



2012-08-09

# Characterization and Variable Expression of the CslF6 Homologs in Oat (*Avena* sp.)

Melissa A. Coon

*Brigham Young University - Provo*

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Animal Sciences Commons](#)

---

## BYU ScholarsArchive Citation

Coon, Melissa A., "Characterization and Variable Expression of the CslF6 Homologs in Oat (*Avena* sp.)" (2012). *All Theses and Dissertations*. 3750.

<https://scholarsarchive.byu.edu/etd/3750>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

Characterization and Variable Expression of the CslF6 Homologs  
in Oat (*Avena* sp.)

Melissa Coon

A thesis submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Master of Science

Eric N. Jellen, Chair  
Peter J. Maughan  
Gongshe Hu

Department of Plant and Wildlife Sciences

Brigham Young University

December 2012

Copyright © 2012 Melissa Coon

All Rights Reserved

## ABSTRACT

Characterization and Variable Expression of the CslF6 Homologs  
in Oat (*Avena* sp.)

Melissa Coon  
Department of Plant and Wildlife Sciences, BYU  
Master of Science

(1,3;1,4)- $\beta$ -D-glucan ( $\beta$ -glucan) is a plant cell wall hemicellulose and a main component of endosperm cell walls. The Cellulose Synthase F family of genes is involved in the synthesis of  $\beta$ -glucan. In this study full-length genomic sequences of CslF6 were obtained from multiple *Avena* species. Three unique alleles were found in each *A. sativa* line. Comparisons of these alleles to diploid *Avena* species allowed for identification of the genomic origin of each allele. The A and D genome alleles had identical amino acid sequences while the C-genome had 13 different amino acids. Global expression of CslF6 was completed at three developmental time point and three tissue types. RNAseq technology was utilized to determine genome specific expression patterns. Differential expression of genome specific-copies of CslF6 was found at all time points tested. Lower levels of C-genome expression of CslF6 were associated with increased levels of B-glucan.

Keywords:  $\beta$ -glucan, differential expression, *Avena sativa*, oat, cslf6

## **ACKNOWLEDGMENTS**

I would like to thank my major professor Dr. Rick Jellen for his support and guidance throughout this project. I am grateful for the support of my family and graduate committee for their help along the way. Gongshe Hu for his patient assistance as I learned the basics of cloning. Robert Campbell for collection of tissue and RNA extraction. Thanks you to Emir Islamovic for his constant guidance. Eric Jackson for his leadership in this project. I would like to thank the staff at the USDA-ARS Aberdeen for their support of the research process. Mike McMullen and Nick Tinker for donation of germplasm for this project. I would like to recognize BYU, USDA-ARS, General Mills, North American Millers Association and Quaker for the generous funding for this project. My appreciation for the Department of Plant and Wildlife Sciences at Brigham Young University for providing academic, administrative, and financial support.

## TABLE OF CONTENTS

TITLE PAGE.....	i
ABSTRACT.....	ii
ACKNOWLEDGMENTS.....	iii
TABLE OF CONTENTS.....	iv
Chapter 1: Characterization and variable expression of the CslF6 homologs in oat ( <i>Avena sp.</i> ) ..	1
Introduction.....	2
Materials and Methods.....	4
Plant Materials.....	4
DNA extraction.....	4
Primer design and PCR amplification.....	4
Cloning and sequencing.....	5
RNA Isolation and cDNA synthesis.....	5
Real-time quantitative PCR.....	6
RNAseq- allele specific analysis.....	6
Predictive 3D structure.....	8
Mapping by sequencing.....	8
Results.....	9
Sequencing and allele assignment.....	9
Gene ontology.....	10
Unique variants.....	11
Global CslF6 expression.....	11
RNAseq- Genome expression analysis.....	12
CslF6 D-genome marker.....	13
Discussion.....	13
Species comparisons.....	13
Global expression.....	14
Genome- specific expression.....	14
References.....	30
Chapter 2: Review of (1,3;1,4)- $\beta$ -D-glucan Synthesis in Cereals.....	32
Introduction.....	32
Oat production.....	32
Oat usage and consumption.....	33
(1,3;1,4)- $\beta$ -D-glucan.....	34
Fine Structure and Function.....	34
Economic Importance.....	36
Biosynthesis.....	37
Cellulose Synthase-Like F6 (CslF6).....	40
Conclusion.....	42
References.....	45

## LIST OF TABLES

Table 1. Sequencing Panel.....	16
Table 2. Cloning and Sequencing Primers.....	17
Table 3. Primers for Relative Global Expression of CslF6 .....	17
Table 4. Global Relative Expression Panel.....	17
Table 5. RNASeq Differential Expression Panel.....	18
Table 6. RNASeq SNP Locations.....	18

## LIST OF FIGURES

Figure 1. Examples of stages of oat embryo development collected for RNA extraction. From left to right: 1-3 days post anthesis (DPA), 4-6 DPA, 7-9 DPA. ....	19
Figure 2. Diagram of CslF6 genomic sequence with exons in blue and introns in red. Arrows mark locations of primers used for cloning.....	19
Figure 3. Splice sites for CslF6.....	20
Figure 4. Unrooted nucleotide tree displaying comparisons between the three CslF6 sequences in oat and other related species. ....	21
Figure 5. Unrooted tree showing comparisons of translated sequences of CslF6 between the three alleles of oat and other related species.....	22
Figure 6. Relative global expression of CslF6 using sybr green chemistry.....	22
Figure 7. Relative global expression of CslF6 using sybr green chemistry. Comparisons across all time points with in each line .....	27
Figure 8. Genome Specific Expression using RNAseq. ....	28

## **Chapter 1: Characterization and variable expression of the CslF6 homologs in oat (*Avena sp.*)**

Melissa A. Coon<sup>1</sup>, Eric W. Jackson<sup>2</sup>, Emir Islamovic<sup>3</sup>, Robert Campbell<sup>3</sup>, Gongshe Hu<sup>3</sup>, P. Jeffery Maughan<sup>1</sup>, Gerard Lazo<sup>4</sup>, Rebekah W. Oliver<sup>2</sup>, Joe Lutz<sup>2</sup>, Robert Reid<sup>6</sup>, Nick Tinker<sup>5</sup>, Michael S. McMullin<sup>7</sup>, Eric N. Jellen<sup>1</sup>

1 Brigham Young University, Department of Plant and Wildlife Sciences

2 General Mills

3 USDA-ARS, National Small Grains and Potato Germplasm Research Unit, Aberdeen ID

4 USDA-ARS, Western Regional Research Center, Albany CA

5 Agriculture and Agri-Food Canada, Ottawa

6 University of North Carolina Charlotte, Department of Bioinformatics and Genomics

7 North Dakota State University, Department of Plant Science



## Introduction

Mixed Linkage (1,3;1,4)- $\beta$ -D-glucan ( $\beta$ -glucan) is a cell wall hemicellulose that is found primarily in the grasses, including the cereals (Poaceae).  $\beta$ -glucan accumulates transiently in walls of vegetative tissue and in the secondary cell walls of the xylem and phloem.  $\beta$ -glucan is a major component of endosperm cell wall, accounting for up to 70% of cell wall weight (Carpita 1996; Fincher and Stone 2004). Cultivated oat (*Avena sativa*) and barley (*Hordeum vulgare*) contain especially high levels of this polysaccharide. The  $\beta$ -glucan content of oat grain ranges from 3- 7% and from 5- 11% in barley (Aman and Graham 1987). Oat  $\beta$ -glucan can be more soluble than barley due to its decreased ratio of (1,4) to (1,3)- $\beta$ -D- glucosyl residues.

Soluble dietary fiber and more particularly  $\beta$ -glucans have been recognized for their positive impact on human health. Consumption of  $\beta$ -glucan accompanied by a diet low in saturated fat has been shown to lower blood serum cholesterol (Kirby et al 1981; Anderson et al. 1984; Shimizu et al. 2008; Collins et al. 2010; Wood 2007; Braaten et al 1994). In 1997 the Federal Drug Administration (FDA) authorized a health claim on the relationship between soluble fiber in whole oats and a reduced risk of cardiovascular or heart disease. The claim stated that products that contain at least .75 grams of  $\beta$ -glucan per serving are authorized to advertise this claim on the packaging of the product. The FDA recently extended this health claim to include barley, oat bran, rolled oats, and whole oat flour as sources of  $\beta$ -glucan.  $\beta$ -glucan is also an important negative component in feed for monogastric animals such as pigs and poultry. In particular, high levels  $\beta$ -glucan in animal feed can have anti-nutritive effects including decreased weight gain, reduced nutritive uptake, and sticky droppings (Hesselman and Aman 1986).  $\beta$ -glucan levels are also one of the most influential characteristics in the malting and brewing of barley, negatively affecting the viscosity of the wort.

Biosynthesis of  $\beta$ -glucan in the Poaceae is mediated, at least in part, by the cellulose-synthase-like F (CslF) family of genes (Farrokhi 2006; Fincher 2009a; Fincher 2009b; Burton and Fincher 2009). *Arabidopsis* plants transformed with rice *CsIF* genes began to accumulate  $\beta$ -glucan in their cell walls. *Arabidopsis* plants do not produce (1,3; 1,4)- $\beta$ -D-glucan in their cell walls and the CslF gene family is absent in their genome (Burton et al. 2006). Among members of the CslF family of genes CslF6 could be of major importance. Transcriptional analysis of the barley *HvCsIF* gene family demonstrated that *CsIF6* had much higher transcript levels throughout endosperm development as compared with other *HvCsIF* genes (Burton et al. 2008). Down-regulation of the *CsIF6* gene in wheat using RNA interference (RNAi) resulted in an average reduction of  $\beta$ -glucan content of 42.4% in mature wheat grain (Nemeth et al. 2010). In addition overexpression of the barley CslF6 gene under the control of an endosperm-specific oat globulin promoter resulted in > 80% increase in  $\beta$ -glucan content in grain of transgenic barley (Burton et al. 2011).

Much of the research published on the CslF gene family and (1,3;1,4)- $\beta$ -D-glucan synthesis has focused on barley. We sought to use barley as model to identify the variants responsible for (1,3;1,4)- $\beta$ - D-glucan in oat. Oat is an allohexaploid ( $2n = 6x = 42$ , AACCCDD subgenomes) with a large, complex genome. Oat research presents challenges in that there are currently limited genetic resources for studying and improving this species. Our study elucidates the genetic variation in oat CslF6 homologs and utilizes a new technique to study subgenome-specific expression.

## **Materials and Methods**

### **Plant Materials**

Germplasm for this project was acquired from the National Small Grains Collection located in Aberdeen, Idaho and from the collection of *Avena* species available at Brigham Young University. Twenty accessions of *A. sativa* were chosen to represent the most important germplasm available for the breeding of oat in North America. A variety of wild accessions of *Avena* were also chosen to represent the three diploid genomes (A, C, D) and their various combinations in allotetraploids and *A. sativa*. A list of the chosen germplasm is provided in Table 1.

### **DNA extraction**

Young leaf tissue was harvested from plants and ground in liquid nitrogen.

Genomic DNA was isolated from the selected lines using “Plant DNA Extraction Protocol for DArT” a CTAB/chloroform extraction protocol (Diversity Arrays Technology Pty Ltd, Australia). DNA was quantified using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, Delaware) in preparation for PCR amplification.

### **Primer design and PCR amplification**

Expressed Sequence Tags (ESTs) from the Collaborative Oat Research Enterprise (CORE) database with predicted protein homology to *CsIF6* were assembled using Sequencher 4.10.1 (Gene Codes Corporation, Ann Arbor, Michigan) or Geneious Pro (Biomatters LTD, New Zealand). The full-length barley coding sequence (CDS) was used as a reference. The EST assembly covered the majority of the *CsIF6* coding sequence not including 200 base pairs at the 5' end. From this assembly a series of primers was designed to amplify the full-length *CsIF6*

gene (Table 2). PCR amplification was done using Phusion® High-Fidelity DNA Polymerase (New England Biolabs Inc., Ipswich, MI) with HF buffer. PCR product was electrophoresed in 1% agarose gel. Bands were excised from the agarose gels and purified using QIAquick gel extraction kit (Qiagen, Valencia, CA) using the spin protocol.

### **Cloning and sequencing**

Cloning of blunt-ended PCR products was done using Strata Clone Blunt PCR Cloning Kit (Agilent Technologies, Cedar Creek, Texas) with kanamycin as selective antibiotic. Colonies were screened for correct insert using PCR. QuickClean Plasmid Miniprep kit (GeneScript Corporation, Piscataway, NJ) was used to isolate high quality plasmid DNA from 5ml bacterial culture. Restriction digest of plasmids using EcoRI and ApaLI (New England Biolabs Inc., Ipswich, MI) was done to verify correct insertion size. Colonies were sequenced at Brigham Young University DNA Sequencing Center using Big Dye v3.1 chemistry and electrophoresed on Applied Biosystems 3730xl DNA Analyzer. Sufficient colonies were sequenced to acquire full-length sequences from each genome. Analysis of sequence data will be done using Geneious Pro (Biomatters LTD, New Zealand).

