UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM FITOTECNIA

# GENOME-WIDE ASSOCIATION FOR β-GLUCAN CONTENT AND CHARACTERIZATION OF THE *Cellulose synthase-like F6* GENE IN OATS

Cristiano Mathias Zimmer
Mestre em Fitotecnia/UFRGS

Tese apresentada como um dos requisitos
à obtenção do Grau de Doutor em Fitotecnia
Área de Concentração em Recursos Genéticos, Biotecnologia e Melhoramento Vegetal

Porto Alegre (RS), Brasil
Fevereiro de 2020

CIP - Catalogação na Publicação

Zimmer, Cristiano Mathias
GENOME-WIDE ASSOCIATION FOR β-GLUCAN CONTENT AND CHARACTERIZATION OF THE Cellulose synthase-like F6 GENE IN OATS / Cristiano Mathias Zimmer. -- 2020.
143 f.
Orientador: Luiz Carlos Federizzi.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Agronomia, Programa de Pós-Graduação em Fitotecnia, Porto Alegre, BR-RS, 2020.

1. Avena sativa L.. 2. Genome-wide association mapping. 3. β-glucan content. 4. Kernel shape. 5. CslF6. I. Federizzi, Luiz Carlos, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

CRISTIANO MATHIAS ZIMMER
Engenheiro Agrônomo – UFPel
Mestre em Fitotecnia - UFRGS

# TESE

Submetida como parte dos requisitos

para obtenção do Grau de

## DOUTOR EM FITOTECNIA

Programa de Pós-Graduação em Fitotecnia

Faculdade de Agronomia

Universidade Federal do Rio Grande do Sul

Porto Alegre (RS), Brasil

Aprovado em: 11.02.2020
Pela Banca Examinadora

Homologado em: 06.05.2020
Por

LUIZ CARLOS FEDERIZZI
Orientador - PPG Fitotecnia
UFRGS

CHRISTIAN BREDEMEIER
Coordenador do Programa de
Pós-Graduação em Fitotecnia

CARLA ANDRÉA DELATORRE
PPG Fitotecnia/UFRGS

MARCELO TEIXEIRA PACHECO
PPG Fitotecnia/UFRGS

ANA BEATRIZ LOCATELLI
Corteva Agriscience

ARTHUR GERMANO FETT NETO
Centro de Biotecnologia/UFRGS

CARLOS ALBERTO BISSANI
Diretor da Faculdade de
Agronomia

**DEDICATION**

To my family, for their love and for teaching me what I would not find in any book. To my wife, Fernanda Leal Zimmer, who has been by my side and supported me during my education.

## ACKNOWLEDGMENTS

To the Federal University of Rio Grande do Sul (UFRGS) and to the Graduate Program in Plant Science, for the opportunity to undertake the doctorate degree.

To my advisor Luiz Carlos Federizzi, for his example, advice, friendship, guidance and unconditional support during my education. I have no words to thank you for all you have done for me.

To professor Marcelo Teixeira Pacheco, who idealized the genome-wide association study, for his example, advice, friendship and unconditional support during my education. I will not forget your dedication and willingness to teach/help.

To all UFRGS professors, especially to Carla Andrea Delatorre and Arthur Germano Fett-Neto that gave me valuable ideas during my qualifying exam.

To the University of Minnesota, especially to my advisor Kevin P. Smith, for the opportunity to study at one of the best plant breeding schools. To the Smith Lab, especially to Ian G. McNish, Dimitri von Ruckert, Joan Manuel Barreto Ortiz, John Hill Price, and Jeffrey Neyhart.

To the PepsiCo Agro & Discovery team for supporting and funding this project. My special thanks to Gabe Gusmini, David P. Eickholt, Mandy Waters, Nickolas Anderson, Jason Cepela, Cesar Velasquez, Laurel Weinkauf, and Lisa Wiggins.

To the Agronomic Institute of Paraná, especially to Klever Márcio Antunes Arruda and Carlos Roberto Riede, for their support and collaboration with field activities in Paraná.

To the University of Passo Fundo, especially to professor Luiz Carlos Gutkoski, Tatiana Oro, and Ricardo Taglietti, for their crucial collaboration with β-glucan content analyses.

To all researchers that collaborated with this project, especially to Kathy Esvelt Klos (USDA-ARS), Stephen Harrison (LSU), and Fernando Fumagalli Miranda (IRGA).

To the National Council for Scientific and Technological Development (CNPq) and Coordination for the Improvement of Higher Education Personnel (CAPES) for providing me a scholarship in Brazil and abroad, respectively.

To the DEPLAV technicians, Fábio Berndt and Gustavo Maia Leão, for their promptness and collaboration with this project. My special thanks to the UFRGS Agronomic Experimental Station assistants Miguel and Paulo, for their help with field experiments.

To my colleagues for their friendship and help during these years. My special thanks to the undergrad students Guilherme Oliveira, Filipe Coelho, and Felipe Krause, for their valuable help during my experiments.

To my parents, Alvete M. Zimmer and Ildemar G. Zimmer (*in memorian*), for all principles and values.

To my family, Alvete M. Zimmer, Aline Zimmer, Greici Zimmer, Arthur Zimmer A. Moreira, João Zimmer Schneid, and Cynthia Leal, for their support, encouragement, and understanding when I was absent.

To my wife, Fernanda Leal Zimmer, for her love and unconditional support.

To everyone who somehow contributed to my education.

# GENOME-WIDE ASSOCIATION FOR β-GLUCAN CONTENT AND CHARACTERIZATION OF THE *Cellulose synthase-like F6* GENE IN OATS[1]

Author: Cristiano Mathias Zimmer
Advisor: Luiz Carlos Federizzi

## ABSTRACT

The β-glucan content has been one of the main goals of oat breeding programs worldwide. However, the loci controlling this trait and their interaction with the environmental conditions are not completely understood to date, especially in subtropical environments. In this sense, the main objectives of this study were: i) to identify genomic regions associated with β-glucan content in elite oat germplasm adapted to subtropical environments; ii) to identify genomic regions controlling kernel shape-related traits and verify its association with β-glucan content on oats; iii) to characterize and quantify the absolute expression of the *Cellulose synthase-like F6* gene in oat genotypes with contrasting β-glucan content; and iv) to characterize the β-glucan formation in primary and secondary oat kernels harvested in different planting dates and years. An oat panel composed by 413 genotypes was genotyped and assessed in different environments. Genome-wide association mappings were performed using individual and multi-environment approaches. Quantitative trait loci (QTL) associated with β-glucan content were identified on Mrg02, Mrg06, Mrg11, Mrg12, Mrg19, and Mrg20 consensus groups. QTL associated with kernel shape-related traits were detected on Mrg06, Mrg13, Mrg21, and Mrg24 consensus groups. Kernel width and kernel thickness traits, which were phenotypically and genetically correlated, share a genomic region in common on Mrg13 and are negatively correlated with β-glucan content. The oat β-glucan content varied between primary and secondary kernels according to the genotype and environmental conditions. Under unfavorable conditions, some genotypes can present higher β-glucan content in secondary kernels than in primary ones. The high number of *CslF6* transcripts was not a guarantee of high β-glucan content in oats. The *CslF6* gene expression was highly influenced by the genotype, environmental conditions, and kernel developmental stage. The third exon of the *CslF6* gene, which corresponds to the CslF6 motif, has 1767 bases pair in oats, regardless of the subgenome. The CslF6 motif of the *CslF6-AA* and *CslF6-DD* genes encodes the same amino acid sequence. The *CslF6-CC* gene showed 13 amino acid changes in the CslF6 motif when compared to the other ones. Single nucleotide polymorphisms were not identified in the *CslF6* homoeologs among the assessed genotypes, indicating that the CslF6 proteins with the same functional properties are encoded in these genotypes. The results presented in this study contribute to improve the understanding of the β-glucan content in oats under subtropical conditions.

[1]Doctoral Thesis in Plant Science, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil. (143p.) February, 2020.

# ASSOCIAÇÃO GENÔMICA AMPLA PARA O TEOR DE β-GLICANAS E CARACTERIZAÇÃO DO GENE *Cellulose synthase-like F6* EM AVEIA[1]

Autor: Cristiano Mathias Zimmer
Orientador: Luiz Carlos Federizzi

## RESUMO

O teor de β-glicanas tem sido um dos principais objetivos dos programas de melhoramento de aveia mundialmente. Contudo, os locos controlando este caráter e a sua interação com as condições ambientais não são completamente compreendidas até o momento, especialmente em ambientes subtropicais. Neste sentido, os principais objetivos deste estudo foram: i) identificar regiões genômicas associadas ao teor de β-glicanas em germoplasma elite de aveia adaptado a ambientes subtropicais; ii) identificar regiões genômicas controlando caracteres relacionados à morfologia de grãos e verificar a sua associação com o teor de β-glicanas em aveia; iii) caracterizar e quantificar de forma absoluta a expressão do gene *Cellulose synthase-like F6* em genótipos de aveia contrastantes para o teor de β-glicanas; e iv) caracterizar a formação de β-glicanas em grãos primários e secundários de aveia colhidos em diferentes épocas de semeadura e anos de cultivo. Um painel de aveia composto por 413 genótipos foi genotipado e avaliado em diferentes ambientes. Mapeamentos de associação genômica ampla foram realizados utilizando análises individuais e conjuntas. *Quantitative trait loci* (QTL) associados ao teor de β-glicanas foram identificados nos grupos consenso Mrg02, Mrg06, Mrg11, Mrg12, Mrg19 e Mrg20. QTL associados à caracteres relacionados a morfologia de grãos foram detectados nos grupos consenso Mrg06, Mrg13, Mrg21 e Mrg24. Os caracteres largura de grãos e espessura de grãos, que foram fenotipicamente e geneticamente correlacionados, compartilham uma região genômica em comum no Mrg13 e estão negativamente correlacionados com o teor de β-glicanas. O teor de β-glicanas variou entre grãos primários e secundários de acordo com o genótipo e as condições ambientais. Em condições desfavoráveis, alguns genótipos podem apresentar maior teor de β-glicanas em grãos secundários em relação aos grãos primários. O elevado número de transcritos do gene *CslF6* não foi uma garantia de alto teor de β-glicanas em aveia. A expressão do gene *CslF6* foi altamente influenciada pelo genótipo, condições ambientais e estádio de desenvolvimento do grão. O terceiro éxon do gene *CslF6*, que corresponde ao *motif* CslF6, tem 1767 pares de base em aveia, independente do subgenoma. O *motif* CslF6 dos genes *CslF6-AA* e *CslF6-DD* codificam a mesma sequência de aminoácidos. O gene *CslF6-CC* apresentou 13 alterações de aminoácidos no *motif* CslF6 quando comparado aos demais. *Single nucleotide polymorphisms* não foram identificados nos homeólogos do gene *CslF6* entre os genótipos avaliados, indicando que proteínas CslF6 com as mesmas propriedades funcionais são codificadas nestes genótipos. Os resultados apresentados neste estudo contribuem para melhorar o entendimento do teor de β-glicanas em aveia em condições subtropicais.

