved in membrane fusion
MLOC_20146	Xylulose metabolic process, xylulokinase activity,
MLOC_10665	Uncharacterised
MLOC_61562	Oxidation-reduction process
4H: 8111198 to 13266457	
MLOC_14146	Uncharacterized
MLOC_36702	Uncharacterized
MLOC_62785	Heat shock protein
MLOC_59382	Uncharacterized
MLOC_53218	Protein phosphorylation, embryo sac development
MLOC_4231	Uncharacterized
MLOC_66158	Uncharacterized
MLOC_57855	Nucleobase-ascorbate transporter LPE1
MLOC_6266	Nucleic acid binding, RNA-binding region RNP-1
MLOC_10096	Putative thioesterase family protein
MLOC_5212	Pentatricopeptide, putative, expressed
MLOC_10935	Golgi organization, Costars family protein
MLOC_10938	Positive regulation of Rab GTPase activity,
MLOC_55179	Glutamine amidotransferase class-I family protein
MLOC_55180	CHCH domain containing protein
MLOC_75098	Response to stress, oxidation-reduction process

MLOC_79660	Regulation of transcription, DNA-templated
MLOC_36424	Gene silencing by RNA, hydrolase activity, acting on ester bonds
MLOC_56677	Transferase activity, transferring glycosyl groups
MLOC_56679	Uncharacterized
MLOC_51353	Putative CXC domain protein
MLOC_64337	tRNA aminoacylation for protein translation, Lysine
MLOC_58755	Cellular cation homeostasis
MLOC_51725	Acute-phase response protein
MLOC_63148	Uncharacterized
MLOC_63149	Uncharacterized
MLOC_11551	Uncharacterized
MLOC_63465	Transferase activity, transferring glycosyl group, Exostosin-2
MLOC_62203	Uncharacterized
MLOC_78576	Protein binding,
MLOC_15805	Uncharacterized
MLOC_7897	Expressed protein; Putative glycine rich protein
MLOC_37911	Pentose-phosphate shunt, abscisic acid biosynthetic process
MLOC_52290	Phosphorylation, kinase activity, Isoamylase N-terminal domain
MLOC_3791	Acid phosphatase activity
MLOC_11464	Uncharacterized
MLOC_55029	LOXB, fatty acid biosynthesis , oxylipin biosynthetic process
MLOC_54031	LOX1.1, fatty acid metabolic process, oxidation-reduction process
MLOC_5268	Hydrolase activity, acting on carbon-nitrogen (but not peptide)
MLOC_5269	Protein, metal ion binding, Zinc finger CCCH domain
MLOC_52027	Protein binding
MLOC_4717	DNA-dependent DNA replication, Cell division cycle protein
MLOC_70449	Protein import into nucleus, transcription coactivator activity
MLOC_19267	Uncharacterized
MLOC_15209	Meiotic chromosome segregation
MLOC_19176	Protein yippee-like
MLOC_15467	PURA, purine nucleotide biosynthetic process
MLOC_74132	Cell wall macromolecule catabolic process, LysM domain
MLOC_63077	Oxidation-reduction process, UDP-glucose 6-dehydrogenase
MLOC_53797	Unidimensional cell growth, ion transmembrane transport
MLOC_53798	Proteolysis involved in cellular protein catabolic process,
MLOC_53799	Uncharacterized
MLOC_80258	Uncharacterized
MLOC_2842	ERS1B, phosphorelay signal transduction system
MLOC_76806	Uncharacterized
4H near 31395904	
MLOC_53722	Putative receptor protein kinase ZmPK1
MLOC_53721	Ribosomal RNA small subunit methyltransferase E

MLOC_49818	Uncharacterized

4H near 321179812
MLOC_21017	APRATAXIN-LIKE, APTX

5H-7491839 to 10176328
MLOC_21074	Uncharacterized
MLOC_42238	Uncharacterized
MLOC_3994	Uncharacterized
MLOC_2781	OS12G0630500
MLOC_20378	OS12G0628600
MLOC_76989	Terpene synthase activity, magnesium ion binding
MLOC_59480	Terpene synthase activity
MLOC_75397	Uncharacterized
MLOC_50740	Uncharacterized
MLOC_60894	Putative bZIP transcription factor superfamily protein
MLOC_60893	Anaphase-promoting complex subunit
MLOC_71885	RNA methylation
MLOC_71887	Lipid metabolic process

Genes near 7H-68292885
MLOC_76756	Rhodanese / Cell cycle control phosphatase superfamily
MLOC_73315	Transferase activity, transferring glycosyl groups
MLOC_36826	CXP;2-2, second-messenger-mediated signaling
MLOC_75506	Zinc ion binding
MLOC_12686	Transmembrane transport
MLOC_62371	Uncharacterized
MLOC_67897	Protein catabolic process, nucleoside-triphosphatase activity