### **RNA Isolation and cDNA synthesis**

Seed from each line was planted in a 1:1:1 (v:v:v) sand, peat, potting soil mix contained in two 35.4 cm pots, eight seed per pot (n = 16 per line). The pots were incubated in two separate growth chambers programmed with a 15-hour photoperiod and a 20°/15° C light/dark temperature cycle. Tissue samples from the selected germplasm were collected at 4 pm and frozen with liquid nitrogen prior to storage at – 80 degrees C. Tissue was collected from roots, shoots, mature embryos and at 3 developmental time points: 1-3 days post-anthesis (DPA), 4-6 DPA and 7-9 DPA (Figure 1). Tissue was finely ground in liquid nitrogen and total RNA was

extracted from tissue samples using UltraClean Plant RNA Isolation Kit (Mobio, Carlsbad, California) according to manufacturer's guidelines with 1% Beta-Mercaptoethanol added to denature RNase. cDNA synthesis was conducted using SuperScript III First-Strand Synthesis for RT-PCR (Invitrogen life technologies, Carlsbad, California) using 1ug total RNA. The provided oligo(dT) primers were used. The quality of RNA and cDNA was checked by gel electrophoresis.

### **Real-time quantitative PCR**

A gene specific primer pair was designed in the third exon of the gene located near the 3' end. The primers were designed in a conserved area in order to amplify all homologs of CslF6. The product size for the selected primers was 350 bp in length. Beta actin was used as an internal reference gene (Table 3). Samples were run in triplicate on a Stratagene Mx3005P QPCR system using Brilliant SYBR<sup>®</sup> Green QPCR Master Mix (Stratagene-Agilent, Carlsbad, CA). Reactions were 25 ul, 60 ng cDNA, 150ng primers. Program 10 min 95 °C, 40 cycles: 30 sec, 95 °C; 60 sec, °C; 60 sec 72 °C. Primers are listed in the appendix. Relative quantification of transcripts was done using the  $2^{-\Delta\Delta C_T}$  method as outlined by Livak and Schmittegen (2001). A list of germplasm used for this analysis can be found in Table 4.

### **RNAseq- allele specific analysis**

RNA from the 3 DPA developmental stages was prepared for HiSeq sequencing using the TruSeq chemistry. The lines used for this analysis were chosen for their varying levels of B-glucan and importance in oat breeding (Table 5). All lines used for this experiment were *A. sativa* lines except for Strimagdo. Strimagdo is a synthetic hexaploid line created by Gideon Ladizinsky from a cross between *A. strigosa* and *A. magna* (Ladizinsky 2000). 'Hifi' and OT3044 are high B-glucan. 'Hifi' is a breeding line developed by Mike McMullin of North

Dakota State University. 88304 is a high  $\beta$ -glucan mutant line. Leggett and Marion have mid-high levels of  $\beta$ -glucan. OT3018 and Strimagdo have low levels of  $\beta$ -glucan.

Raw RNAseq reads went through a filtering process. First, duplicated reads were discarded. A read was considered a duplicate if it had a 100% identity to another read. Low quality bases were trimmed from the end of the reads using a Q-score of 20 as the threshold. Lastly reads were screened for the presence of adapter sequence and subsequently removed. Filtered reads from all 33 conditions were aligned to a representative consensus CslF6 gene using the Bowtie alignment program (Langmead et al. 2009). Matching RNA SEQ reads from Bowtie were then assembled *de novo*, to generate transcripts using Velvet / Oases (kmer size= 49, reads = short, cov\_cutoff = 8) (Schulz et al. 2012). The number of reads per transcript were calculated and then used for the subsequent SNP analysis step. The composition for each nucleotide {A,C,G,T} was then determined for each of the putative SNP locations within CslF6.

To determine the expression of each homeolog of CslF6, SNPs were selected that could distinguish reads of each genome. SNPs were selected based on an alignment of multiple sequences of each genome. Only SNPs present in 100% of sequences of a single homeolog were chosen for analysis. Three SNPs were chosen for each homeolog distributed one SNP per exon when possible. There was not a SNP to distinguish the A genome from C and D in exon 1. Similarly there was not a SNP that could distinguish the D genome from A and C in exon 3. The SNP locations can be seen in Table 6. Read counts of each nucleotide were recorded at each SNP location. This process was repeated for every variety and condition.

To determine relative expression of each allele, normalization and analysis of reads counts was completed. First, the percentage of reads that correspond to each allele (A,C, or D) were calculated for each SNP. These calculations were completed with respect to exon; when there

were two SNPs representing a single exon an average percentage was used. This created a 3x3 table for each time point and line with alleles across the top and exon 1, 2, or 3 down the side. Next the percentages of reads in each exon were added giving 3 totals, one for each exon. The proportion of each allele's contribution to each exon total is then calculated and recorded into a new 3x3 table. The average proportion for each allele is then calculated. The sum of all the averages is then used to calculate the final percentage of expression for each allele. This final value is calculated by dividing the average proportion for each allele by the sum of all proportions.

### **Predictive 3D structure**

Full-length coding sequences of CslF6 were obtained from genomic sequences by splicing out intron sequences. Splice sites were determined through comparison of our generated oat genomic sequences with coding sequences of CslF6 from barley (*Hordeum vulgare* L.), rice (*Oryza sativa* L.) and oat available on NCBI. Translations of the three coding sequences showed that the A and D had identical amino acid sequences while the C genome sequence had approximately 13 changes. The 2 unique amino acid sequences were put through I-TASSER or The iterative threading assembly refinement server. I-TASSER is a unified platform for automated protein structure and function prediction (Roy et al. 2010). This program first generates three-dimensional (3-D) atomic models from multiple threading alignments. The function of the protein is then inferred by structurally matching the 3-D models with other know proteins.

### **Mapping by sequencing**

To create a marker for the CslF6 D genome allele specific primers were designed to amplify a short region of intron 2 (Table 7). The product size was approximately 250 bp. This region was then amplified in mapping parents 'Hifi' and 'SolFi' and all 52 RILs using HotStart Taq Master

Mix Kit (Qiagen, Valencia, CA). PCR products were purified with exonucleaseI and antarctic phosphatase (New England Biolabs Inc., Ipswich, MI) and directly sequenced as mentioned previous. Sequences were assembled in Geneious (Biomatters LTD, New Zealand) and genotypes were determined by individual inspection. The map was created using JoinMap (Kyazma B.V., Netherlands) with a LOD join score of 4.5 and regression analysis.

## **Results**

### **Sequencing and allele assignment**

Multiple primer sets were used to amplify and clone CslF6 genomic sequence. The largest fragment cloned was 5.5 kb long and included 30 bp upstream of the start codon and 160bp downstream of the stop codon. A smaller fragment of approximately 3 kb was also cloned. This smaller fragment began at the start of the second intron and extended 160 bp beyond the stop codon. Approximately 5- 10 clones were sequenced at random by Sanger sequencing per variety. Alignment of sequences from each variety showed that three distinct alleles of CslF6 were present. Diploid oat accessions were sequenced to determine if the three alleles of CslF6 sequenced are homologs originating from the three genomes in oat. *Avena ventricosa* (CC), *A. canariensis* (putative DD), *A. strigosa* (AA) and *A. wiestii* (AA) were the chosen diploid species for comparison. *Avena strigosa* and *A. wiestii* sequences were exactly identical to each other, confirming the traditional assignment of both taxa to subgenome A<sub>s</sub> A<sub>s</sub>. Comparison of the diploid sequences with the sequences from the hexaploid oat allowed for a determination of genome origin.

The C-genome ortholog in particular was sequenced at a very low frequency as compared to the A- and D- genome sequences. In some cases the screening of a large number of clones was



insufficient to obtain a full-length sequence of the C genome. In most cases a shorter partial sequence could be obtained, confirming the existence and identity of the allele.

### **Gene ontology**

The genomic sequence of CslF6 was approximately 5.2 kb. Due to large differences in intronic regions between each homolog the gene length can differ. The A-genome ortholog was 5,268 bp, the C-genome 5,162 bp, and the D-genome sequence was 5,162 bp in length. CslF6 had two introns, the first measuring 1,600 bp and the second 748 bp. A graphical representation is provided in Figure 2.

Splice sites of CslF6 were determined through comparison of the *A. sativa* genomic sequences with coding sequences (CDS) of CslF6 from *H. vulgare* and *A. sativa*. The splice site of the second intron could have been located in two different spots. Both splice site locations in the A and D genome produced the same coding sequence (Figure 3). For the C-genome a splicing event at the first location will result in a single base pair change in the coding sequence due to the difference in sequence from ‘AAGGT’ in the A- and D-genomes to ‘AGGGT’ in the C-genome. This single base change alters the amino acid residue 356 in the C-genome from a Leucine to an Arginine. Splicing the C-genome at the second site produced a coding sequence similar to the A- and D-genomes. Sequence data suggests splicing can occur at both sites but there is no conclusive evidence that splicing occurs at one site over the other. Splicing of the C-genome was conducted in concurrence with A- and D-genomes for this study. The coding sequences were very highly conserved among homoeologs. The A- and D-genome full-length CDS were 2,845 bp, while the C-genome full length CDS was one amino acid shorter at 2,842 bp long. C-genome sequences have a 3 bp deletion at position 57 to 59 in exon 1. This deletion results in the removal of a single serine residue while maintaining the same open reading frame.

Translations of each of the sequences revealed that the amino acid sequences for A and D genomes were identical while the C genome had 13 amino acid changes.

### **Unique variants**

Through comparison of all the sequences it was found that there were some unique insertions found in two of the cultivars. In 'Hifi' clone 10 a 352 bp insertion was found near the start of the first intron. A BLASTn search of this sequence found that this sequence has some similarity to a group of *Hordeum vulgare* subsp. *vulgare* Talisker transposons. The insertion does not appear to interfere with the splicing of the mRNA. In Marion clone 7 a small insertion of 118 bp was found in the second intron. Within this insertion a 51 bp region is a direct repeat of an adjacent region down stream. Searches reveal this region bears similarity to a microRNA.

### **Global CslF6 expression**

Global expression of the CslF6 gene was completed using sybr green chemistry on seven *Avena* species (Table 4). Expression data was collected from roots, shoots, mature embryos and 3 developmental time points: 1-3 days post-anthesis (DPA), 4-6 DPA and 7-9 DPA. Comparisons were made between the three developmental time points in each line and time point 1-3 DPA was used as a calibrator for calculations (Figure 6). Comparisons were also made between lines at each developmental stage and tissue types. For these calculations the cultivar Leggett was arbitrarily chosen to be the calibrator (Figure 7). Expression is reported as  $2^{-\Delta\Delta C_T}$  values or fold change in gene expression normalized to B-actin and relative to the calibrator sample. Graphs of relative gene expression show the calibrator lines with a fold change of one.

Expression of CslF6 is generally greater in the first two developmental stages and decreases by 6-9 DPA. Expression of CslF6 was consistently less in mature embryos than in any embryo

development stage. An exception to this trend was *A. strigosa* that had intermediate expression of CslF6 in mature embryos.

When lines were compared to each other it was evident that expression in the *A. sativa* lines was greater in the three developmental time points. *Avena strigosa* and Bam49-2 both had the lowest expression of CslF6 at all developmental time points. Expression of Cslf6 in shoot tissue was relatively consistent among lines except for ‘Schwartzhafer’ that had nearly a 2-fold increase in CslF6 expression.

### **RNAseq- Genome expression analysis**

The null hypothesis for this experiment was that expression of CslF6 was equal among all homeologs or genomes. Our results show that there was differential expression in most lines and time points. At 1-3 days (Figure 8a) DPA C-genome expression was around 25% with the exception of ‘Strimagdo’, which had approximately 35% expression. Expression of the A and D genomes was more similar, with each accounting for 35% to 45% of the total expression. At 4-6 DPA ‘Strimagdo’ had a marked increase in the proportion of D-genome expression with the necessary decrease in C- genome expression, the proportion of A genome expression remained relatively constant (Figure 8b). Expression in OT3018 also changed substantially with the expression of all genomes being equivalent. At the 7-9 DPA time point there was a greater equalization of expression between genomes; some differential expression can still be seen but the magnitude was not as pronounced. Leggett and 88304 had almost equal expression of each genome at this time point.

OT3018 is a low B-glucan line that showed a unique expression pattern. At the first time point, 1-3 DPA C- genome expression was the lowest out of any line at 22% but increased by 7-9 DPA to have the highest C-genome expression at 39%. This large fluctuation in C- genome expression

also changed the expression of the D-genome while expression of the A-genome remained relatively constant.

### **CslF6 D-genome marker**

In the second intron of the D genome of CslF6 there was a variable region of a 5 bp repeat of ATGTG. The D-genome allele in ‘Hifi’ had a 5 unit repeat while the ‘SolFi’ allele had a 6 unit repeat. The addition of this marker to the ‘Hifi’ x ‘SolFi’ linkage brought the total number of markers to 403. JoinMap placed the CslF6-D marker in a linkage group with a large number of closely placed markers. The markers adjacent to the CslF6\_D marker are located on chromosome 9D of the *A. sativa* consensus map (personal communication). The flanking markers were G\_ES15\_c10291\_118 and G\_ES01\_c9396\_338. The CslF6-D placement on 9D confirmed the subgenome identification of this ortholog/homeoallele in the D genome.