[1]Tese de Doutorado em Fitotecnia, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil. (143f.) Fevereiro, 2020.

## LIST OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# 1 INTRODUCTION

Oats (*Avena sativa* L.) became one of the most recommended foods for human consumption. One of the main factors contributing for the great nutritional quality of oat kernels is the higher β-glucan content when compared to other crops, such as wheat, rice, and corn. The ingestion of β-glucans in proper quantities provides many benefits for human health, including the improvement of the immune function, reduction of cholesterol levels, reduction of blood pressure, and reduction of glycemic index in diabetics.

Although associated with many benefits for human health, the allocation of resources for the Brazilian oat breeding is still scarce when compared to other small grains. The first research projects involving oats in Brazil were performed in the 1930s, when varieties from Uruguay and Argentina were evaluated in Southern Brazil. In the 1960s, oat fields used to be cropped using foreign varieties, mainly from United States. The first Brazilian oat varieties were released in the 1970s.

The first selection criteria aimed exclusively adaptation to subtropical conditions, considering mainly flowering time adequacy and pathogen resistance. After many selection cycles, the Brazilian oat breeders developed genotypes associating earliness, high biomass accumulation, pathogen resistance, and great yield potential. On the other hand, the selection focusing on chemical quality traits, including the β-glucan content, was not effectively performed to date. The selection of genotypes with high grain quality requires many chemical analyses, which are usually expensive, time-consuming, and demand expert analysts.

The quantitative inheritance of the β-glucan content and the high environmental influence exercised on this trait turn the development of superior genotypes even more difficult. Currently, the selection for oat grain quality in Brazil is performed indirectly, in advanced self-fertilization generations, considering visual assessment of the grain shape. In this sense, even without any direct parameter for grain quality, such as chemical analyses, for example, the Brazilian oat germplasm shows suitable levels of β-glucans when compared to the average of genotypes from other countries. In this context, the grain shape may be directly associated with the β-glucan content trait in oats.

The genetic and molecular mechanisms involved with the β-glucan formation in oats are poorly understood. The identification of these mechanisms, as well as the understanding of the environmental influence on them, will assist the development of new oat varieties with high β-glucan content. Many molecular tools are available to assist plant breeders nowadays, including genome-wide association and marker-assisted selection approaches. Using a collection of genotypes contrasting for the β-glucan content, previously assessed phenotypically and genotyped, it is possible to identify the genomic regions associated with the trait and develop new tools for oat breeders. In this way, the main objectives of this study were:

i) to identify genomic regions associated with β-glucan content in elite oat germplasm adapted to subtropical environments;

ii) to identify genomic regions controlling kernel shape-related traits and verify its association with β-glucan content in oats;

iii) to characterize the β-glucan formation in primary and secondary oat kernels harvested in different planting dates and years;

iv) to characterize and quantify the absolute expression of the *Cellulose synthase-like F6* gene in oat genotypes with contrasting β-glucan content.

# 2 LITERATURE REVIEW

## 2.1 Oat: from weed species to functional food

Oat belongs to the *Poaceae* botanical family, *Poideae* subfamily, *Aveneae* tribe, and *Avena* genus. The different oat species form a polyploid series, represented by diploid, tetraploid, and hexaploid species, with the basic chromosome number equal to seven. The cultivated oat (*Avena sativa* L.) is an allohexaploid species constituted by six sets of basic chromosomes (2n = 6x = 42), formed from the aggregation of three subgenomes (AA, CC, and DD) from ancestral diploid species (Rajhathy & Thomas, 1974).

Oat evolution occurred from interspecific crossings and chromosomal duplication events. The crossing between the diploid species *Avena ventricosa* (CC) and *Avena canariensis* (AA) originated the *Avena magna* (AACC) species, which is the putative tetraploid ancestral of hexaploid oats (Rajhathy & Thomas, 1974). After, the crossing between *A. magna* and *Avena insularis* (CCDD), which is the putative subgenome DD donor, originated the hexaploid species *Avena sterilis* (AACCDD) (Loskutov, 2008). *A. sterilis* is likely the ancestral of hexaploid oats, including *A. sativa* L. (Coffman, 1977).

The oat origin center is associated with the Minor Asia region (Loskutov, 2008), being that the domestication of this species was related to the Center-North of Europe. Later, oat seeds were dispersed to temperate areas as weed species in other small grains, such as wheat (*Triticum* ssp.) and barley (*Hordeum* ssp.) (Coffman, 1961). As consequence of their great adaptability, oats changed from wild to cultivated stage (Leggett and Thomas, 1995) and was widely disseminate due to their great agricultural aptitude, especially in temperate

and wet environments, such as North America and Europe (Murphy and Hoffman, 1992). The oat introduction in Latin America is associated with the Spanish immigrants' arrival. The first studies involving oats in Brazil occurred in the 1930s and the breeding of this species began in 1974 (Federizzi *et al.*, 2005).

Oat grains have great nutritional aptitude for human consumption and, when ingested in proper quantities, are considered a functional food with many benefits for human health. The functional food denomination aimed to associate oat consumption to human health and was stablished by the American Food and Drug Administration Agency on June 29$^{th}$, 1995 (Beber *et al.*, 2002). The main factors contributing to the high nutritional quality of oat kernels are protein content, with special and complete amino acid composition, and β-glucan content. The β-glucan [(1→3);(1→4)-β-D-glucan] ingestion is associated with different benefits for human health, including immune function improvement, reduction of cholesterol levels in the blood, and reduction in glycemic index in diabetics (Cho *et al.*, 2013; Whitehead *et al*., 2014; Ye *et al*., 2012).

### 2.2 Characterization and formation of β-glucans

β-glucans found in oat kernels are part of a mixed-linkage β-glucan family. Among many β-glucans found in plants, with β(1→2), β(1→3), β(1→4) or β(1→6) linkages, only grasses and lichens have β(1→3);(1→4) mixed-linkage β-glucans. Oat β-glucans are linear non-amylaceous polysaccharides composed by β-D- glucopyranosyl units connected with β(1→3) and β(1→4) mixed-linkages. β(1→3) linkages represent 30%, while β(1→4) ones represent 70% of the total linkages (Butt *et al*., 2008). Described as hemicelluloses, β-glucans are found in a wide range of organisms, from fungi to higher plants (Pettolino *et al*., 2009). The hemicelluloses, which are long linear chains of polysaccharides, usually with short lateral branches, are characterized for binding to the cellulose surface. In this sense, hemicelluloses can either connect cellulose microfibrils to form a cohesive network or act as a coating that prevents the microfibril surfaces from sticking together.

The formation of β-glucans is widely verified in monocots, including evolutionarily basal families such as *Flagellariaceae*. In dicot species, on the other hand, the formation of β-glucans was not detected (Smith and Harris, 1999). Considering the 16 families that belongs to the *Poales* order, which usually present the formation of β-glucans, the *Poaceae* one is the most important (Trethewey *et al.*, 2005). The formation of β-glucans is associated with many functions in plants, including the definition of physical properties of the cell wall, basically interacting with other polymers (Scheller and Ulvskov, 2010), and the carbohydrate storage in seeds (Meier and Reid, 1982; Reid, 1985).

There are many factors acting on the formation and deposition of β-glucans in plant cells. Among different enzymes involved on these processes, glycosyltransferases are classed as a key component of the β-glucan biosynthesis pathway. Glycosyltransferases are associated with the Golgi complex membrane, where β-glucans are synthetized. After synthetized in the Golgi complex, β-glucans are transported and incorporate in the cell wall. This transport occurs through subcellular vesicles, more precisely secretory vesicles or multivesicular bodies (Figure 1).

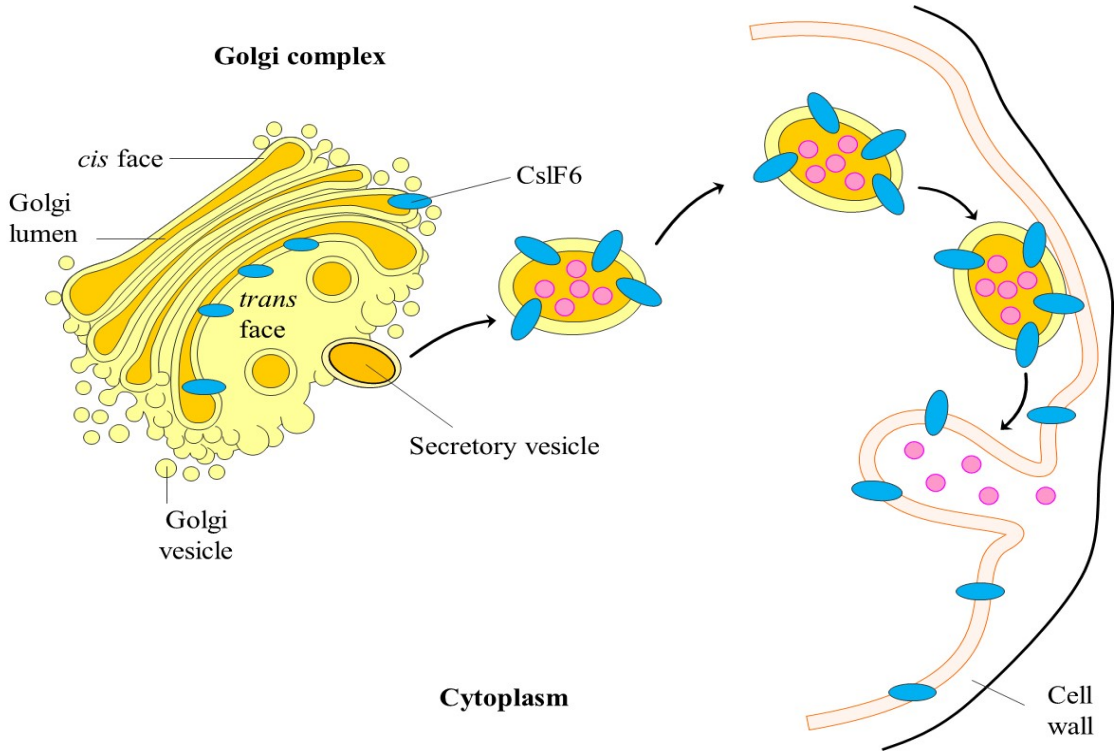



FIGURE 1. Biosynthesis of β-glucans in grasses. Adapted from Wilson *et al.* (2015).