## **Discussion**

### **Species comparisons**

A tree comparison of CDS sequences was done between the three homeologs of CslF6 in *A. sativa* and available sequences from other species including barley, wheat, and rice (Figure 4). Only a partial CDS for rice was publically available while full- length coding sequences were used for all other species. The three homeologous sequences from *A. sativa* were the most similar to each other. The barley and wheat sequences were the next most similar grouping. The sequences from rice were the most dissimilar when compared to all other species. A comparison of the protein sequences was also done (Figure 5). The relationship between clades was not changed. These results confirm what is known of the evolutionary relationship between these species.

## **Global expression**

‘Marion’ had relatively high expression of CslF6 in all three of the developmental time points.

‘Marion’ is an interesting cultivar in that the  $\beta$ -glucan is deposited uniformly throughout the seed and not concentrated in the endosperm, as is usually seen in *A. sativa* lines (personal communication). A more consistent expression of CslF6 during seed development could account for this difference in  $\beta$ -glucan distribution.

Burton et al. published the transcriptional profiles of the CslF family in barley (2008). In barley the expression of CslF6 was relatively high throughout endosperm development. Their data also showed that expression of CslF6 was highest at 1 and 4 days after pollination but expression decreased to its lowest point by 8 days after pollination. These results are similar to our results, suggesting that expression of CslF6 is the highest for the first six days after pollination. The barley data showed a marked increase in CslF6 expression late in endosperm development peaking 20 days after pollination. Our expression data do not extend to 20 DPA to confirm this trend in *Avena*.

## **Genome- specific expression**

Lower C- genome expression of CslF6 during embryo development was associated with lines with the highest content of B-glucan. High B-glucan lines ‘Hifi’ and OT3044 consistently showed low proportions of C-genome expression during all stages of embryo development usually approximately 25% or below. This observation, combined with information that the amino acid sequence of C genome is different than the A and D, leads us to hypothesize that the enzyme produced by the C-genome transcript of CslF6 is less functional than the enzyme produced by A- and D- genome transcripts.

‘Strimagdo’ is a synthetic hexaploid line with low levels of B-glucan that was made from the cross of *A. sativa* and *A. magna* (Ladizinsky 2000). This line resembles most *A. sativa* lines in that it has a similar chromosome constitution but has not gone through thousands of years of evolutionary change and domestication. The average DPA expression shows that this line has equivalent expression of the D and C genome while the A remains high. The overall expression in most other *A. sativa* lines differs in that expression of the D- and A- genome alleles are about equal while the C genome is much lower. It is possible that domestication of oat has favored a decrease in C- genome expression in order to support an increase in  $\beta$ -glucan production. Low C- genome expression appears to be necessary for high levels of  $\beta$ -glucan production but might not be as crucial after 6 DPA. Many high to average  $\beta$ -glucan producing lines- including ‘Leggett’, ‘Marion’, and 88304 had reached close to equal expression of each genome by 7-9 DPA. The global expression data showed in most cases by 7-9 DPA that overall expression of CslF6 is decreasing. This is also consistent with the report of *H. vulgare* in which expression decreased to its lowest point at 8 days after pollination (Burton et al. 2008).

## Figures

**Table 1.** Sequencing Panel

Species	Name	Genome	Notes
<i>A. sativa</i>	Sun II-1	AACCDD	Monosomic background
<i>A. sativa</i>	Hurdal	AACCDD	Fusarium resistant
<i>A. sativa</i>	Morgan_AC	AACCDD	
<i>A. sativa</i>	Marion (Canada)	AACCDD	High Fiber line; KM map parent
<i>A. sativa</i>	Goslin	AACCDD	
<i>A. sativa</i>	Asencao	AACCDD	
<i>A. sativa</i>	Ajay	AACCDD	High Fiber line
<i>A. sativa</i>	Ogle	AACCDD	KO, OT map parent
<i>A. sativa</i>	Rigodon_AC	AACCDD	
<i>A. sativa</i>	Gem	AACCDD	Low Fiber line
<i>A. sativa</i>	Morton	AACCDD	
<i>A. sativa</i>	Assiniboia/S42	AACCDD	
<i>A. sativa</i>	Buckskin	AACCDD	
<i>A. sativa</i>	HiFi	AACCDD	High Fiber line; SolFI x HIFi map parent
<i>A. sativa</i>	TAM O-301	AACCDD	OT map parent
<i>A. sativa</i>	Coker 227	AACCDD	
<i>A. sativa</i>	Kanota	AACCDD	KO, KM map parent; monosomic background
<i>A. sativa</i>	Kangaroo	AACCDD	
<i>A. sativa</i>	CI 4706-2	AACCDD	BT map parent
<i>A. sativa</i>	Schwartzhafer	AACCDD	Low BG line, stem rust resistant
<i>A. sativa</i>	Tardis	AACCDD	
<i>A. sativa</i>	Leggett	AACCDD	Mid-high BG
<i>A. magna</i> x <i>A. strigosa</i>	Strimagdo	AACCDD	Domesticated magna x Saia
<i>A. magna</i>	BA 13-13	CCDD	Domesticated parent of BAM mapping population
<i>A. magna</i>	#169	CCDD	Wild parent of BAM mapping population
<i>A. strigosa</i>	CI 3815 / PI 131641	AA	Parent of SW mapping population
<i>A. wiestii</i>	CI 1994 / PI 53626	AA	Parent of SW mapping population
<i>A. ventricosa</i>	BYU 709/PI 657338	CC	
<i>A. canariensis</i>	BYU 682/CN 23017	DD	

**Table 2.** Cloning and Sequencing Primers

Primer Name	Sequence 5' -- 3'	Genomic Location
CsIF6_F1	GAGTGAGTGCCTGCATTGAG	-30
CsIF6_F2	CTCCCGCACCGAGAAGATTA	279
CsIF6_R2	GTGGTGACGAAGATGTCGAG	2335
CsIF6_F4	TGCTGATCTTCGTTCTGTCTG	1948
CsIF6_IntronR1	CTCCATTAATGCGGTGGTCT	1561
CsIF6_IntronF2	TGGCGCCGTTTCATAATAAAT	864
CsIF6_R4	CGAGGATGGAGAGCACTGA	2259
CsIF6_F5	CTACGTCTCCGACGACAGTG	2291
CsIF6_R10	CCCAGCATGAAGCAGATACC	3769
CsIF_F11	GTGTCCCATGCTCGTCTAC	3611
CsIF_F12	ACCTCTACGCCAACCACAAC	3835
CsIF_R12	GAGCCCAGGCTTCTGGTACT	4043
CsIF_F14	GATGTTCCGCAAGACCAAGT	4007
CsIF_R14	GATCCCCTCAGCAGTGTGT	4214
CsIF_R15	GCGGGTAGATGGAGCAGTAG	4392
CsIF_F18	TCAAGCTCACATCCAAGCTG	4843
CsIF_R18	ACTATGCATGTGGGGGATGT	5189
CsIF_R20	TATACGTGCCCCAGCATGTA	+250

**Table 3.** Primers for Relative Global Expression of CsIF6

Primer Name	Sequence 5'-- 3'	Genomic Location
CsIF_F16	AGGACGTCGTCACAGGCTAC	4324
CsIF_R16	ACGATGCCCAGGTAGACGTA	4675

**Table 4.** Global Relative Expression Panel

Species	Name	Description	% $\beta$ - glucan
<i>A. sativa</i>	Leggett	Mid-high BG	4.3-5.0
<i>A. sativa</i>	Marion	Mid-high BG	4.7-4.8
<i>A. sativa</i>	Hifi	High BG	7.0-7.3
<i>A. sativa</i>	Schwarzhafer	Low BG, stme rust resistant	2.8-3.0
<i>A. sativa</i>	Strimagdo	Synthetic Hexaploid, <i>A. strigosa</i> x <i>A. magna</i>	3.1
<i>A. strigosa</i>	Strigosa	High BG, high protein	4.9-5.6
<i>A. magna</i>	Bam49-2	Low BG, high protein (hexaploid)	2.2



**Table 5.** RNASeq Differential Expression Panel

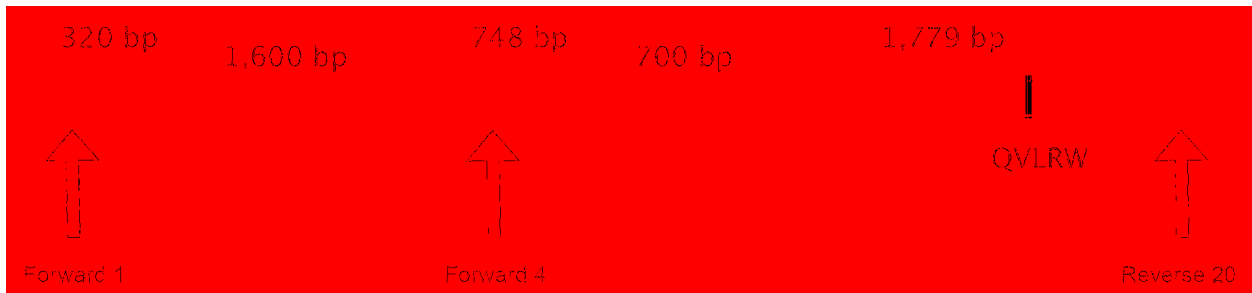
Species	Name	Description	% $\beta$ - glucan
<i>A. sativa</i>	Leggett	Mid-high BG	4.3-5.0
<i>A. sativa</i>	Marion	Mid-high BG	4.7-4.8
<i>A. sativa</i>	Hifi	High BG	7.0-7.3
<i>A. sativa</i>	88304	High BG mutant	5.8
<i>A. sativa</i>	Strimagdo	Synthetic Hexaploid, <i>A. strigosa</i> x <i>A. magna</i>	3.1
<i>A. sativa</i>	OT3044	High BG	4.8
<i>A. sativa</i>	OT3018	Low BG	3.0-3.1

**Table 6.** RNASeq SNP Locations

	A	C	D
Exon 1	-	124, 126	213
Exon 2	990	489	612, 978
Exon 3	2178, 2409, 2493	1455	-



**Figure 1.** Examples of stages of oat embryo development collected for RNA extraction. From left to right: 1-3 days post anthesis (DPA), 4-6 DPA, 7-9 DPA.



**Figure 2.** Diagram of CslF6 genomic sequence with exons in blue and introns in red. Arrows mark locations of primers used for cloning.

## Intron 1 Splice Site



## Intron 2 Splice Site



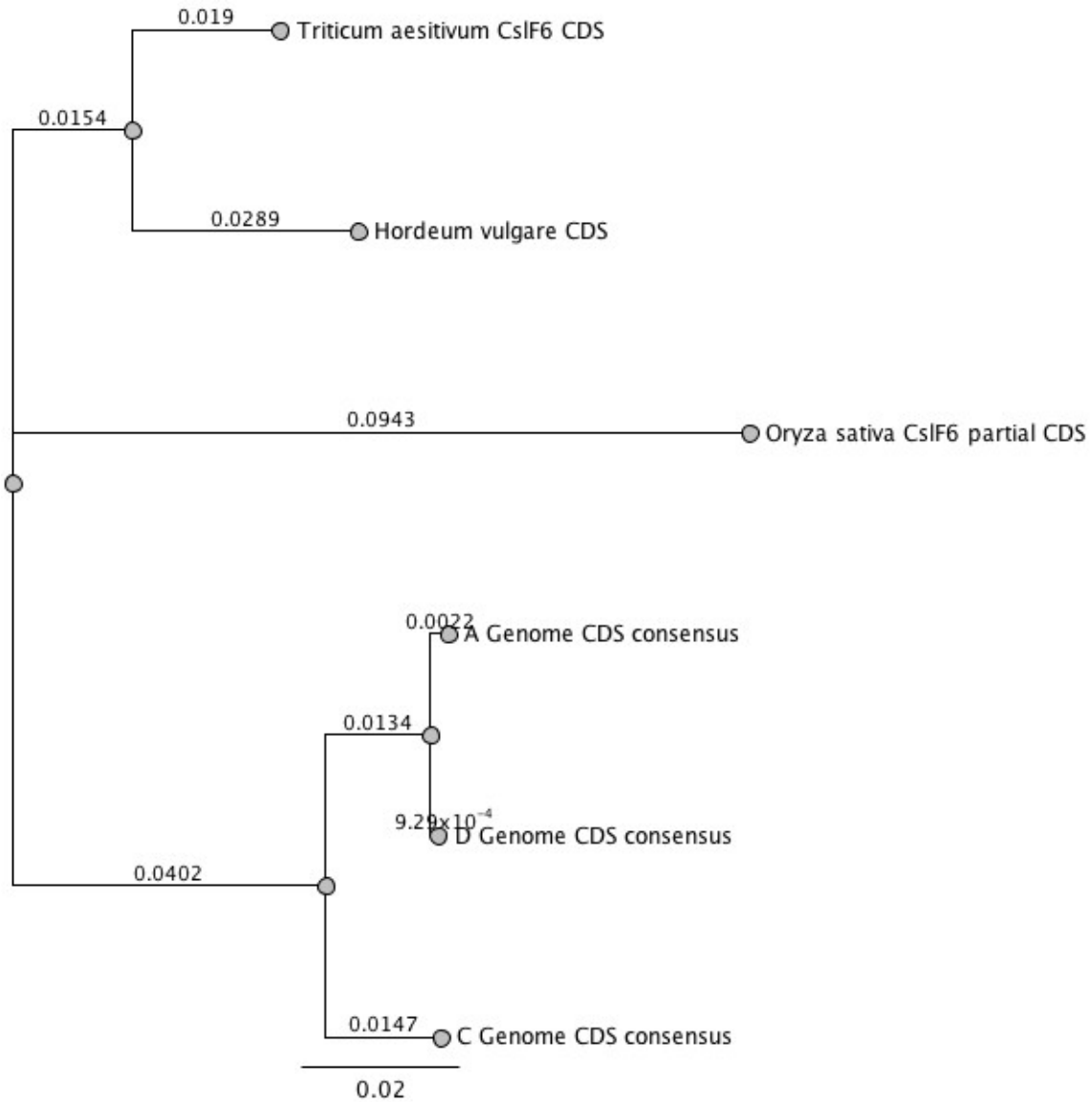
A and D  
Genomes



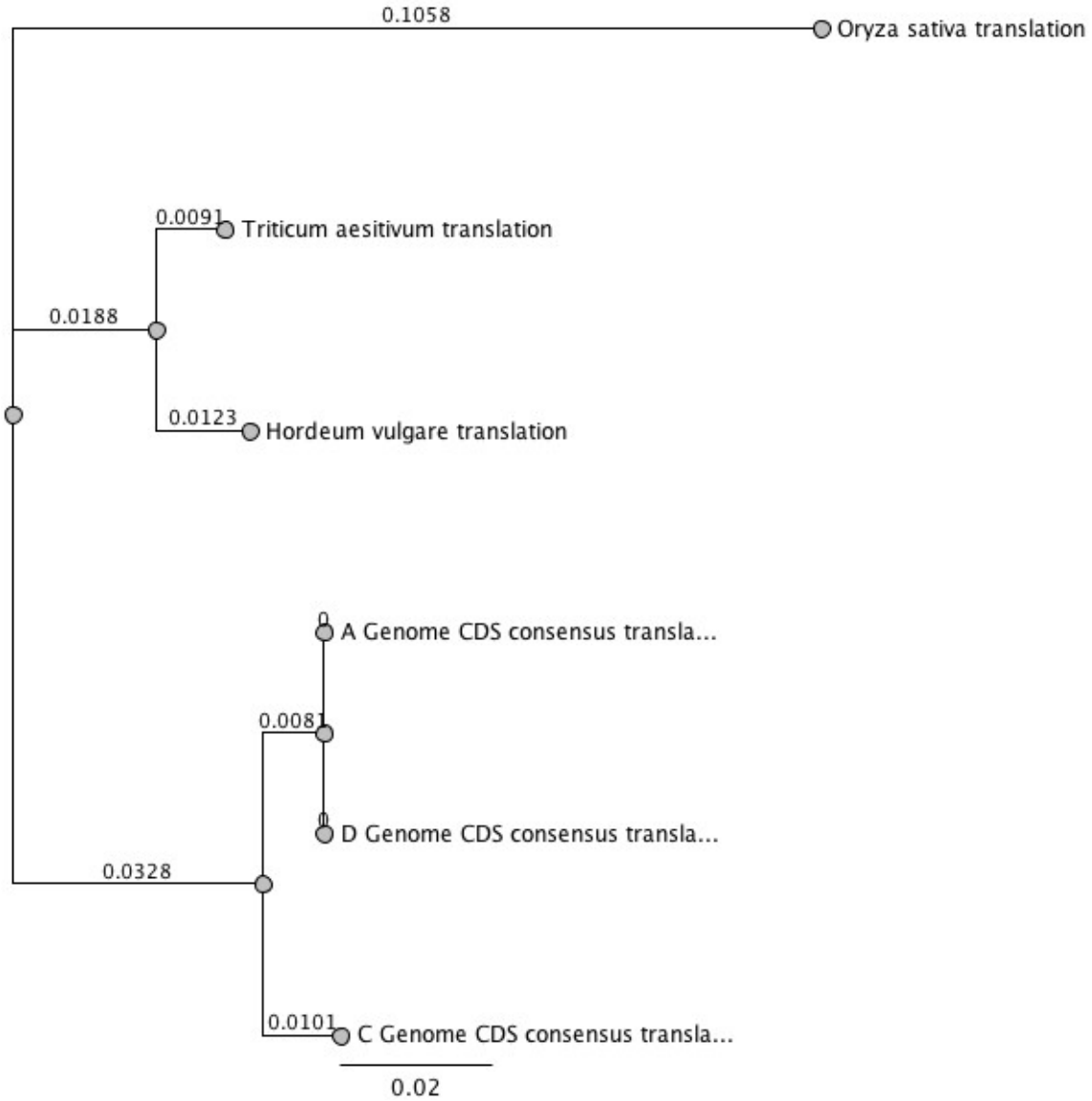
C Genome



**Figure 3.** Splice sites for CslF6. Intron 2 in the A and D genomes could be spliced at 2 sites in the sequence 4 bp apart from each other both resulting in the same coding sequence. For the C genome cutting intron 2 at the first splice site 1 results in a different coding sequence due to the A to G Mutation. This Mutation resulted in an amino acid change from Leucine to an Arginine. Splicing at the second location produces a coding sequence identical to the A and D.

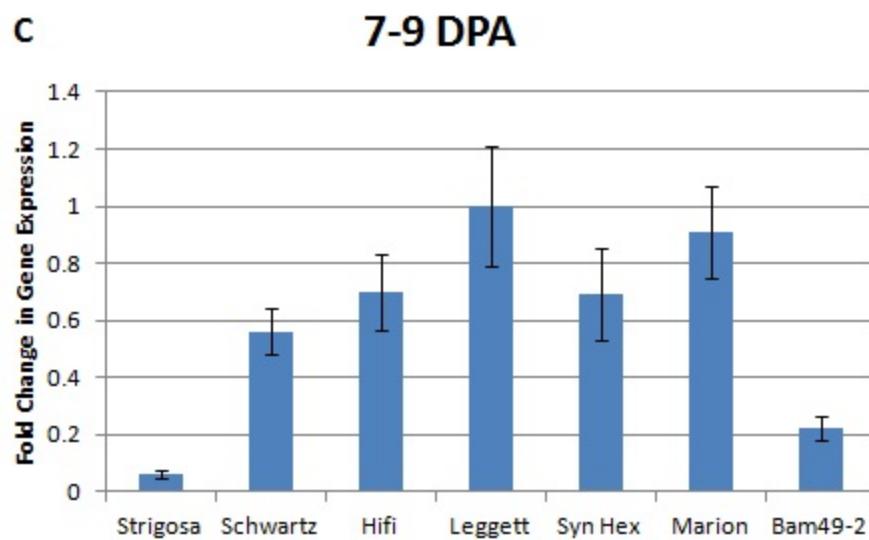
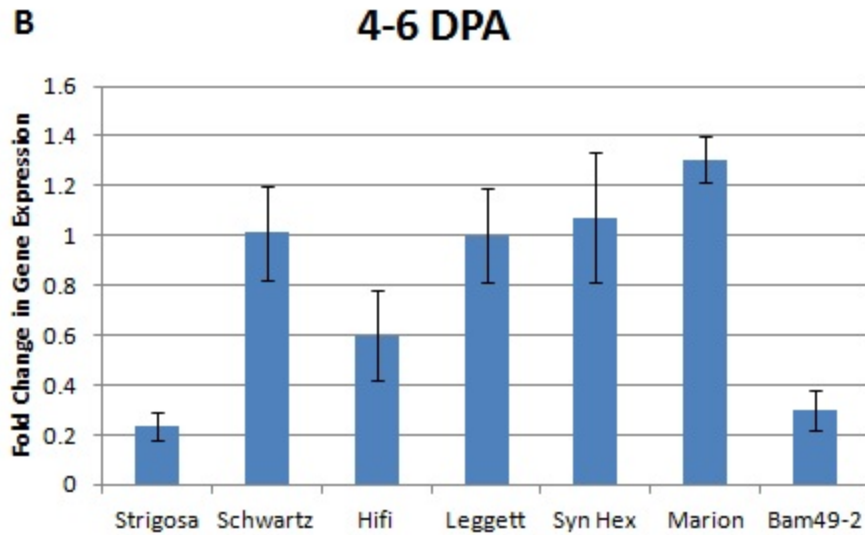
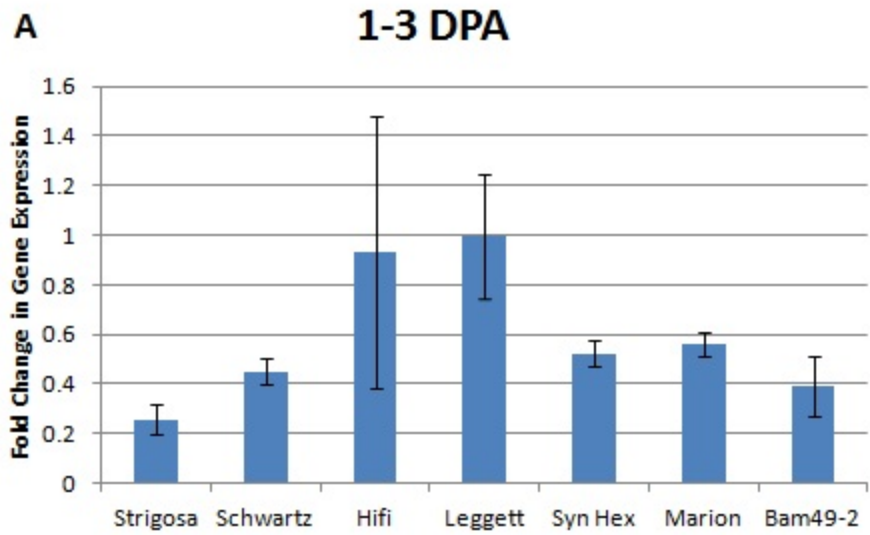


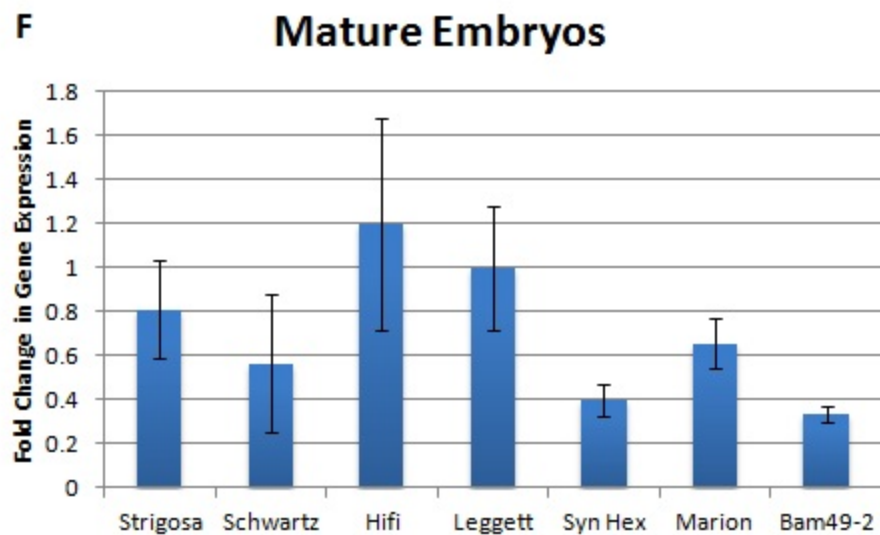
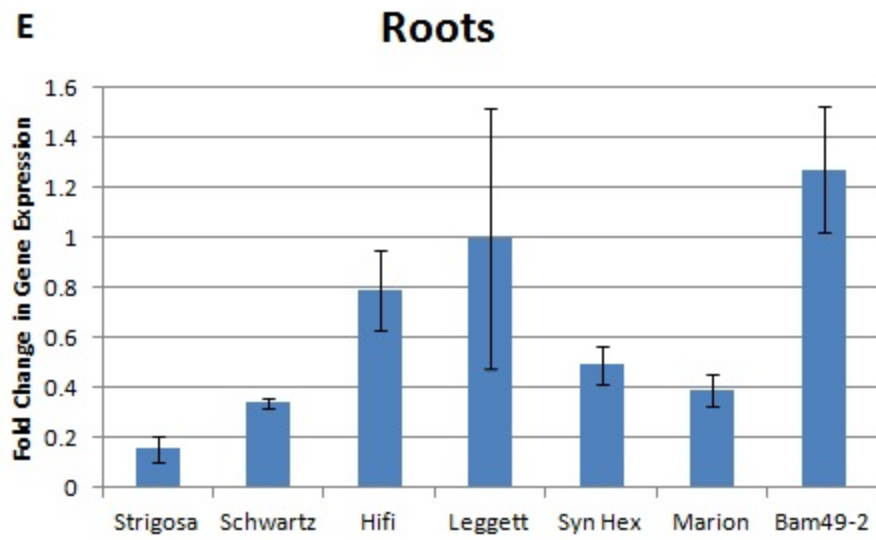
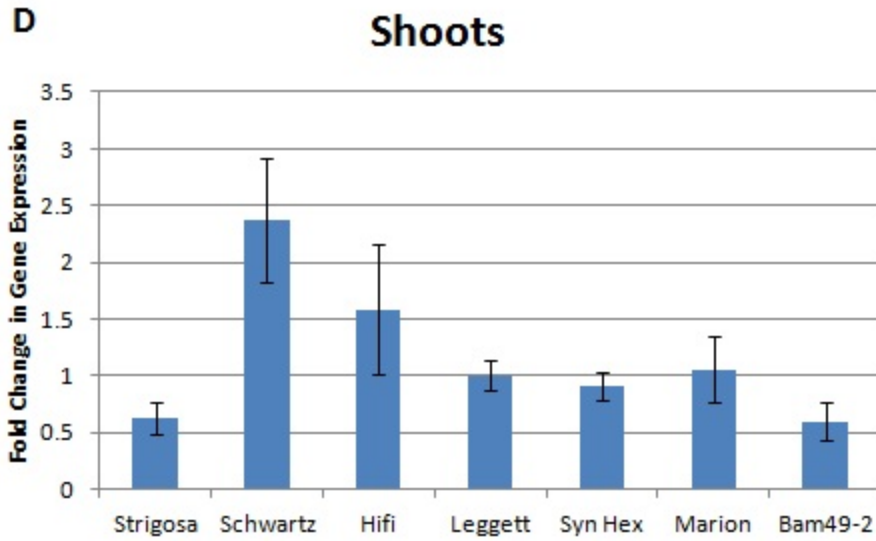
**Figure 4.** Unrooted nucleotide tree displaying comparisons between the three CslF6 sequences in oat and other related species.