Once carried into secretory vesicles, the transport of these vesicles to cell wall will be performed using actin microfibrils and tubulin microtubes. After transported and close to the plasmatic membrane, cell turgor pressure acts “flattening” the vesicles and forcing the transference of molecules carried inside them, including β-glucans, initially to the plasmatic membrane and later to the cell wall.

### 2.3 β-glucans in oats and other grasses

The formation of β-glucans is abundant in primary and secondary cell walls of species that belong to the *Poaceae* family. Among many species with agricultural importance, the formation of β-glucans is more pronounced in oats, barley (*Hordeum vulgare* L.) and rye (*Secale cereale*). The β-glucan content in kernels of these grasses range from 3 to 7%, from 3 to 11%, and from 1 to 2%, respectively (Skendi *et al.*, 2003). Conversely, the formation of β-glucans is less expressive in wheat (*Triticum aestivum* L.), rice (*Oryza sativa* L.), and corn (*Zea mays*) (Fincher and Stone, 2004). Studying a collection of 230 tetraploid wheat genotypes, including different subspecies, Marcotuli *et al.* (2016) verified a phenotypic variation for the β-glucan content trait ranging from 0.18 to 0.89% in kernels. The β-glucan content in rice and corn kernels is negligible (Wood and Beer, 2000).

Great effort has been performed to restrict the β-glucan formation in barley kernels, once β-glucans have a negative influence on brewing procedures (Vis and Lorenz, 1997). On the other hand, aiming the production of healthy products for the industry, the selection of genotypes with high β-glucan content became one of the main objectives of oat breeding programs worldwide (Peterson, 1991). The oat variety ‘HiFi’, which was developed by the North Dakota State University Oat Breeding Program, has great β-glucan content, ranging from 6 to 7%, varying among crop seasons (McMullen *et al.*, 2005). These values represent approximately 1.5 – 2% more β-glucans than the oat average. The

oat variety ‘IAC 7’, which was developed by the Agronomic Institute of Campinas, is always reported as a check for β-glucan content in Brazil. However, the phenotyping of this variety for β-glucan content is usually performed using near infrared spectroscopy (NIRS).

Oat kernels are structurally composed by several layers. Starting from the external layer, oat kernels are coated by epidermis, hypodermis, cross cells, tube cells, seed coat (testa), nucellar tissue, aleurone layer, and endosperm (Figure 2). β-glucans are preferentially concentrated in the aleurone layer, sub aleurone layer, and starchy endosperm adjacent to the embryo. The richest region for β-glucans in oat kernels correspond to the junction between aleurone layer and endosperm (Wood *et al*., 1991; Wood, 1993).

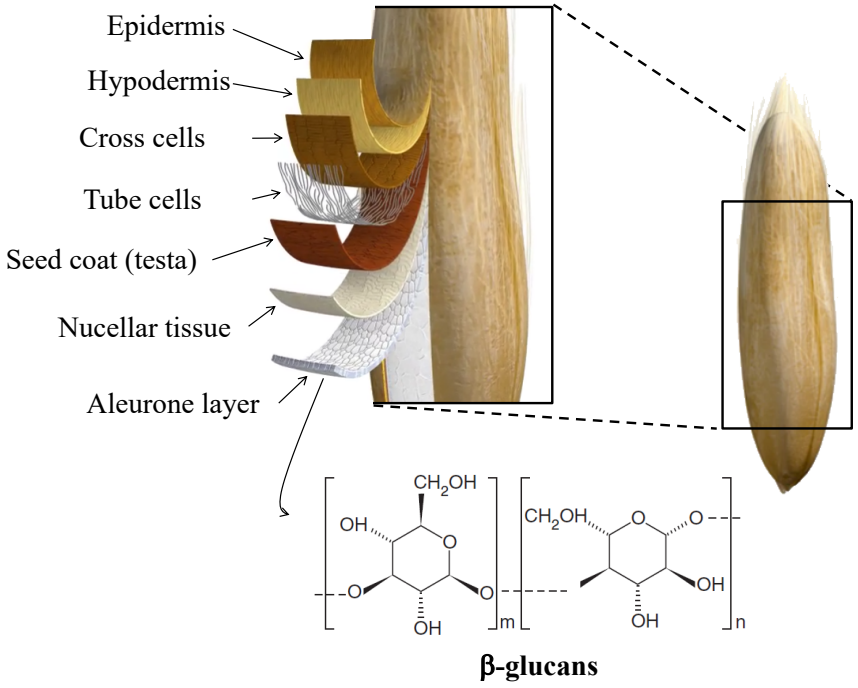



FIGURE 2. Parts of an oat kernel and preferred β-glucan deposition site. Adapted from Kampffmeyer (2016).

There are major limitations for the elucidation of the β-glucan content in oats and selection of genotypes with great performance for this trait. Among the main limitations are: i) absence of advanced studies indicating which genomic regions control the synthesis and deposition of β-glucans in oats; ii) absence of molecular markers associated with the

trait and useful for marker assisted selection (MAS), which would allow the selection of oat genotypes in a fast and precise way; iii) scarcity of fast, low-cost, and reliable methods for β-glucan content quantification in oats; and iv) influence of environmental conditions on the β-glucan content trait, biasing the breeder's selection.

### 2.4 Genotype by environment influence on β-glucan content

The cell wall composition and polysaccharides involved on its formation are not static. During plant development, cell wall components can be modified by the synthesis and action of enzymes that act trimming lateral branches and digesting pectin and hemicelluloses (Gibeaut *et al*., 2005). In this sense, the genetic mechanisms that control the synthesis of cell wall components, including β-glucans, must interact and are influenced by the environmental conditions of each crop season.

Many studies have been published indicating the environmental influence on β-glucan content in oats (Saastamoinen *et al*., 1992; Doehlert *et al*., 2001; Beber *et al*., 2002; Saastamoinen *et al*., 2004). Studying the β-glucan content in 15 Brazilian oat genotypes, evaluated in Southern Brazil, Crestani *et al.* (2010) verified a significant interaction among genotype, environment, and year influencing the trait. Similarly, studying four Canadian oat varieties, Humphreys *et al.* (1994) identified that different planting dates also impact the β-glucan content in oats. In addition, the occurrence of successive rainfalls during grain filling tend to reduce the β-glucan content in oat kernels (Miller *et al*., 1993). On the other hand, rainfalls during the vegetative stage are positively correlated with the β-glucan content in oats (Herrera *et al*., 2016). High temperature conditions during the formation of kernels also tend to increase the β-glucan content in oats (Saastamoinen *et al*., 1992; Miller *et al*., 1993; Saastamoinen *et al*., 1995). In barley, the occurrence of drought stress at the maturity stage also influenced the β-glucan content (Fastnaught *et al*., 1996; Hang *et al*., 2007).

Using multiple regression analysis with different oat genotypes and environments, Andersson and Börjesdotter (2011) estimated that 23% of the observed phenotypic variation for the β-glucan content is due to the genotype, 42% is due to environmental conditions, and 11% is due to genotype by environmental condition interaction. Similarly, Brunner and Freed (1994) verified that the β-glucan content trait is highly influenced by crop year and less influenced by genotype. On the other hand, Saastamoinen *et al.* (2004) emphasized that the β-glucan content is determined mainly by the genotype. Recently, Herrera *et al.* (2016) verified that year is the major factor contributing to the phenotypic variation of β-glucan content in oats, followed by genotype. In this context, the main challenges for oat breeders and researchers are to identify and localize the genetic mechanisms involved on the synthesis and deposition of β-glucans, as well as to clarify how these mechanisms are regulated according to the specific conditions of the environment.

### 2.5 Genetic basis of the β-glucan content in grasses

The information available in the literature about the genetic mechanisms involved on the synthesis and accumulation of β-glucans in oats is limited. The β-glucan content is a quantitative trait, and the loci controlling this trait show basically additive effects (Holthaus *et al*., 1996; Crestani *et al*., 2012). There are many enzymes involved on the synthesis and deposition of polysaccharides in the cell wall. Among the main enzymes involved on the β-glucan biosynthesis pathway are the glycosyltransferases, which act modifying carbohydrates (Coutinho and Henrissat, 1999). The glycosyltransferases are divided in two main groups: i) polysaccharide synthases (type I); and ii) glycosyltransferases (type II).

The type I enzymes are coded by the Cellulose Synthase gene family, which is divided in subfamilies. Among these subfamilies are the Cellulose synthase (CesA) and

Cellulose synthase-like (Csl) ones (Richmond and Somerville, 2000). The Csl subfamily genes are classed from 'A' to 'H'. In Arabidopsis and rice 30 and 37 members of this subfamily were identified, respectively (Hazen *et al*., 2002; Somerville *et al*., 2004). Studies involving comparative mapping showed that a cluster of CslF genes from rice is co-located in a genomic region corresponding to a major quantitative trait loci (QTL) for β-glucan content in barley (Burton *et al*., 2006). The insertion and expression of CslF genes from rice allowed the detection of β-glucans in the cell walls of Arabidopsis (Burton *et al*., 2006), indicating that these genes play a major role in the β-glucan biosynthesis pathway. In wheat, the silencing of the *Cellulose synthase-like F6* (*CslF6*) gene using an interference RNA reduced on average 42.4% of the β-glucan content in kernels (Nemeth *et al*., 2010). On the other hand, when the *CslF6* gene was overexpressed in barley, under the control of an oat endosperm specific promoter, an increasing greater than 80% in the β-glucan content was observed in kernels (Burton *et al*., 2011). In addition, the *betaglucanless* (*bgl*) mutants in barley have a point mutation in the *CslF6* gene which alter a highly conserved amino acid in the coded protein (Taketa *et al.,* 2012).

Studying the expression of seven CslF gene members in the barley genotypes 'Himalaya' and 'Sloop', which are contrasting for the β-glucan content, Burton *et al.* (2008) verified gene expression profiles completely different for the *Cellulose synthase-like F9* (*CslF9*) and *CslF6* genes. These genes were the only CslF ones expressed during the grain formation and filling in barley. The *CslF9* gene showed an expression peak at eight days after anthesis (DAA) in the 'Sloop' genotype, which has reduced β-glucan content. On the other hand, the *CslF6* gene showed an increasing in the transcript levels between 12 and 20 DAA, reaching an expression peak at 20 DAA in the 'Himalaya' genotype, which has high β-glucan content (Burton *et al*., 2008). These results suggest a distinct and opposite temporal expression of the *CslF9* and *CslF6* genes in barley, wherein only the *CslF6* one acts positively on the β-glucan content in this species.

The results described previously indicate that CslF genes are strongly associated with the β-glucan content in grasses. These genes encode polysaccharide synthases that, along with a complex regulatory pathway, act on the synthesis and accumulation of β-glucans. Among the CslF subfamily members, the *CslF6* gene plays a major role on the β-glucan biosynthesis in grasses. The involvement of this gene on the β-glucan synthesis may occur in different ways, either on the expression level (number of transcripts) or on specific mutations in the coding region which could result in enzymes with different functional properties. The *CslF6* gene was identified and cloned in oats (Chawade *et al*., 2010; Jobling, 2015). However, studies associating the occurrence of genetic polymorphisms on this gene to the β-glucan content, as well as the characterization of his expression profile, are not available in the literature to date. The only study involving the *CslF6* gene expression in oats was performed under controlled conditions and with developing kernels until nine days after pollination (Coon, 2012), not comprehending the crucial stage for the β-glucan formation verified in other grasses.

### 2.6 Genomic regions associated with β-glucan content in grasses

The first studies involving the identification of genomic regions associated with β-glucan content in grasses were performed in barley. A major QTL on the barley β-glucan content was detected on the 1H chromosome, contributing with approximately 15% of the phenotypic variation observed for the trait (Han *et al*., 1995). Recently, loci acting on the increasing of the β-glucan content were identified on the barley chromosomes 3H, 4H, 5H, 6H, and 7H, being that the major ones were located on the chromosomes 4H, 5H, and 7H (Islamovic *et al*., 2013). The *CslF6* gene was mapped on the barley chromosome 7H (Burton *et al*., 2008), in a genomic region that corresponds to a major QTL for β-glucan content on the ‘Derkado’ x ‘B83-12/21/5’ and ‘Beka’ x ‘Logan’ biparental populations (Igartua *et al*., 2002; Molina-Cano *et al*., 2007).

Assessing a collection composed by 230 tetraploid wheat genotypes, Marcotuli *et al.* (2016) identified 12 markers positively associated with the β-glucan content. These markers were distributed in seven QTL located on the wheat chromosomes 1A, 2A, 2B, 5B, and 7A. In most of these regions, genes encoding glycosyltransferases or glycosyl hydrolases were identified. In addition, nucleotide sequences showing high homology to genes encoding the enzymes starch synthase II, isoamylase, and hydroxylases were detected on these QTL. Most of these sequences were syntenic between wheat and barley (Marcotuli *et al*., 2016).

Molecular mapping approaches have been performed to identify genomic regions associated with the β-glucan content in oats (Kianian *et al*., 2000; De Koeyer *et al*., 2004; Herrmann *et al*., 2014). These studies were performed using biparental populations, which correspond to a set of individuals obtained by the single crossing between two contrasting parents. However, the loci identified on these studies are rarely used for breeding purposes due to many reasons, such as: i) the association between marker and QTL may be significant only for the studied biparental population; and ii) the identified QTL may show minor effects on the phenotype, being highly influenced by environmental conditions and not explaining in a satisfactory way the observed phenotypic variation.

The genome-wide association emerges as an alternative to overcome the limitations presented previously. The genome-wide association, differently of the molecular mapping using biparental populations, is performed using a collection of unrelated genotypes and with a wide genetic variability. As consequence, the genome-wide association analysis has great power to detect major QTL acting on the phenotype, not being restricted to the alleles from two parents, as verified in biparental populations.

Genome-wide association studies for the β-glucan content have been performed in oats (Newell *et al*., 2012, Asoro *et al.*, 2013). Using an elite oat panel composed by 431 genotypes, Newell *et al.* (2012) identified three markers associated with the β-glucan

content. One of these markers shows high nucleotide sequence homology to genes located on the rice chromosome seven, in an adjacent region to the CslF subfamily members (Newell *et al*., 2012). Studying 446 oat genotypes previously genotyped with 1005 diversity array technology (DArT) markers, Asoro *et al.* (2013) identified 37 main markers explaining the β-glucan content trait in oats. These markers were identified using individual tests for each marker and the least absolute shrinkage and selection operator (LASSO) regression method. However, both studies used North American oat germplasm, which is completely distinct and not adapted to subtropical environments. In this sense, the present study was proposed to identify the genomic regions that control the β-glucan content in subtropical conditions, as well as to elucidate the action of specific genes located on these regions.

## 2.7 References

ANDERSSON, A. A. M.; BÖRJESDOTTER, D. Effects of environment and variety on content and molecular weight of β-glucan in oats. **Journal of Cereal Science**, London, v. 54, p. 122-128, 2011.

ASORO, F. G. *et al.* Genome-wide association study for beta-glucan concentration in elite North American oat. **Crop Science**, Madson, v. 53, p. 542-553, 2013.

BEBER, R. C. *et al*. Caracterização química de genótipos brasileiros de aveia (*Avena sativa* L.). **Acta Científica Venezolana**, Caracas, v. 53, p. 202-209, 2002.

BRUNNER, B. R.; FREED, R. D. Oat grain β-glucan content as affected by nitrogen level, location and year. **Crop Science**, Madison, v. 34, p. 473-476, 1994.

BURTON, R. A. *et al*. *Cellulose synthaselike CslF* genes mediate the synthesis of cell wall (1,3;1,4)-b-D-glucans. **Science**, Washington, v. 311, p. 1940-1942, 2006.

BURTON, R. A. *et al.* The genetics and transcriptional profiles of the Cellulose Synthase-Like HvCslF gene family in barley. **Plant Physiology**, Rockville, v. 146, p. 1821-1833, 2008.

BURTON, R. A. *et al.* Overexpression of specific *HvCslF cellulose synthase-like* genes in transgenic barley increases the levels of cell wall (1,3;1,4)-β-D-glucans and alters their fine structure. **Plant Biotechnology Journal**, Oxford, v. 9, p. 117-135, 2011.

BUTT, M. S. *et al*. Oat: unique among the cereals. **European Journal of Nutrition**, Darmstadt, v. 47, p. 68-79, 2008.

CHAWADE, A. *et al*. Development and characterization of an oat TILLING-population and identification of mutations in lignin and beta-glucan biosynthesis genes. **BMC Plant Biology**, London, v. 10, p. 1-13, 2010.

CHO, S. S. *et al*. Consumption of cereal fiber, mixtures of whole grains and bran, and whole grains and risk reduction in type 2 diabetes, obesity, and cardiovascular disease. **American Journal of Clinical Nutrition**, New York, v. 98, p. 594-619, 2013.

COFFMAN, F. A. **Oats and oats improvement**. Madison: American Society of Agronomy, 1961. 650p.

COFFMAN, F. A. **Oat history, identification and classification**. Washington: USDA-ARS, 1977. 356p.

COON, M. A. Characterization and variable expression of the *CslF6* homologs in oat (*Avena* sp.). **Thesis (Master)** - Department of Plant and Wildlife Sciences, Brigham Young University (BYU), Utah, f. 47, 2012.

COUTINHO, P. M.; HENRISSAT, B. Carbohydrate-active enzymes: an integrated database approach. In: GILBERT, H. J.; DAVIES, G. J.; HENRISSAT, B.; SVENSSON, B. (Eds.). **Recent advances in carbohydrate bioengineering: Carbohydrate-active enzymes: an integrated database approach.** The Royal Society of Chemistry, Cambridge, 1999. p. 3-12.

CRESTANI, M. *et al*. β-glucan content in white oat cultivars grown in different environments. **Pesquisa Agropecuária Brasileira**, Brasília, v. 45, p. 261-268, 2010.

CRESTANI, M. *et al.* Combining ability for grain chemistry quality traits in a white oat diallelic cross. **Euphytica**, Dordrecht, v. 185, p. 139-156, 2012.

DE KOEYER, D. L. *et al*. A molecular linkage map with associated QTLs from a hulless x covered spring oat population. **Theoretical and Applied Genetics**, Berlin, v. 108, p. 1285-1298, 2004.

DOEHLERT, D. C.; MCMULLEN, M. S.; HAMMOND, J. J. Genotypic and environmental effects on grain yield and quality of oat grown in North Dakota. **Crop Science**, Madison, v. 41, p. 1066-1072, 2001.

GIBEAUT, D. M. *et al*. Changes in cell wall polysaccharides in developing barley (*Hordeum vulgare*) coleoptiles. **Planta**, Berlin, v. 221, p. 729-738, 2005.

FASTNAUGHT, C. E. *et al*. Genetic and environmental variation in beta-glucan content and quality parameters of barley for food. **Crop Science**, Madison, v. 36, p. 941-946, 1996.

FEDERIZZI, L. C. *et al*. Melhoramento da aveia. In: BORÉM, A. (Ed.). **Melhoramento de espécies cultivadas**, 2ed. Viçosa: UFV, 2005. p. 141-169.

FINCHER G. B.; STONE B. A. Chemistry of nonstarch polysaccharides. In: WRIGLEY, C.; CORKE, H.; WALKER, C. E. (Eds.). **Encyclopedia of Grain Science**, Elsevier. Oxford, 2004. p. 206-223.

HAN, F. *et al*. Mapping of beta-glucan content and beta-glucanase activity loci in barley grain and malt. **Theoretical and Applied Genetics**, Berlin, v. 91, p. 921-927, 1995.

HANG, A. Barley amylose and beta-glucan: their relationships to protein, agronomic traits, and environmental factors. **Crop Science**, v. 47, p. 1754-1760, 2007.

HAZEN, S. P.; SCOTT-CRAIG, J. S.; WALTON, J. D. *Cellulose synthase-like* genes of rice. **Plant Physiology**, Rockville, v. 128, p. 336-340, 2002.

HERRMANN, M. H. *et al*. Quantitative trait loci for quality and agronomic traits in two advanced backcross populations in oat (*Avena sativa* L.). **Plant Breeding**, Berlin, v. 133, p. 588-601, 2014.