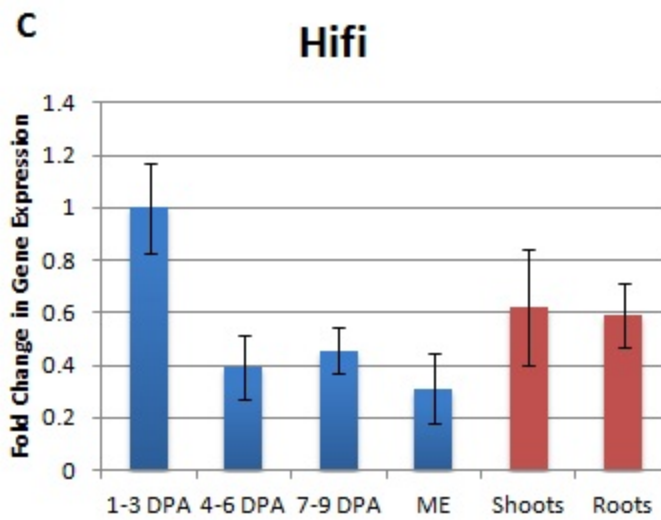
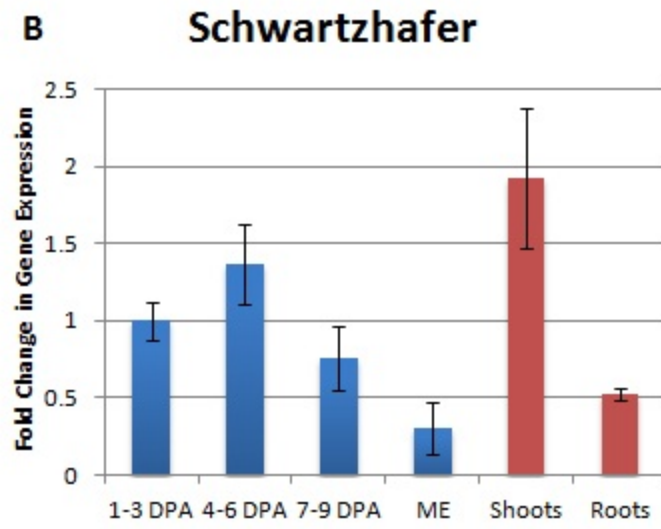
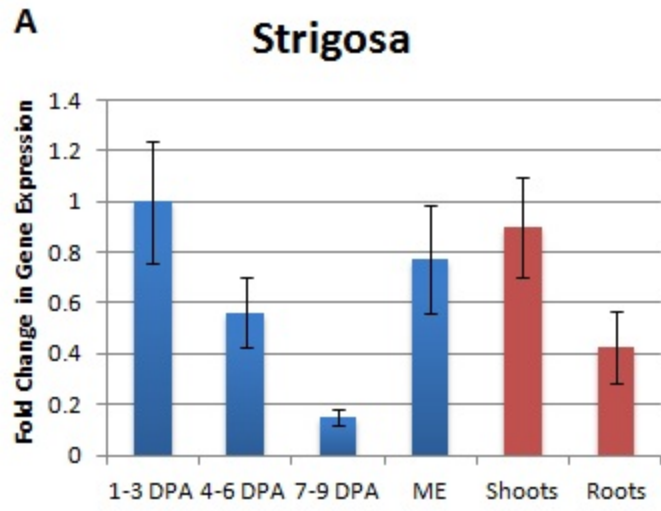


**Figure 5.** Unrooted tree showing comparisons of translated sequences of CsIF6 between the three alleles of oat and other related species.

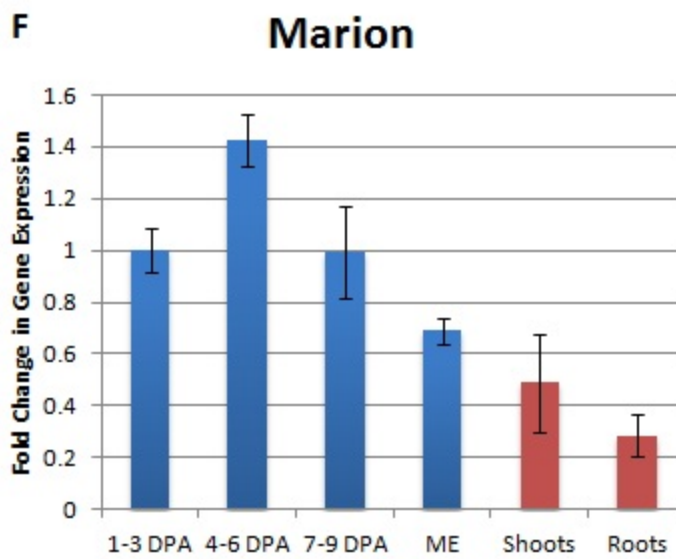
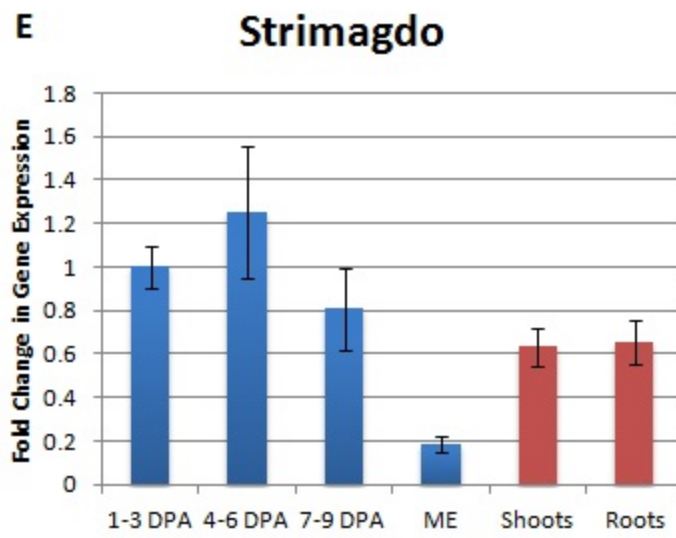
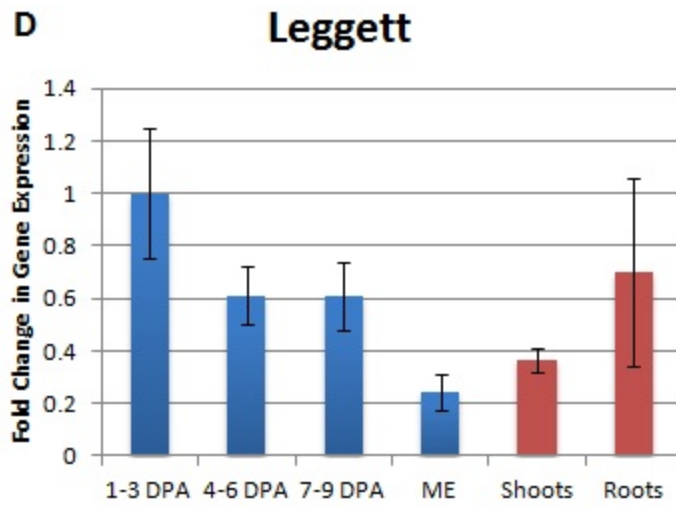
**Figure 6.** Relative global expression of CsIF6 using sybr green chemistry. Comparisons of all lines at each time point using Leggett was as calibrator. Data are presented as fold changes in gene expression normalized to beta actin and relative to calibrator line. 1-3 DPA (A). 4-6 DPA (B). 7-9 DPA (C). Shoots (D). Roots (E.) Mature Embryos (F). Error bars are SD