HERRERA, M. P. *et al.* β-Glucan content, viscosity, and solubility of Canadian grown oat as influenced by cultivar and growing location. **Canadian Journal of Plant Science**, Ottawa, v. 96, p. 183-196, 2016.

HOLTHAUS, J. F. *et al.* Inheritance of β-glucan content of oat grain. **Crop Science**, Madison, v. 36, p. 567-572, 1996.

HUMPHREYS, D. G.; SMITH, D. L.; MATHER, D. E. Nitrogen fertilizer and seeding date induced charges in protein, oil and β-glucan contents of four oat cultivars. **Journal of Cereal Science**, London, v. 20, p. 283-290, 1994.

IGARTUA, E. *et al.* Genetic control of quantitative grain and malt quality traits in barley. **Journal of Crop Production**, New York, v. 5, p. 131-164, 2002.

ISLAMOVIC, E. *et al*. Genetic dissection of grain beta-glucan and amylose content in barley (*Hordeum vulgare* L.). **Molecular Breeding**, Dordrecht, v. 31, p. 15-25, 2013.

JOBLING, S. A. Membrane pore architecture of the *CslF6* protein controls (1-3,1-4)-β-glucan structure. **Science Advances**, Washington, v. 1, p. 1-9, 2015.

LEGGETT, J. M.; THOMAS, H. Oat evolution and cytogenetics. In: WELCH, R. W. (Ed.). **The oat crop**: production and utilization. London: Chapman & Hall, 1995. p. 120-149.

LOSKUTOV, I. G. On evolutionary pathways of *Avena* species. **Genetic Resources and Crop Evolution**, Amsterdam, v. 55, p. 211-220, 2008.

KAMPFFMEYER. **Oat kernel layer - Animation**. [2016] Available in: http://grain-gallery.com/en/oat/animations. Access in 14 July 2016.

KIANIAN, S. F. *et al*. Quantitative trait loci influencing β-glucan content in oat (Avena sativa, 2n=6x=42). **Theoretical and Applied Genetics**, Berlin, v. 101, p. 1039-1048, 2000.

MARCOTULI, I. *et al*. Genetic diversity and genome wide association study of β-glucan content in tetraploid wheat grains. **Plos One**, San Francisco, v. 11, p. 1-15, 2016.

MCMULLEN M. S.; DOEHLERT, D. C.; MILLER, J. D. Registration of 'HiFi' oat. **Crop Science**, Madison, v. 45, p. 1664, 2005.

MEIER, H; REID, J. S. G. Reserve polysaccharides other than starch in higher plants. In: LOEWUS, F. A.; TANNER, W. (Eds.). **Encyclopedia of Plant Physiology**, Springer. Berlin, 1982. p. 418-71.

MILLER, S. S. *et al.* Oat β-glucans: an evaluation of eastern canadian cultivars and unregistered lines. **Canadian Journal of Plant Science**, Ottawa, v. 73, p. 429-436, 1993.

MOLINA-CANO, J. L. *et al.* 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**, Dordrecht, v. 19, p. 275-284, 2007.

MURPHY, J. P.; HOFFMAN, L. A. Origin, history and production of oat. In: MARSHALL, H. G.; SORRELS, M. E. (Eds.). **Oat Science and Technology**. Madison: Crop Science Society of American, p. 1-28, 1992.

NEMETH, C. *et al*. Down-regulation of the *CslF6* gene results in decreased (1,3;1,4)-β-D-glucan in endosperm of wheat. **Plant Physiology**, Rockville, v. 152, p. 1209-1218, 2010.

NEWELL, M. A. *et al.* Genome-wide association study for oat (*Avena sativa* L.) beta-glucan concentration using germplasm of worldwide origin. **Theoretical and Applied Genetics**, Berlin, v. 125, p. 1687-1696, 2012.

PETERSON, D. M. Genotype and environment effects on oat beta-glucan concentration. **Crop Science**, Madison, v. 31, p. 1517-1520, 1991.

PETTOLINO, F. A. *et al*. Hyphal cell walls from the plant pathogen *Rhynchosporium secalis* contain (1,3; 1,6)-beta-D glucans, galacto and rhamnomannans, (1,3; 1,4) beta-D-glucans and chitin. **FEBS Journal**, Oxford, v. 276, p. 4122-4133, 2009.

RAJHATHY, T.; THOMAS, H. **Cytogenetics of oats**. Ottawa: Genetics Society of Canada, 1974. 90p.

REID, J. S. G. Cell wall storage carbohydrates in seeds - biochemistry of the seed "gums" and "hemicelluloses." **Advances in Botanical Research**, New York, v. 11, p. 125-155, 1985.

RICHMOND, T.; SOMERVILLE, C. The cellulose synthase superfamily. **Plant Physiology**, Rockville, v. 124, p. 495-498, 2000.

SAASTAMOINEN, M.; PLAAMI, S.; KUMPULAINEN, J. Genetic and environmental variation in β-glucan content of oats cultivated or tested in Finland. **Journal of Cereal Science**, London, v. 16, p. 279-290, 1992.

SAASTAMOINEN, M. Effects of environmental factors on the β-glucan content of two oat varieties. **Acta Agriculturae Scandinavica**, Stockholm, v. 45, p. 181-187, 1995.

SAASTAMOINEN, M. *et al.* β-glucan contents of groats of different oat cultivars in official variety, in organic cultivation, and nitrogen fertilization trials in Finland. **Agricultural and Food Science**, Jokioinen, v. 13, p. 68-79, 2004.

SCHELLER, H. V.; ULVSKOV, P. Hemicelluloses. **Annual Reviews of Plant Biology**, Palo Alto, v. 61, p. 263-289, 2010.

SKENDI, A. *et al*. Structure and rheological properties of water soluble β-glucans from oat cultivars of *Avena sativa* and *Avena bysantina*. **Journal of Cereal Science**, London, v. 38, p. 15-31, 2003.

SMITH, B. G; HARRIS, P. J. The polysaccharide composition of *Poales* cell walls: *Poaceae* cell walls are not unique. **Biochemical Systematics and Ecology**, Amsterdam, v. 27, p. 33-53, 1999.

SOMERVILLE, C. *et al*. Toward a systems approach to understanding plant cell walls. **Science**, Washington, v. 306, p. 2206-2211, 2004.

TAKETA, S. *et al*. Functional characterization of barley betaglucanless mutants demonstrates a unique role for CslF6 in (1,3;1,4)-β-D-glucan biosynthesis. **Journal of Experimental Botany**, Oxford, v. 63, p. 381-392, 2012.

TRETHEWEY, J. A. K.; CAMPBELL, L. M.; HARRIS, P. J. (1→ 3),(1→ 4)-ß-d-Glucans in the cell walls of the *Poales* (sensu lato): an immunogold labeling study using a monoclonal antibody. **American Journal of Botany**, Baltimore, v. 92, p. 1660-1674, 2005.

VIS, R. B.; LORENZ, K. β-Glucans: importance in brewing and methods of analysis. **LWT-Food Science and Technology**, London, v. 30, p. 331-336, 1997.

WHITEHEAD, A. *et al*. Cholesterol-lowering effects of oat β-glucan: a meta-analysis of randomized controlled trials. **American Journal of Clinical Nutrition**, New York, v. 100, p. 1413-1421, 2014.

WILSON, S. M. *et al.* Determining the subcellular location of synthesis and assembly of the cell wall polysaccharide (1, 3; 1, 4)-β-d-glucan in grasses. **The Plant Cell**, Rockville, v. 27, p. 754-771, 2015.

WOOD, P. J. Physicochemical characteristics and physiological properties of oat (1-3), (1-4)-β-D-Glucan. In: WOOD, P. J. (Ed). **Oat Bran**, American Association of Cereal Chemists. Saint Paul, 1993. p. 83-112.

WOOD, P.; WEISZ, J.; FEDEC, P. Potential for β-glucan enrichment in brans derived from oat (*Avena sativa* L.) cultivars of different (1-3), (1-4)-β-D-glucan concentrations. **Cereal Chemistry**, Saint Paul, v. 68, p. 48-51, 1991.

WOOD, P. J.; BEER, M. U. Productos funcionales de avena. In: MAZZA, G. (Ed.). **Alimentos funcionales: aspectos bioquímicos y de procesado**. Zaragoza: Acribia, 2000. p. 1-38.

YE, E. Q. *et al*. Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease and weight gain. **Journal of Nutrition**, Rockville, v. 142, p. 1304-1313, 2012.

# 3 CHAPTER 1: GENOME-WIDE ASSOCIATION FOR β-GLUCAN CONTENT, POPULATION STRUCTURE, AND LINKAGE DISEQUILIBRIUM IN ELITE OAT GERMPLASM ADAPTED TO SUBTROPICAL ENVIRONMENTS


## Abstract

High β-glucan content is one of the main goals of oat breeding programs worldwide. However, the genomic regions and genes controlling β-glucan content in oats are not fully understood to date. In this sense, the objectives of this study were: i) to characterize the Federal University of Rio Grande do Sul (UFRGS) Oat Panel population structure; and ii) to identify genomic regions associated with oat β-glucan content in germplasm adapted to subtropical environments. An oat panel with 413 genotypes was evaluated for β-glucan content under subtropical conditions in different years and genotyped using genotyping by sequencing. Population structure, linkage disequilibrium, and genome-wide association (GWA) analyses were carried out. GWA mapping was performed for each year separately and also considering in a multi-environment model. The UFRGS Oat Panel showed weak population structure and has great potential to elucidate many agronomic traits in subtropical conditions. Seven QTL associated with β-glucan content were identified. These QTL are located on Mrg02, Mrg06, Mrg11, Mrg12, Mrg19, and Mrg20. The QTL located on Mrg02, Mrg06, and Mrg11 seem to be genomic regions syntenic with barley. The use of these QTL may be useful to increase the genetic progress of oat β-glucan content in subtropical conditions.

### 3.1 Introduction

Oats (*Avena sativa* L.) have been described as one of the healthiest foods for human consumption. Oat kernels present complete nutritional properties and its consumption is associated with many benefits for human health (Chu *et al.*, 2013). One of the most important factors for the high nutritional quality of oat kernels is the β-glucan content. Oat kernels are characterized by high β-glucan levels when compared to wheat, maize, and rice (Fincher and Stone, 2004). β-glucan, which is a dietary fiber found in oat kernels, shows great benefits to human health when ingested in recommended quantities. Among these benefits are the improvement of immune system function and reduction of the glycemic index, cholesterol levels in the blood, obesity, and cardiovascular disease risk (Cho *et al.*, 2013; Whitehead *et al*., 2014; Ye *et al*., 2012).