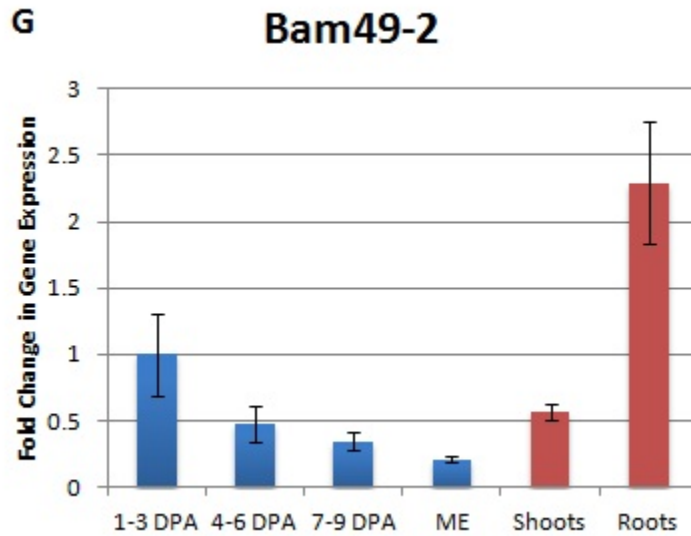






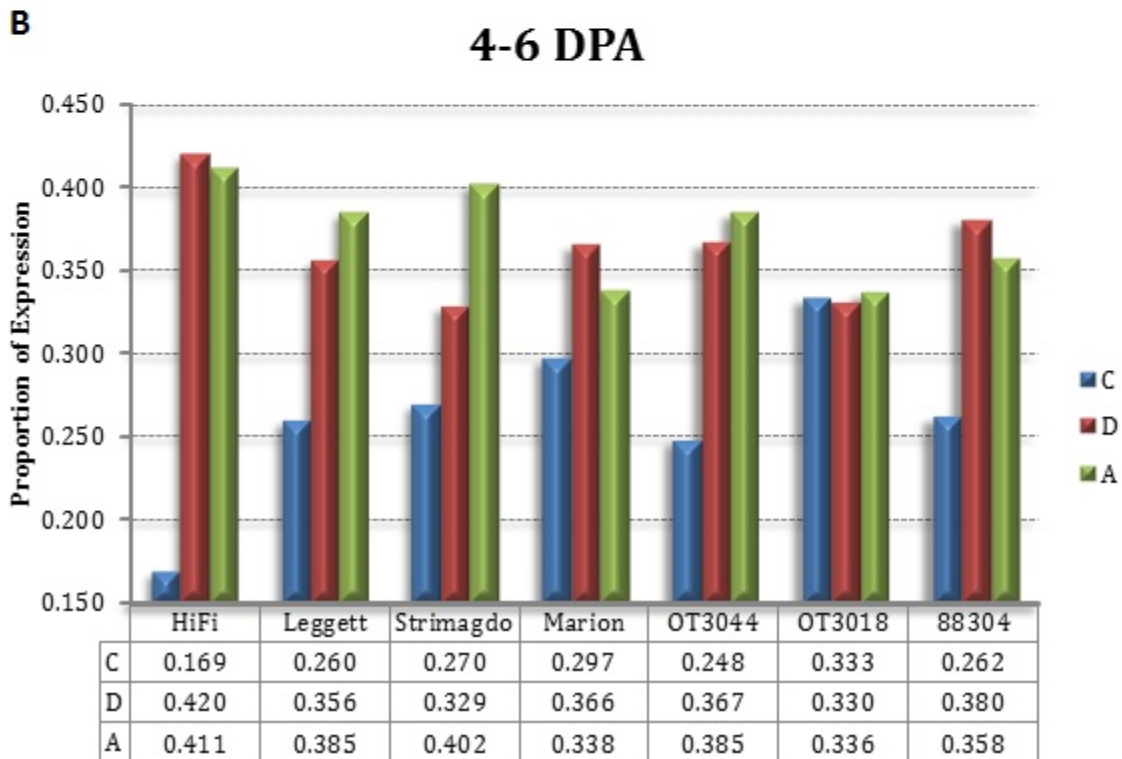
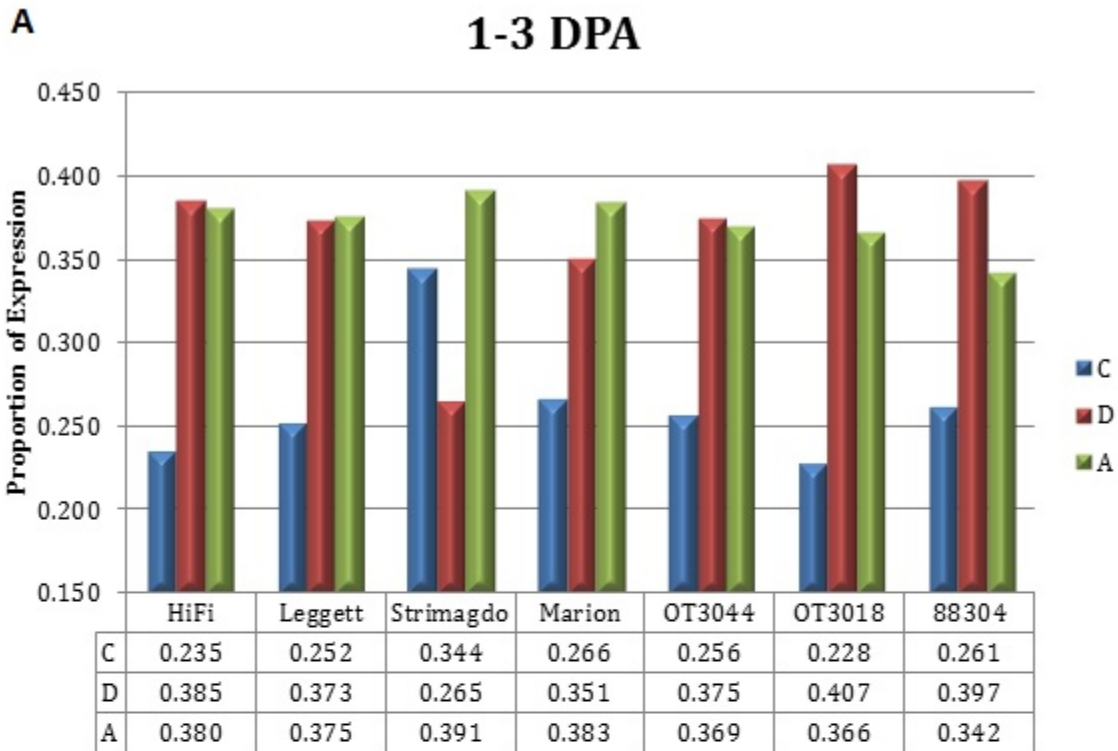






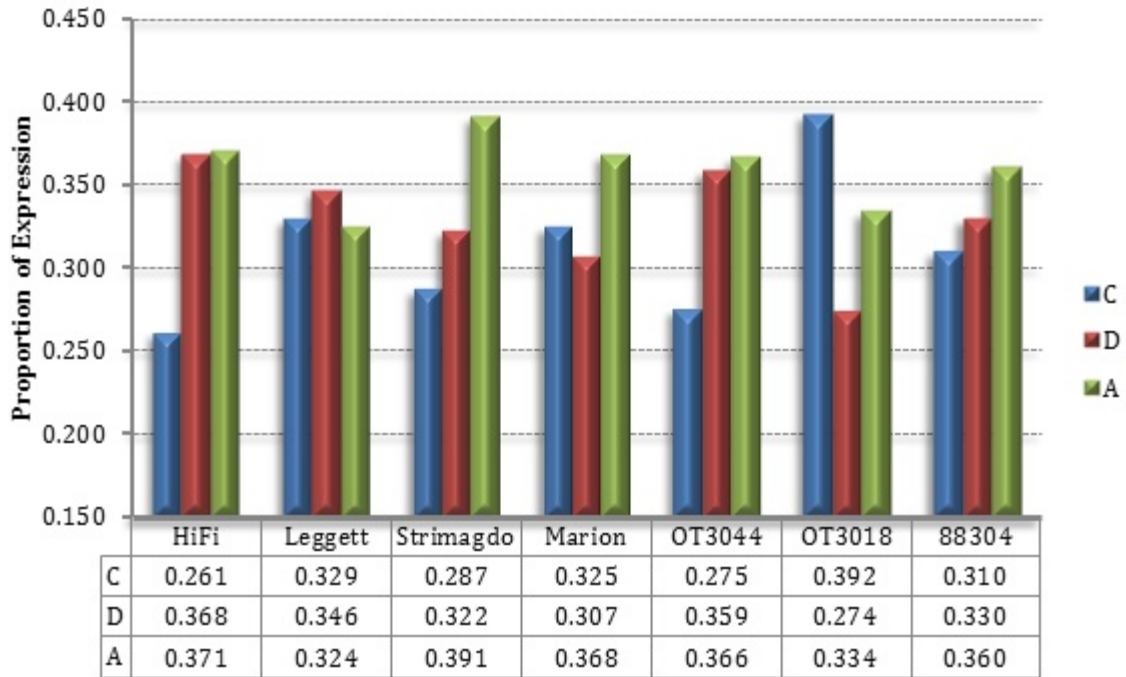
**Figure 7.** Relative global expression of CslF6 using sybr green chemistry. Comparisons across all time point with in each line using time point 1-3 DPA as calibrator. Data are presented as fold changes in gene expression normalized to beta actin and relative to calibrator line. Strigosa (A). Schwartzhafer (B). ‘Hifi’ (C). Leggett (D). Strimagdo (E). Marion (F). Bam49-2 (G). Error bars are SD.

**Figure 8.** Genome Specific Expression using RNAseq. Data are reported as a proportion of expression contributed to each allele. 1-3 DPA (A). 4-6 DPA (B). 7-9 DPA (C). Average DPA (D).



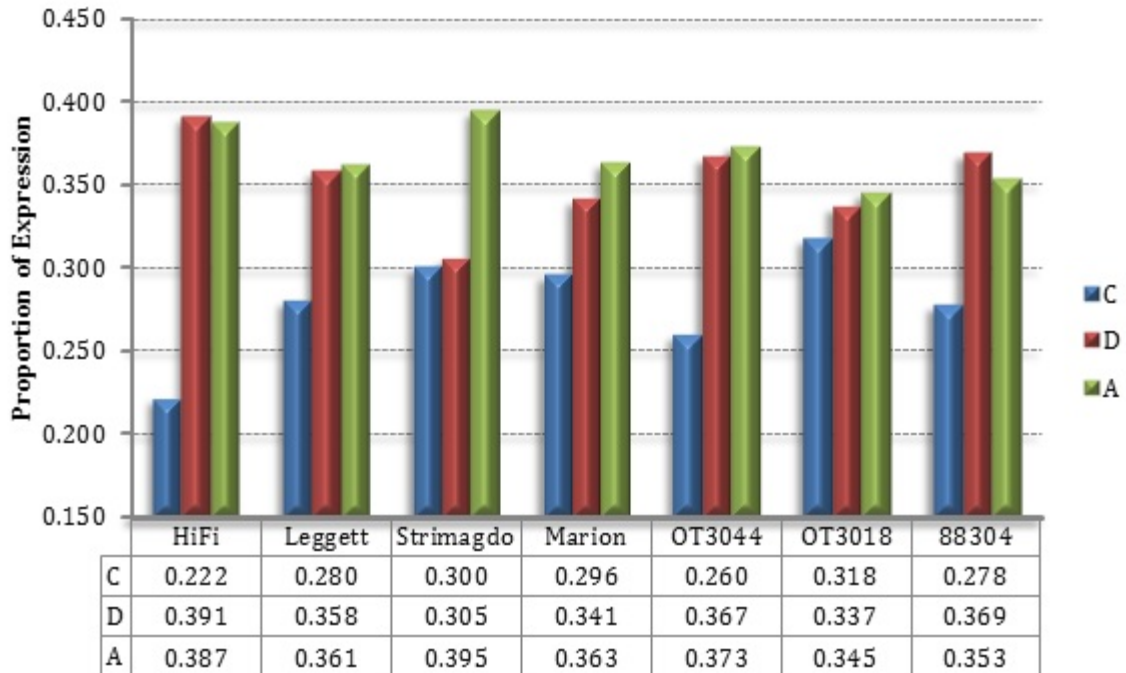
C

### 7-9 DPA



D

### Average DPA



## References

- Anderson JW, Story L, Sieling B, Chen WL, Petro MS, Story J (1984) Hypocholesterolemic effects of oat-bran or bean intake for hypercholesterolemic men. *Am J Clin Nutr* 40:1146-1155
- Braaten JT, Wood PJ, Scoot FW, Wolynetz MS, Lowe MK, Bradley-White P, Collins MW (1994) Oat B-glucan reduces blood cholesterol concentration in hypercholesterolemic subjects. *Eur J Clin Nutr* 48:465-474
- 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 Biotech J* 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. *Mol 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 Physiol* 146:1821-1833
- Burton RA, Wilson SM, Hrmova M, Harvey AJ, Shirley AM, Stone BA, Newbigin EJ, Bacic A, Fincher GB (2006) Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)- $\beta$ -D-glucans. *Science* 311:1940-1942
- Carpita NC (1996) Structure and biogenesis of the cell walls of grasses. *Annu Rev Plant Physiol Plant Mol Biol* 47:445-447
- Collins HM, Burton RA, Topping DL, Liao ML, Bacic A, Fincher GB (2010) Variability in fine structure of noncellulosic cell wall polysaccharides from cereal grains: potential importance in human health and nutrition. *Cereal Chem* 87:272-282
- Farrokhi N, Burton RA, 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 Biotech J* 4:145-167
- FDA (1997) 21 CFR Part 101. Food labeling, health claims: soluble dietary fiber from certain foods and coronary heart disease. *Federal Register* 62 15:3584-3601
- Fincher GB (2009a) Exploring the evolution of (1,3;1,4)- $\beta$ -D-glucans in plant cell walls: comparative genomics can help! *Curr Opin Plant Biol* 12:140-147
- Fincher GB (2009b) Revolutionary times in our understanding of cell wall biosynthesis and remodeling in the grasses. *Plant Physiol* 149:27-37

Fincher GB, Stone BA (2004) Chemistry of nonstarch polysaccharides. In: Wrigley C, Corke H, Walker CE (eds) Encyclopedia of Grain Science, Elsevier Academic Press, Oxford, pp 206-223

Hesselman K, Aman P (1986) The effect of  $\beta$ -glucanase on the utilization of starch and nitrogen by broiler chickens fed on barley of low- or high-viscosity. *Anim Feed Sci Technol* 15:83-93

Kirby RW, Anderson JW, Sieling B, Rees ED, Chen WL, Miller RE, Kay RM (1981) Oat-bran intake selectively lowers serum low-density lipoprotein cholesterol concentrations of hypercholesterolemic men. *Am J Clin Nutr* 34:824-829

Langmead B, Cole T, Pop M, Salzberg S. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Bio* 10:R25

Marshall HG, Sorrells ME (1992) Oat science and technology. American Society of Agronomy Inc. & Crop Science Society of America Inc., Madison

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

Shimizu C, Kihara M, Aoe S, Araki S, Ito K, Hayashi K, Watari J, Sakata Y, Ikegami S (2008) Effect of high  $\beta$ -glucan barley on serum cholesterol concentrations and visceral fat area in Japanese men- a randomized, double-blinded, placebo controlled trial. *Plant Foods Hum Nutr* 63:21-25

Schulz MH, Zerbino DR, Vingron M, Birney E (2012) Oases: Robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28(8):1086-92

Wood PJ (2007) Cereal  $\beta$ -glucans in diet and health. *J Cereal Sci* 46:230-238

## Chapter 2: Review of (1,3;1,4)- $\beta$ -D-glucan Synthesis in Cereals

### Introduction

*Avena sativa* L. ( $2n = 6x = 42$ ) or common oat is the sixth most important cereal crop worldwide behind wheat, maize, rice, barley and sorghum. The area dedicated to oats has fallen sharply in the past decades from 38.2 million hectares in 1960 to just 9 million in 2010 (Faostat). This downward trend can be contributed in part to draft horses being replaced by mechanized farm equipment. Oats have also suffered due to increased production of alternate crops that are more profitable such as corn and soybean. The rapid rise in soybean production in the United States has diminished the value of oats as a protein source in feed rations (Welch 1995).