Although β-glucan content has been described as one of the main goals of oat breeding programs worldwide (Peterson, 1991), improvement of this trait has been slow. There are many limitations to the development of oat varieties with desirable agronomic performance and high β-glucan content, including: i) the absence of a fast and reliable method for β-glucan content assessment; ii) the genetic factors associated with β-glucan content are not fully understood; and iii) the high influence exercised by environmental conditions on this trait. Considering these limitations, many oat breeding programs have not made substantial progress for β-glucan content to date. In southern Brazil, for example, great progress has been made for oat adaptation to subtropical conditions by selecting for flowering date and disease resistance. Since 1974, with the beginning of the UFRGS Oat Breeding Program, lines combining flowering earliness, crown rust resistance, and high grain yield were developed. On the other hand, selection for kernel-related traits has been successfully performed mainly for physical aspects to meet milling requirements.

The identification of the genetic loci associated with β-glucan content and their location in the oat genome may represent a key step for the genetic progress of this trait.

Some studies have been published involving molecular mapping for β-glucan content in oats. Quantitative trait loci (QTL) associated with β-glucan content were identified in two biparental populations, 'Kanota' x 'Ogle' and 'Kanota' x 'Marion', assessed in multiple environments in the United States (Kianian *et al*., 2000). These QTL explained from 7 to 20% of the phenotypic variance and two of the QTL were found to be in common for both mapping populations. Studying a hulless by covered spring oat biparental population, which was assessed in North America using two phenotyping methods, De Koeyer *et al*. (2004) detected five QTL associated with β-glucan content. The QTL detected by De Koeyer *et al*. (2004) explained 23% of the phenotypic variance for β-glucan content and some of the QTL were detected previously by Kianian *et al*. (2000). Two additive loci were identified using a double haploid oat mapping population 'Aslak' x 'Matilda' (AM), explaining over one third of the phenotypic variation of the β-glucan content trait (Tanhuanpää *et al.,* 2010). Finally, four QTL were detected for β-glucan content on the AM oat mapping population, accounting from 20 to 41% of the phenotypic variation (Tanhuanpää *et al.,* 2012). Two of these QTL were environment-specific and two of them appear to have been previously detected by Kianian *et al.* (2000). Three loci affecting β-glucan content were identified on two mapping populations assessed at three sites in Germany over a three-year period (Herrmann *et al.,* 2014). The most influential QTL was *QBg1.jki.A-1*, explaining 31% of the phenotypic variance, which is putatively homologous to KO_6 of the 'Kanota' x 'Ogle' map (Herrmann *et al.,* 2014).

Most studies involving identification of QTL associated with β-glucan content were conducted using biparental mapping populations. As consequence, the association between markers and QTL may be significant for only a specific genetic background. In addition, the QTL identified in these studies may also present a small effect on the phenotype, not fully explaining the phenotypic variation observed in non-related germplasm. In this sense, genome-wide association (GWA) is a new tool to potentially overcome these limitations.

GWA mapping is performed using a collection of genetically diverse inbred lines. As consequence, GWA mapping shows high statistical power for the detection of QTL with major effects on the phenotype (Huang *et al.*, 2014), exploiting alleles from more than two individuals as compared to biparental populations.

Some GWA mapping for β-glucan content have been published in oats (Newell *et al*., 2012, Asoro *et al.*, 2013). Using an oat panel composed by 431 genotypes, Newell *et al.* (2012) identified three markers associated with β-glucan content in oats. One of these markers showed high sequence homology to genes located on rice chromosome seven, in an adjacent region to CslF gene family members (Newell *et al*., 2012). Studying 446 oat lines genotyped with 1005 diversity array technology (DArT) markers, Asoro *et al.* (2013) identified 37 markers as the most important explaining β-glucan content. These markers were identified using single-marker tests and least absolute shrinkage and selection operator (LASSO). However, both studies used mostly North American oat germplasm, which is distinct and non-adapted to subtropical conditions. In this sense, the objectives of this study were: i) to characterize the UFRGS Oat Panel population structure; and ii) to identify genomic regions associated with oat β-glucan content in germplasm adapted to subtropical conditions.

## 3.2 Materials and methods

### *3.2.1 Germplasm - UFRGS Oat Panel*

An oat panel of 413 genotypes, including commercial cultivars and inbred lines, was assembled based on previous pedigree analysis. Prioritizing different pedigrees, it is expected that this panel shows a wide genetic variability. This oat panel, hereafter termed as UFRGS Oat Panel, is currently being assessed for many traits to undergo future association mapping. The UFRGS Oat Panel is composed of 384 hulled lines and 29 naked lines from different oat breeding programs. A total of 362 genotypes were developed by

the UFRGS Oat Breeding Program, from 1974 to 2015; 11 were developed by the Agronomic Institute of Paraná (IAPAR); 3 were developed by other Brazilian oat breeding programs (Agronomic Institute of Campinas, Federal University of Pelotas, and University of Passo Fundo); and 37 were developed by foreign oat breeding programs, mainly from North America (Louisiana State University, University of Florida, and North Dakota State University). All lines are part of the UFRGS Oat Breeding Program germplasm.

*3.2.2 Field experiments*

The field experiments were carried out in 2017 and 2018. In 2017, the UFRGS Oat Panel was assessed at the UFRGS Experimental Station (30° 07' S and 51° 41' W), located in Eldorado do Sul city, Rio Grande do Sul state, and at the IAPAR Experimental Station (23° 21' S and 51° 10' W), located in Londrina city, Paraná state, both in Southern Brazil. The experiments were conducted using a randomized complete block design, with two replications. Each experimental unit was composed of a single 1 m long row, spaced 0.25 m apart each other. Planting was carried out manually at a rate of 50 and 60 seeds per plot for hulled and naked genotypes, respectively. The base fertilization in Londrina was 275 kg $ha^{-1}$ of a 5-20-20 N–P–K formula, and topdressing nitrogen, in the form of dry urea, was applied once, when plants showed four extended leaves, at a rate of 20 kg $ha^{-1}$ of N. In Eldorado do Sul, the base fertilization was 300 kg $ha^{-1}$ of a 5-30-15 N-P-K formula, and topdressing nitrogen, in the form of dry urea, was applied twice, when plants showed four and six extended leaves, at a rate of 33 kg $ha^{-1}$ of N per application.

In 2018, the UFRGS Oat Panel was assessed only at the UFRGS Agronomic Station, in Eldorado do Sul city, Rio Grande do Sul state, Southern Brazil. The experiment was conducted in a randomized complete block design, with two replications. Each experimental unit was composed by two 2 m long rows, spaced 0.2 m apart each other and 0.4 between plots. Planting was carried out mechanically at a rate of 60 and 80 seeds per linear m for hulled and naked genotypes, respectively. The base fertilization was 300 kg

$ha^{-1}$ of a 5-30-15 N-P-K formula, and topdressing nitrogen, in the form of dry urea, was applied twice, when plants showed four and six extended leaves, at a rate of 33 kg $ha^{-1}$ of N per application. Pests, fungal diseases, and weeds were chemically controlled in all experiments. At the field maturity stage, approximately 25 panicles of each experimental unit were collected for phenotypic analysis.

*3.2.3 Sample preparation and β-glucan content analyses*

Panicles were threshed and 50 g of kernels were dehulled using a Wintersteiger LD 180 laboratory thresher (Wintersteiger, Ried, Austria). A subsample of 15 g of dehulled kernels per line was ground in a Willey type mill (Marconi, Piracicaba, Brazil) using a 0.5 mm sieve. After milling, each whole flour sample was well homogenized using a set of sieves and stored at -20 ºC until β-glucan quantification. Chemical analyses for β-glucan were carried out at the Center for Food Research of the University of Passo Fundo, Passo Fundo city, Rio Grande do Sul state, Southern Brazil. β-glucan content was quantified for 408 samples from a single plot from Londrina using the mixed-linkage β-glucan assay kit (Megazyme International Ireland Ltd, Wicklow, Ireland). The β-glucan content analysis was performed in duplicates, based on McCleary and Codd (1991), according to AACC 32-23.01 and AOAC 995.16 methods, with results expressed in dry basis.

The laboratory data obtained using enzymatic analyses was used to build a near-infrared spectroscopy (NIRS) model for β-glucan content prediction, which was used for the phenotypic screening of all experiments. Samples were kept at room temperature (25 °C) for 4 h before NIRS prediction to balance the moisture and temperature to avoid any influence of these factors on the reflectance and absorbance. About 15 g of oat flour from each genotype, obtained from a single plot, was screened using a NIRS™ DS2500 analyzer (Foss, Hillerod, Denmark).

*3.2.4 DNA isolation*

Fifteen seeds of each genotype were obtained from a randomly selected panicle. The seeds were germinated in wet germination paper at the Department of Crop Science of UFRGS. The seeds were placed in a Biochemical Oxygen Demand (BOD) chamber at 25 ºC without light. After seven days, 100 mg of fresh tissue were collected of each genotype and immediately dry-freeze lyophilized for 48 h and stored at -20 ºC. Samples were then shipped to the University of Minnesota Genomics Center (UMGC) for DNA isolation. Genomic DNA was isolated using mechanical tissue disruption and the DNeasy® Plant DNA isolation kit (Qiagen, Valencia, USA). DNA quantification was performed using the fluorescent dye PicoGreen® and NanoDrop™ spectrophotometer.

*3.2.5 Genotyping by sequencing*

The genotyping by sequencing (GBS) was performed at the UMGC, following a two-enzyme protocol adapted from Poland *et al.* (2012). Two GBS libraries were constructed in 384-plex, with four random blank wells for quality control purposes. The DNA was digested with the restriction enzymes PstI (CTGCAG) and MspI (CCGG) and barcode adapters were ligated to individual samples. The libraries were amplified by a polymerase chain reaction and cleaned before sequencing. Each library was sequenced using a NovaSeq™ 6000 (Illumina, San Diego, USA), producing individual 100-bp single-end reads. The sequence data was trimmed using the Trimmomatic software (Bolger *et al.,* 2014), removing all adapters. The quality control of reads was verified before and after trimming using the FastQC software (Andrews, 2010).

*3.2.6 SNP calling and filtering*

Genotypes were called using a *de novo* pipeline, which operates without the requirement of a reference genome. The single-nucleotide polymorphisms (SNP) calling was performed using the Haplotag software (Tinker *et al.*, 2016). Initially, the Universal Network Enabled Analysis Kit (UNEAK) pipeline (Lu *et al.*, 2013), available on TASSEL

3 software (Bradbury *et al.*, 2007), was used to demultiplex FastQ files into tag count files for each taxa. Following this, the tag counts and merged tag count files produced by the UNEAK pipeline were used in a Haplotag discovery pipeline with default parameters (Tinker *et al.*, 2016). All SNPs were aligned to most recent oat consensus map (Bekele *et al.* 2018) using a Basic Local Alignment Search Tool (BLAST; Altschul *et al.*, 1990). SNP markers were filtered using the following criteria: i) biallelic; ii) missing data < 50%; iii) minor allele frequency (MAF) > 2%; and iv) heterozygosity < 10%. The Pearson's correlation coefficient ($r$) was calculated for each SNP pairwise combination of SNPs. Markers highly correlated ($r > 0.99$) were removed to avoid redundant information and any bias in future analyses. After filtering steps, 10,169 SNPs were retained in the final matrix. Missing genotypes were imputed with the 'expectation-maximization' (EM) method, using the 'A.mat' function from the rrBLUP package (Endelman, 2011), available on R software (R Development Core Team, 2008).

### *3.2.7 Population structure and linkage disequilibrium analyses*

Analyses were performed to characterize the UFRGS Oat Panel and to determine parameters for GWA. Population structure was characterized by principal component analysis (PCA) using the base R function 'prcomp'. The amount of variance explained by the first ten principal components was visualized in a scree plot and the first two principal components were visualized in a scatterplot. Analysis of molecular variance (AMOVA), described by Excoffier *et al.* (1992), was conducted to characterize the amount of genetic variation explained by the division of the UFRGS Oat Panel into 'UFRGS' and 'other' subpopulations. AMOVA was performed using the pegas R package (Paradis, 2010) with 200 permutations.

Pairwise linkage disequilibrium (LD) was estimated as the squared allele frequency correlation ($r^2$) based on non-imputed genotype data using the R package LDcorSV (Mangin *et al.*, 2012). Markers from each of the twenty-one linkage groups of the oat

consensus map (Bekele *et al.,* 2018) were assumed to be independent from markers placed in other linkage groups. In this sense, $r^2$ was estimated for each linkage group separately, allowing a useful comparison of LD among the linkage groups.

*3.2.8 Genome-wide association mapping*

GWA mapping was conducted for β-glucan content for each trial separately and in a multi-environment model. β-glucan content assessed by the enzymatic method was not included in the multi-environment model because this phenotype was obtained in a different way, which could lead some bias to the analysis. Genome-wide association analyses were performed using the 'GWAS' function from the rrBLUP R package (Endelman, 2011). A mixed-model was employed accounting for population structure and kinship (Q + K model). Population structure was accounted for using the first three principal components. Kinship was accounted using a kinship matrix calculated using the 'A.mat' function from the rrBLUP R package (Endelman, 2011). Population structure, kinship, and trial (in the multi-environment model) were fit as fixed effect. Quantile-quantile (Q-Q) plots of the observed versus expected p-values under the null hypothesis were visualized for each analysis to verify signs of statistical test inflation. The statistical significance threshold was established as p-value < 0.001. This threshold was determined because this study is the first exploratory analysis of β-glucan content in subtropical conditions. In addition, the β-glucan content is genetically complex and presents quantitative inheritance. QTL were reported only if significant in at least two single-environment analyses.

*3.2.9 Comparative mapping*

Considering that a small fraction of GBS-SNP markers is located on genome coding regions, the use of these markers directly for comparative mapping is not usual. However, GBS-SNP markers are in high LD with other genetic markers which are located

in coding regions and, therefore, are useful for comparative mapping among species. High LD is easily observed on the oat consensus map (Bekele *et al.*, 2018). Many GBS-SNP markers and markers from the 6K BeadChip Illumina Infinium array share the same genetic position in the oat consensus map (Bekele *et al.*, 2018). In this sense, considering the high LD between SNP markers detected in the present study and other markers, an alternative approach was performed to identify syntenic regions associated with β-glucan content between barley and oat.

SNP markers from the 6K BeadChip Illumina Infinium array, which are located in oat coding regions (Bekele *et al.*, 2018), and share the same genetic position of the SNP markers presented in this study, were aligned to the *Hordeum vulgare* subsp. *vulgare* genome, available at https://plants.ensembl.org/Hordeum_vulgare/Info/Index. BLAST (Altschul *et al.*, 1990) was used to verify the nucleotide similarity between oat and barley, considering query cover > 90% and e-value < $10^{-20}$. Candidate genes associated with β-glucan content in barley were also identified using nucleotide sequences available at GenBank (https://www.ncbi.nlm.nih.gov/genbank) and morexGenes - Barley RNA-seq Database (https://ics.hutton.ac.uk/morexGenes). In addition, QTL associated with β-glucan content identified by Islamovic *et al.* (2013), Houston *et al*. (2014), and Mohammadi *et al.* (2014) were located in the barley genome. Oat and barley marker sequences were obtained from the T3/Oat (https://triticeaetoolbox.org/oat) and T3/Barley (https://triticeaetoolbox.org/barley) databases. Based on the alignment results, barley chromosomes were represented to illustrate the synteny between oat and barley using the MapChart software (Voorrips, 2002).

## 3.3 Results

### *3.3.1 Phenotypic variation for β-glucan content*

The UFRGS Oat Panel show wide phenotypic variation for β-glucan content in all assessed environments, regardless of the quantification method (Figure 1). Average β-glucan content varied from 4.2 to 4.4% among the environments (Figure 1). An important difference was observed between the results of the enzymatic and NIRS methods in Londrina. Both methods showed approximately the same average, but the NIRS prediction reduced the phenotypic variance of the genotypes (Figure 1).

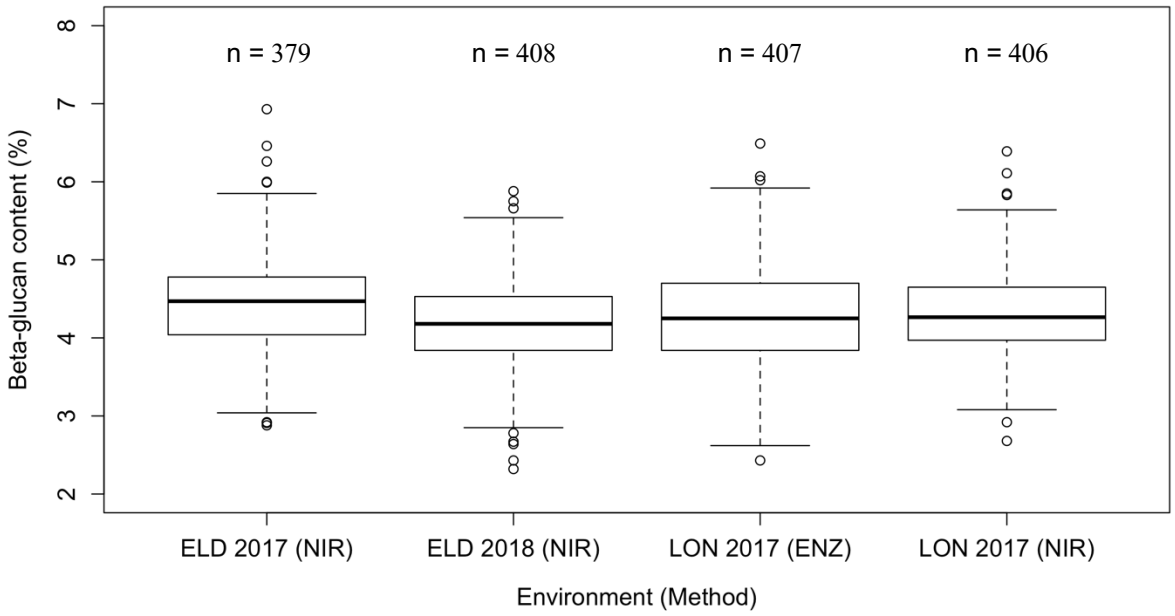



FIGURE 1. Phenotypic variation for β-glucan content in Eldorado do Sul and Londrina considering enzymatic and NIRS methods. ELD, Eldorado do Sul; LON, Londrina, ENZ, enzymatic method; and NIRS, near-infrared spectroscopy.

In 2017, the environmental conditions in Eldorado do Sul were stressful due to dry conditions during the field establishment and later with irregular rainfalls during the reproductive stage. The environmental conditions observed in Eldorado do Sul during 2018 were favorable to oat cultivation, resulting in lower β-glucan content average when compared to 2017 (Figure 1).

*3.3.2 Population structure of the UFRGS Oat Panel*

PCA conducted on the UFRGS Oat Panel reveals a weak population structure. A scree plot of the principal components revealed that the first ten principal components account for 32.65% of the genetic variance. The principal components have smooth decrease in the variation explained after the third principal component (Figure 2). These results indicate that three principal components are a good cutoff to account for population structure in genome-wide association analyses using the UFRGS Oat Panel.

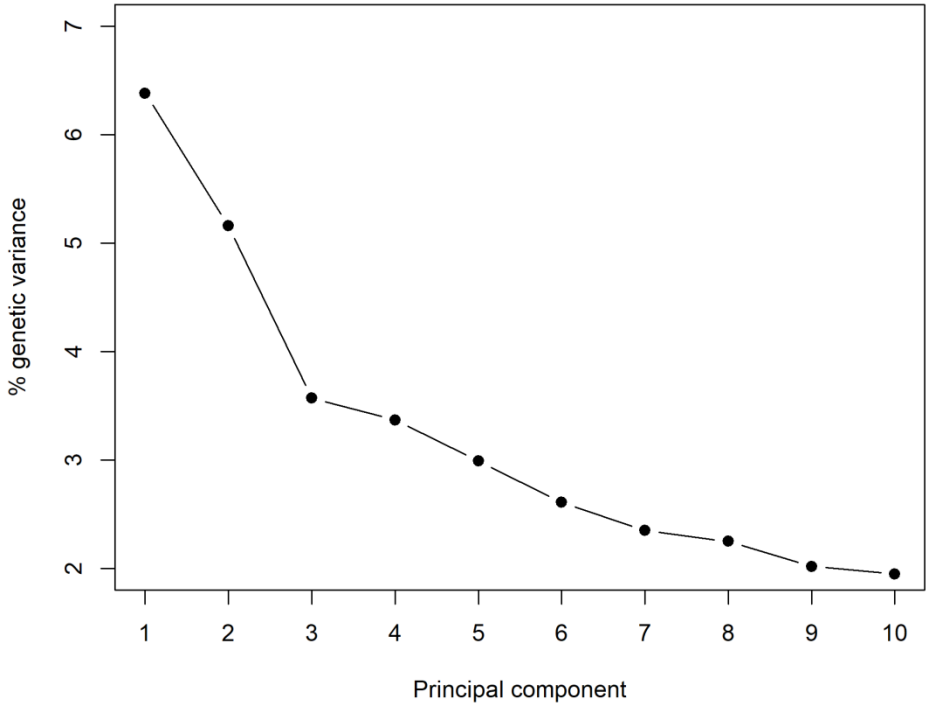



FIGURE 2. Genetic variance explained by the first ten principal components of the UFRGS Oat Panel.