### Oat production

Oats grow best in cool moist climates and are sensitive to hot dry weather particularly from head emergence to maturity. For these reasons, oat production is concentrated between latitudes 35°-50° N including Finland and Norway, and 20°-40° S. Canada, Russia, Australia and the United States are the leading producers of oat grain. Oats can be grown on many soil types, but produce better on acidic soil than other small grains. Most of the world's production comes from spring-sown cultivars because oats lack winter hardiness (Marshall and Sorrells 1992; Welch 1995). Diploid oat *A. strigosa* L. ( $2n = 2x = 14$ ) is increasingly being planted as a forage and winter cover crop in subtropical and temperate areas of Brazil, Uruguay, Argentina and Chile (Reynolds 2004).

## **Oat usage and consumption**

Approximately 75% of oat grain in the world is consumed as animal feed. Oats are fed to dairy cattle, horses, mules and turkeys and to lesser extent to beef cattle, hogs and sheep (Marshall 1992). Oats are the preferred feed of horses and breeding stock of cattle- or animals that need to be kept in optimal health for long periods of time (Welch 1995). Oats are a good source of protein, fiber and minerals and have a greater protein content than corn per pound but with fewer calories. The lower energy content of oats makes it less effective than corn at fattening livestock cattle (Welch 1995).

The second most common use for oat grain is for food. Oats are commonly used in a wide variety of breakfast cereals (Welch 1995). These include porridge type hot cereals, muesli, granola and extruded cereal products such as Cheerios®. The use of oat grain in specialty bakery products is becoming increasingly popular. Processed oat products are also found in baby foods, meat extenders, thickeners and stabilizers. Oats are increasingly being incorporated in more products due to the hypocholesterolemic properties of (1,3;1,4)- $\beta$ -D-glucan found in oats.

Oats have also found a niche in the cosmetics and supplement industry. Oat extracts have been found to have anti-itch properties and have been incorporated into many skin products. American skin care company Aveeno®, named for the scientific name of oats, produces a wide variety of products that contain colloidal oatmeal and oat oil. Oat hulls are also used in the industrial production of chemicals furfural and furan. These chemicals have excellent solvent properties and are an important intermediate in the manufacture of various chemicals (Welch 1995).



### **(1,3;1,4)- $\beta$ -D-glucan**

(1,3;1,4)- $\beta$ -D-glucan is an unbranched, unsubstituted, non-cellulosic plant cell wall hemicellulose. (1,3;1,4)- $\beta$ -D-glucan is found primarily and almost exclusively in grasses and cereals of the Poaceae (Carpita 1996; Fincher and Stone 2004). (1,3;1,4)- $\beta$ -D-glucan is a major component of cells walls in the starchy endosperm and aluerone layer of the grain. (1,3;1,4)- $\beta$ -D-glucan is also found in vegetative tissues, especially in growing cells of young developing tissue such as the coleoptile and the base of young leaves, while being virtually absent in mature tissues (Burton and Fincher 2009; Wilson et al. 2006).

Within the Poaceae there are differences in (1,3;1,4)- $\beta$ -D-glucan content. Barley (*Hordeum vulgare* L.), oat (*Avena sativa* L.), and rye (*Secale cereale* L.) are all rich sources while wheat (*Triticum aestivum* L.), maize (*Zea mays* L.), and rice (*Oryza sativa* L.) produce low levels of the polysaccharide (Fincher and Stone 2004, Welch 2000). In the starchy endosperm of barley and oat (1,3;1,4)- $\beta$ -D-glucan can account for up to 70% by weight of the walls, while in wheat and rice it accounts for only approximately 20% (Fincher and Stone 2004).

### **Fine Structure and Function**

(1,3;1,4)- $\beta$ -D-glucan consists of (1,4) and (1,3)- $\beta$ -D- glucosyl residues in a 2.2-2.6:1 ratio. The degree of polymerization can be greater than 1000 units (Fincher 2009b). The two linkage types are not arranged in regular, repeating patterns, but neither are they randomly arranged. (1,3)- $\beta$ -D-glucosyl residues occur singly between two or more (1,4)- $\beta$ -D-glucosyl residues. In cereals (1,3)- $\beta$ -D-glucosyl residues never occur adjacent to each other and are rarely separated by just one (1,4)- $\beta$ -D-glucosyl residue. (1,4)- $\beta$ -D-glucosyl residues are usually seen as (1,4)-oligoglucosyl residues or cellodextrin units consisting of two or three (1,4) linkages. Longer cellodextrin units ranging from 5-20 (1,4) linkages occur and can account for 10% of the

polysaccharide chain. The ratio of celotriosyl [two (1,4) linkages] to celotetraosyl [three (1,4) linkages] can vary from 1.8-4.5:1 among species (Collins et al. 2010). The introduction of (1,3) linkages in the otherwise linear cellulosic chain introduces molecular kinks. The irregular distribution of celotetraosyl and celotriosyl linkages creates irregularly spaced molecular kinks (Fincher and Stone 2004; Burton and Fincher 2009; Fincher 2009b). These asymmetric polysaccharide molecules are unable to align over extended regions and thus cannot aggregate into well structured microfibrils. The greater the irregularity in the distribution of linkages the more asymmetric the polysaccharide will be, further increasing its solubility. These polysaccharide chains remain in aqueous solution and form a gelatinous matrix. The extended celodextrin units allow for (1,3;1,4)- $\beta$ -D-glucan to interface with cellulose microfibrils and other cell wall polysaccharides. The gelatinous matrix provides a degree of additional support to a growing cell wall while remaining sufficiently flexible and porous to permit the flow of water and other small molecules and nutrients needed for cell growth (Fincher 2009a). This function agrees with the distribution of (1,3;1,4)- $\beta$ -D-glucan in growing cells.

(1,3;1,4)- $\beta$ -D-glucan also serves as a storage polysaccharide. High levels of this polysaccharide are present in the starchy endosperm and aleurone layers of grain, contributing to the total amount of glucose stored in the seed. During germination the glucose in (1,3;1,4)- $\beta$ -D-glucan will be utilized by the growing seedling. It has also been shown that (1,3;1,4)- $\beta$ -D-glucan can be metabolized as an energy source in vegetative tissue during periods of insufficient glucose. Barley seedlings transferred to the dark showed increased levels of  $\beta$ -D-glucan glucohydrolase and endohydrolase enzymes, consistent with (1,3;1,4)- $\beta$ -D-glucan being metabolized for glucose. After two days in the dark a 30% decrease in (1,3;1,4)- $\beta$ -D-glucan was

observed in leaf tissue (Roulin et al. 2002). (1,3;1,4)- $\beta$ -D-glucan also contributes to a stiffer grain texture.

### **Economic Importance**

(1,3;1,4)- $\beta$ -D-glucan is an important component of soluble dietary fiber and has been recognized for its role in decreasing the risk of many serious human health conditions including high blood serum cholesterol, obesity, non-insulin dependent diabetes, and colorectal cancer (Wood 2007; Braaten et al. 1994; Anderson et al. 1984; Shimizu 2008). Starting in 1997 the Federal Drug Administration (FDA) authorized a health claim on the relationship between soluble fiber in whole oats accompanied by a diet low in saturated fat with a reduced risk of coronary heart disease. Products that contain at least .75 grams of (1,3;1,4)- $\beta$ -D-glucan per serving are authorized to advertise this health claim on their packaging (FDA 1997). Raised plasma low-density lipoprotein (LDL) cholesterol is a major but modifiable risk for cardiovascular disease. Cereals high in (1,3;1,4)- $\beta$ -D-glucan can effectively lower total and LDL cholesterol. (1,3;1,4)- $\beta$ -D-glucans and other soluble fibers are thought to modulate digestion through increasing the viscosity of the aqueous media in the intestine. The high viscosity is believed to slow the flow of digesta and decrease the absorption of fat and cholesterol, while impeding the reabsorption of bile acids (Collins et al. 2010). The absorption of other nutrients including glucose can be slowed, leading to a lowering of the glycemic response. The lowering of the glycemic response can be a benefit in managing and reducing the risk of diabetes.

Conversely (1,3;1,4)- $\beta$ -D-glucan is considered an antinutritive for the feed of monogastric animals such as pigs and poultry, resulting in slower weight gain. The high viscosity in the gut conferred by (1,3;1,4)- $\beta$ -D-glucan can prevent nutritive uptake. Gut viscosity also leads to sticky droppings, creating potential microbial infection issues (Hesselman and Aman

1986; Fincher and Stone 2004). High levels of (1,3;1,4)- $\beta$ -D-glucan are also not considered favorable in malting and brewing processes due to their delaying effects on endosperm modification.

It has been estimated that 20% of the land surface of earth is currently covered by species from the family Poaceae (Fincher 2009). These include some of the world's most important crops such as rice, maize, wheat, barley, sugar cane (*Saccharum officinarum* L.), and oat. Considering the importance of these crops and the large impact (1,3;1,4)- $\beta$ -D-glucan has on their potential use, much research has been done to elucidating the genetics of (1,3;1,4)- $\beta$ -D-glucan synthesis. This knowledge can be applied to the development of specialized oat and other cereal varieties with high or low (1,3;1,4)- $\beta$ -D-glucan content.

### **Biosynthesis**

The cellulose synthase like (*Csl*) family of genes has been implicated in the biosynthesis of (1,3;1,4)- $\beta$ -D-glucans. There are currently nine subfamilies of *Csl*, designated as *CslA* through *CslJ*, with *CslI* omitted (Farrokhi 2005; Fincher 2009a). The *CslF*, *CslH* and *CslJ* families of genes are unique to the Poaceae. The *CslF* and *CslH* families are found in all investigated species within the Poaceae while the *CslJ* subfamily is only seen in some of the grasses including barley, wheat, and maize, but not in rice or *Brachypodium* (Fincher 2009b; Hazen et al. 2002). The *CslJ* group was formally proposed to be recognized as a new subgroup by Fincher in 2009. It is thought that these three groups of genes could possibly be involved in (1,3;1,4)- $\beta$ -D-glucan synthesis because they are grass-specific gene subfamilies.

Comparative genomics first indicated that the cellulose-synthase gene family F (*CslF*) was involved in synthesis of (1,3;1,4)- $\beta$ -D-glucans. A quantitative trait locus (QTL) with a large effect on (1,3;1,4)- $\beta$ -D-glucan content was mapped to barley chromosome 2H. DNA markers

flanking this region were used to locate a syntenic region of about 3.5 megabases on chromosome 7 of rice. Examination of this corresponding region in rice revealed a cluster of six *CsIF* genes, namely *OsCsIF1-4* and *OsCsIF8-9* (Burton et al. 2006). Constructs were created that contained full-length open reading frames of *OsCsIF2*, *OsCsIF4*, and *OsCsIF9* and were used to transform *Arabidopsis* plants. *Arabidopsis* plants do not produce endogenous (1,3;1,4)- $\beta$ -D-glucan nor are *CsIF* genes present in their genome. This allows for a gain-of-function experiment in which deposition of (1,3;1,4)- $\beta$ -D-glucan could be attributed to the function of *CsIF* genes. Immunogold labeling using monoclonal antibodies against barley (1,3;1,4)- $\beta$ -D-glucan revealed that transformants began to produce (1,3;1,4)- $\beta$ -D-glucans in their cell walls. Generally low levels of (1,3;1,4)- $\beta$ -D-glucan were detected in cell walls despite high *OsCsIF* transcript levels. This finding suggests that there are other limiting components that could be inhibiting increased levels of (1,3;1,4)- $\beta$ -D-glucan synthesis and or transfer to the cell wall. No other phenotypic differences were reported in the transformants.

A similar gain-of-function experiment was used to demonstrate that the *HvCsIH* subfamily from barley could also mediate (1,3;1,4)- $\beta$ -D-glucan synthesis (Doblin et al. 2009). Barley and other members of the Poaceae subfamily Pooideae only contain one *CsIH* gene. Other Poaceae subfamilies contain more than one *CsIH* gene (Fincher 2009b). Constructs containing the single *HvCsIH1* gene under control of the 35S promoter were used to transform *Arabidopsis*. Immunogold labeling revealed that (1,3;1,4)- $\beta$ -D-glucan was detectable in cell walls of transgenic lines, although different labeling patterns were observed among positive transgenic lines (Doblin et al. 2009). In agreement with *CsIF*, *CsIH* transformants did not exhibit any phenotypic differences from the wild type. The lack of detectable phenotype could be due to the relatively low levels of (1,3;1,4)- $\beta$ -D-glucan produced. Enzymatic digestion of (1,3;1,4)- $\beta$ -D-

glucan from transgenic *Arabidopsis* plants revealed that there were differences in linkage ratios as compared to wild type barley. Specifically, transgenic (1,3;1,4)- $\beta$ -D-glucan contained high levels of (1,4)- $\beta$ -D-glucosyl residues between two (1,3)- $\beta$ -D-glucosyl residues; this G3G4G3G4 pattern is seen in very low abundance in the cereals. This study also showed that the *HvCslH1* protein was localized to the endoplasmic reticulum (ER) and to Golgi-derived vesicles but not at the plasma membrane.

*CsIF* and *CsIH* families of genes are both responsible for (1,3;1,4)- $\beta$ -D-glucan synthesis but probably do so independently of one another. No significant transcriptional correlation was seen between these two gene groups in barley tissue to indicate they were possibly involved in the same complex (Doblin et al. 2009). The observation that *CsIF* and *CsIH* genes were able to support (1,3;1,4)- $\beta$ -D-glucan production alone could support the hypothesis that initial evolution of (1,3;1,4)- $\beta$ -D-glucan only required one novel gene. This could also mean that (1,3;1,4)- $\beta$ -D-glucan synthesis might utilize preexisting cellular machineries, such as the *CESA* enzymes involved in cellulose synthesis. Other members of these gene families are believed to have evolved later, allowing for variation in fine structure and function. It should be mentioned that no gene had been identified as the first novel gene. It has been proposed that (1,3;1,4)- $\beta$ -D-glucan synthesis in other families evolved independently via convergent evolution (Fincher 2009a).

Previously there has been some dispute as to the cellular location of (1,3;1,4)- $\beta$ -D-glucan synthesis. Most non-cellulosic polysaccharides are known to be synthesized at the Golgi and exported to the plasma membrane via exocytosis. Isolates of Golgi membranes have been able to synthesize (1,3;1,4)- $\beta$ -D-glucan with the addition of UDP-Glc. But prior to July 2010 (1,3;1,4)- $\beta$ -D-glucan monoclonal antibodies had failed to locate (1,3;1,4)- $\beta$ -D-glucan at the Golgi despite heavy labeling at the cell wall (Burton and Fincher 2009). Carpita and McCann demonstrated

that (1,3;1,4)- $\beta$ -D-glucan was in fact synthesized at the Golgi in maize (2009). Monoclonal antibodies were able to detect (1,3;1,4)- $\beta$ -D-glucan at the periphery of the Golgi and in exocytotic vesicles bound for the plasma membrane. Golgi membranes isolated from maize seedlings pulse labeled with [ $^{14}\text{C}$ ]O $_2$  were found to contain (1,3;1,4)- $\beta$ -D-glucan. Further, when the pulse was followed by a chase period, labeled polysaccharides were seen to be depleted from the Golgi and transferred to the plasma membrane.

### **Cellulose Synthase-Like F6 (CslF6)**

Of the seven *HvCslF* genes identified *CslF6* has emerged as a possible major player. *HvCslF6* transcripts were relatively abundant in all barley tissues examined, especially in developing tissues such as the first basal leaf and coleoptiles. Transcripts of the other *HvCslF* genes were less much less abundant than *HvCslF6*, but were increased in developing tissues as well (Burton et al. 2008). Publicly available expressed sequence tag (EST) databases also confirm that *HvCslF6* transcripts predominate. On the protein level all *HvCSLF* enzymes appear to be similar in size and structure, with the exception that *HvCSLF6* has a unique 54 amino acid loop in the cytosolic region. The functional significance of this feature has yet to be determined (Burton et al. 2008).

Down-regulation of the *CslF6* gene in wheat using RNA interference (RNAi) resulted in an average reduction of (1,3;1,4)- $\beta$ -D-glucan content of 42.4% in mature wheat grain (Nemeth et al. 2010). When 100 grain weights of five transgenic lines were compared to the wild type control no significant difference was found ( $P = 0.14$ ). The authors concluded that reduction in (1,3;1,4)- $\beta$ -D-glucan content does not have a significant effect on grain filling.

Immunolocalization confirmed that the reduction in (1,3;1,4)- $\beta$ -D-glucan that occurred in transgenic plants was localized to the starchy endosperm while having little effect on the

aleurone layers. The effects of decreased (1,3;1,4)- $\beta$ -D-glucan on plant growth were not reported (Nemeth et al. 2010). As mentioned previously wheat contains lower levels of (1,3;1,4)- $\beta$ -D-glucan, contributing only 20% to endosperm cell walls as compared to the >70% that is present in oat and barley. Given that (1,3;1,4)- $\beta$ -D-glucan is not the major polysaccharide of wheat endosperm cell walls, other polysaccharides including arabinoxylan and glucomannans could compensate for the loss of (1,3;1,4)- $\beta$ -D-glucan and still maintain grain integrity.

Burton et al. conducted an experiment in which specific HvCslF genes were over-expressed in transgenic barley (2011). Of particular interest was their report that transgenic barley plants over-expressing the barley CslF6 gene under the control of an endosperm-specific oat globulin promoter resulted in a greater than 80% increase in (1,3;1,4)- $\beta$ -D-glucan content in the grain. When CslF6 was placed under control of the constitutive Pro35S promoter transgenic plants experienced an increase in (1,3;1,4)- $\beta$ -D-glucan in the leaves while levels the grain remained similar to control plants. Leaves of Pro35S:CslF6 seedlings were thicker than control leaves and that their cells were generally larger than those in control leaves.

A monofactorial recessive mutant gene for (1,3;1,4)- $\beta$ -D-glucanless barley grain was created using ethyl methanesulfonate (EMS) and was designated as *bgl* for (1,3;1,4)  $\beta$ -D-glucanless grain (Tonooka et al. 2008). Near-isogenic lines (NILs) were created with this mutant using Nishinohoshi (Ni), a Japanese two-rowed hulled barley cultivar, as the recurrent parent to test its effect on grain quality and chemistry. The Ni(*bgl*) NILs grew normally in the field and showed normal seed fertility, but had shorter culms and awns. Some of the plants also displayed partial chlorosis, especially on the awns. This observation is consistent with (1,3;1,4)- $\beta$ -D-glucan's function in growing cells. Chemical analysis of the barley grain revealed that it "utterly lacked" (1,3;1,4)- $\beta$ -D-glucan, but contained significantly higher levels of arabinoxylan as



compared with the Ni recurrent parent (Tonooka et al. 2008). The increased levels of arabinoxylan were likely there to maintain cell wall stability in lieu of (1,3;1,4)- $\beta$ -D-glucan. Scanning electron microscopy revealed the Ni(*bgl*) lines had markedly thinner endosperm cell walls with no difference in aleurone cell walls when compared with the Ni recurrent parent. The Ni(*bgl*) grain was much softer and had a much shorter pearling time, but also resulted in a higher percentage of broken grains when compared to Ni.

Linkage mapping placed the *bgl* gene in the centromeric region of chromosome 7H. Phenotypes in *bgl* cosegregated with polymorphisms in *HvCslF6*, leading researchers to conclude the *bgl* mutation is caused by a mutation and loss of function of the *HvCslF6* gene. In a follow-up study three of the Ni(*bgl*) NILs were sequenced. Each of the *bgl* mutants had a unique single-nucleotide substitution in the coding region of the *HvCslF6* gene resulting in a change of a highly conserved amino acid residue (Taketa et al. 2012). Microsomal membranes isolated from the developing endosperm of the *bgl* mutants lacked detectable (1,3;1,4)- $\beta$ -D-glucan indicating the *HvCslF6* protein was inactive. Further *Nicotiana benthamiana* plants were transformed with the wild-type and mutant *HvCslF6* genes. Wild-type *HvCslF6* proteins were able to direct the synthesis of high levels of (1,3;1,4)- $\beta$ -D-glucan while the mutants proteins were incapable of synthesizing (1,3;1,4)- $\beta$ -D-glucan (Taketa et al. 2012). These results and other suggest that *HvCslF6* has a unique role and are essential for (1,3;1,4)- $\beta$ -D-glucan synthesis in barley.

## **Conclusion**

In recent years much progress had been made in identifying genes and families of genes involved in (1,3;1,4)- $\beta$ -D-glucan synthesis. But to date a (1,3;1,4)- $\beta$ -D-glucan synthase complex has not be purified nor is it known which enzymes are members of this complex. Evidence suggests that *CslF6* is potentially a major player in (1,3;1,4)- $\beta$ -D-glucan synthesis. Although

CslF6 appears to be essential for (1,3;1,4)- $\beta$ -D-glucan synthesis other genes, such as CslH and CslJ are probably involved in altering the fine structure and deposition of (1,3;1,4)- $\beta$ -D-glucan. Little research has been done concerning (1,3;1,4)- $\beta$ -D-glucan synthesis in oat due to the greater complexity hexaploid genetics. We therefore have used barley and to a lesser extent rice as cereal models for investigating the genetics of (1,3;1,4)- $\beta$ -D-glucan synthesis in oat.

## References

- Anderson JW, Story L, Sieling B, Chen WL, Petro MS, Story J (1984) Hypocholesterolemic effects of oat-bran or bean intake for hypercholesterolemic men. *Am J Clin Nutr* 40:1146-1155
- Braaten JT, Wood PJ, Scoot FW, Wolynetz MS, Lowe MK, Bradley-White P, Collins MW (1994) Oat B-glucan reduces blood cholesterol concentration in hypercholesterolemic subjects. *Eur J Clin Nutr* 48:465-474
- 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 Biotech J* 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. *Mol 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 Physiol* 146:1821-1833
- Burton RA, Wilson SM, Hrmova M, Harvey AJ, Shirley AM, Stone BA, Newbigin EJ, Bacic A, Fincher GB (2006) Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)- $\beta$ -D-glucans. *Science* 311:1940-1942
- Carpita NC (1996) Structure and biogenesis of the cell walls of grasses. *Annu Rev Plant Physiol Plant Mol Biol* 47:445-447
- Carpita NC, McCann MC (2010) The maize mixed-linkage (1,3;1,4)- $\beta$ -D-glucan polysaccharide is synthesized at the Golgi membrane. *Plant Physiol* 153:1362-1371
- Collins HM, Burton RA, Topping DL, Liao ML, Bacic A, Fincher GB (2010) Variability in fine structure of noncellulosic cell wall polysaccharides from cereal grains: potential importance in human health and nutrition. *Cereal Chem* 87:272-282
- 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. *Proc Nat Acad Sci USA* 106:5996-6001
- Farrokhi N, Burton RA, 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 Biotech J* 4:145-167
- FDA (1997) 21 CFR Part 101. Food labeling, health claims: soluble dietary fiber from certain foods and coronary heart disease. *Federal Register* 62 15:3584-3601

- Fincher GB (2009a) Exploring the evolution of (1,3;1,4)- $\beta$ -D-glucans in plant cell walls: comparative genomics can help! *Curr Opin Plant Biol* 12:140-147
- Fincher GB (2009b) Revolutionary times in our understanding of cell wall biosynthesis and remodeling in the grasses. *Plant Physiol* 149:27-37
- Fincher GB, Stone BA (2004) Chemistry of nonstarch polysaccharides. In: Wrigley C, Corke H, Walker CE (eds) *Encyclopedia of Grain Science*, Elsevier Academic Press, Oxford, pp 206-223
- Hazen SP, Scott-Craig JS, Walton JD (2002) Cellulose synthase-like genes of rice. *Plant Physiol* 128:336-340
- Hesselman K, Aman P (1986) The effect of  $\beta$ -glucanase on the utilization of starch and nitrogen by broiler chickens fed on barley of low- or high-viscosity. *Anim Feed Sci Technol* 15:83-93
- Kirby RW, Anderson JW, Sieling B, Rees ED, Chen WL, Miller RE, Kay RM (1981) Oat-bran intake selectively lowers serum low-density lipoprotein cholesterol concentrations of hypercholesterolemic men. *Am J Clin Nutr* 34:824-829
- Marshall HG, Sorrells ME (1992) Oat science and technology. American Society of Agronomy Inc. & Crop Science Society of America Inc., Madison
- Nemeth C, Freeman J, Jones HD, Sparks C, Pellny TK, Wilkinson MD, Dunwell J, Andersson A, Aman P, Guillon F, Saulnier L, Mitchell R, Shewry PR (2010) Down-regulation of the CslF6 gene results in decreased (1,3;1,4)- $\beta$ -D-glucan in endosperm of wheat. *Plant Physiol* 152:1209-1218
- Reynolds SG (2004) Background to fodder oats worldwide. In Suttie JM, Reynolds SG (eds) *Fodder Oats: a world overview*. Plant Production and Protection Series, No. 33, FAO, Rome, pp 1-5
- Roulin S, Buchala AJ, Fincher GB (2002) Induction of (1,3;1,4)- $\beta$ -D-glucan hydrolases in leaves of dark-incubated barley seedlings. *Planta* 215:51-59
- Shimizu C, Kihara M, Aoe S, Araki S, Ito K, Hayashi K, Watari J, Sakata Y, Ikegami S (2008) Effect of high  $\beta$ -glucan barley on serum cholesterol concentrations and visceral fat area in Japanese men- a randomized, double-blinded, placebo controlled trial. *Plant Foods Hum Nutr* 63:21-25
- 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. *J Exp Bot* 63:381-392
- 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. *Breed. Sci* 59:47-54

Welch RW (1995) The oat crop, production and utilization. Chapman & Hall, London

Welch RW, Brown JCW, Leggett JM (2000) Interspecific and intraspecific variation in grain and groat characteristics of wild oat (*Avena*) species: very high groat (1-->3), (1-->4)- $\beta$ -D-glucan in an *Avena atlantica* genotype. *J Cereal Sci* 31:273-279

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

Wood PJ (2007) Cereal  $\beta$ -glucans in diet and health. *J Cereal Sci* 46:230-238