The first two principal components, which account for 11.54% of the genetic variance, are presented in Figure 3. Visualization of the two principal components reinforce that the UFRGS Oat Panel has a weak structure, with genotypes from different oat breeding programs scattered among UFRGS genotypes. Most genotypes from the IAPAR and other Brazilian oat breeding programs are centered, surrounded by the UFRGS lines and the lines from foreign breeding programs (Figure 3).

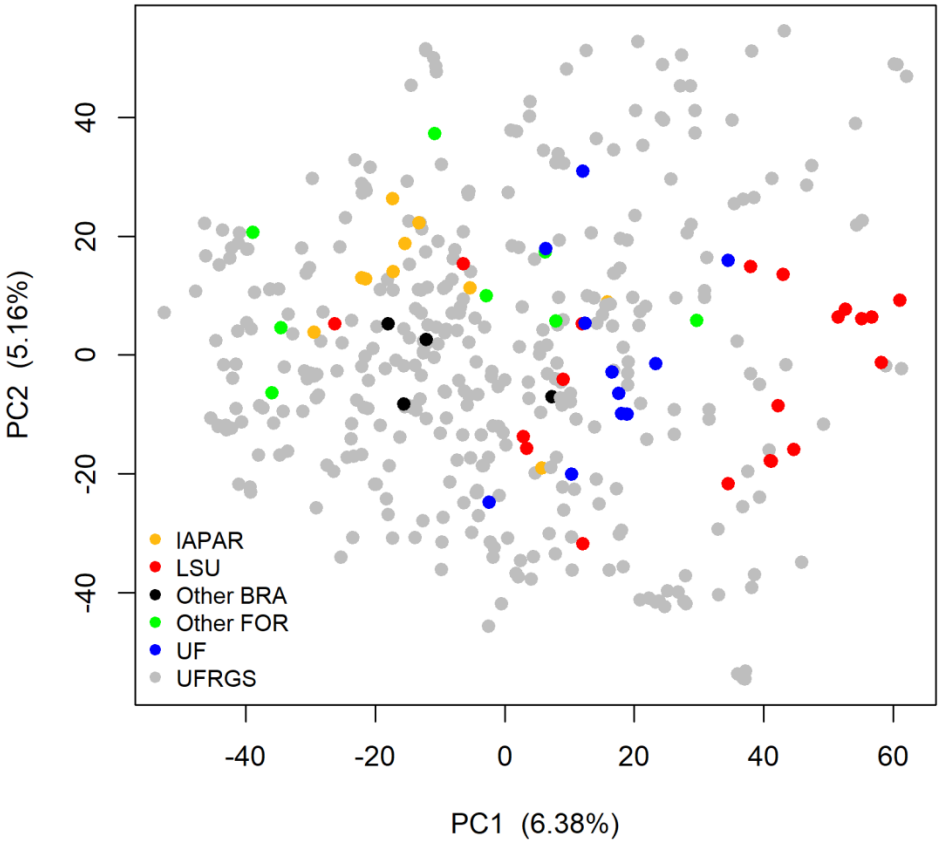



FIGURE 3. Principal component analysis of the UFRGS Oat Panel. The first two principal components account for 11.54% of the genetic variance. IAPAR, Agronomic Institute of Paraná; LSU, Louisiana State University; Other BRA, other Brazilian oat breeding programs; Other FOR, other foreign oat breeding programs; UF, University of Florida; UFRGS, Federal University of Rio Grande do Sul; PC, principal component.

The lines developed by the University of Florida and Louisiana State University, which are part of the North America southern oat subpopulation, are closer to each other and clustered mostly with UFRGS genotypes. AMOVA of the UFRGS Oat Panel revealed that the division between 'other' and 'UFRGS' subpopulations accounts for just 2.78% ($p < 0.001$) of the genetic diversity.

### *3.3.3 Linkage disequilibrium*

LD analyses were performed to identify different patterns of LD decay among oat linkage groups. In this sense, LD information would be useful to support GWA mapping results and also to clarify some findings. LD decays of all oat consensus linkage groups were plotted together (Figure 4). LD for each linkage group separately is available in the

Supplementary Figure 1. As expected, LD decay as the genetic distance on the oat consensus map increases (Figure 4).

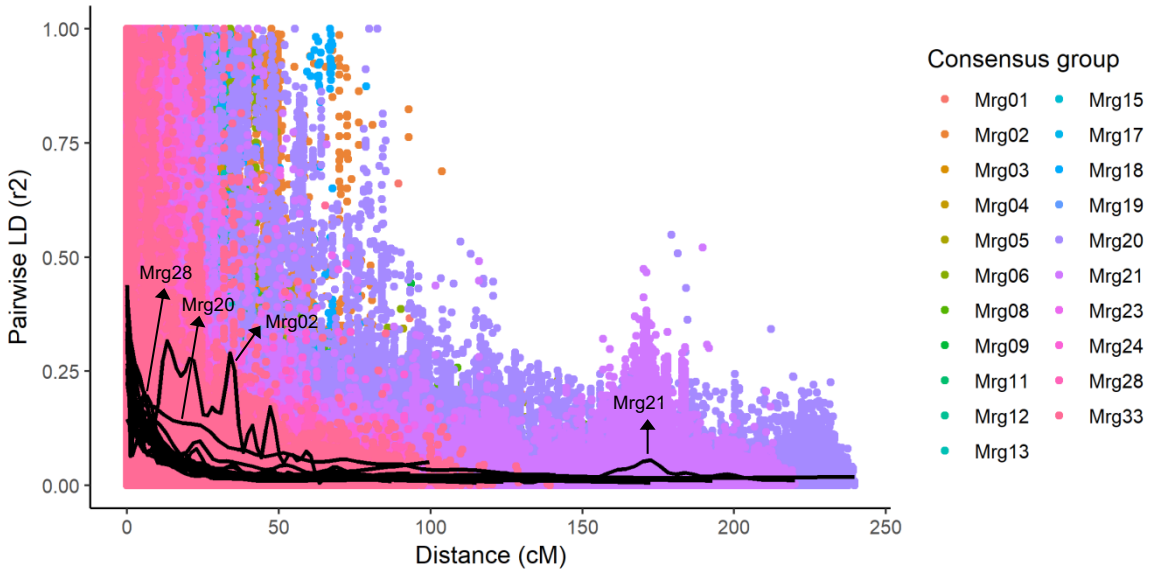



FIGURE 4. Linkage disequilibrium (LD) scatter plot showing correlations ($r^2$) between marker pairs inside each consensus group as a function of genetic position; a smoothing spline (black lines) is fit to each consensus group data.

Most of the consensus linkage groups presented similar behavior for LD, with a quick decay after a few cM. Considering the genome-wide disequilibrium decay, LD between linked markers decayed quickly to $r^2 = 0.2$ for marker pairs with a map distance of 2.2 cM. However, some consensus linkage groups do not follow this pattern. For example, the linkage group Mrg02 showed many blocks of markers in high LD between 0 cM and 50 cM positions, as well as on the Mrg28 between 0 and 10 cM (Figure 4 and Supplementary Figure 1). On the Mrg20, a delay on LD decay was observed, while on Mrg21 a block of markers in high LD was observed between the 150 cM and 200 cM positions (Figure 4 and Supplementary Figure 1).

*3.3.4 Genome-wide association for β-glucan content*

GWA mapping for β-glucan content was performed for each single environment and also considering a multi-environment model. β-glucan content was assessed using enzymatic and NIRS prediction methods. The phenotypic data from each method was

analyzed separately. The results are presented initially for the trial assessed by the enzymatic method, followed by all other single trials assessed using NIRS. Finally, the multi-environment analysis using NIRS data will be presented. QTL were reported if significant in at least two individual models.

Two QTL associated with β-glucan content were identified using the enzymatic data from Londrina 2017. These QTL are located on the Mrg02 and Mrg20 linkage groups (Figure 5a). The QTL located on the Mrg02 is associated with the marker avgbs2_50181.1.51, which is located at the position 37.5 cM (Table 1). The QTL detected on Mrg20 is associated with the marker avgbs_cluster_35845.1.61, which is located at the position 79.7 cM (Table 1). Four additional β-glucan content QTL were identified using the NIRS data from Londrina 2017, using the same samples employed for the enzymatic method (Figure 5b). These QTL are located on the Mrg02, Mrg11, Mrg19, and Mrg20 oat consensus groups. The QTL identified on the Mrg02 is associated with the marker avgbs_cluster_30914.1.60, which is located at the position 85.2 cM (Table 1). Seven markers are associated with the QTL located on the Mrg11, covering a region of 5.1 cM. The markers avgbs2_86064.1.51, avgbs2_86064.1.6, avgbs_200620.1.64, and avgbs2_44135.1.36 are located at the position 3.7 cM on the Mrg11 (Table 1). In this same consensus group, the marker avgbs_42869.1.46, avgbs_104028.1.28, and avgbs2_25762.1.10 were also associated with β-glucan content in oats. The marker avgbs_42869.1.46 is mapped at the position 7.2 cM on the Mrg11, while the other ones are located at the 8.8 cM position in the same oat consensus group (Table 1). The QTL identified on the Mrg19 is associated with the marker avgbs_cluster_32091.1.36, which is mapped at the position 76.3 cM. On the Mrg20, the marker avgbs_cluster_35845.1.61, which was reported previously acting on β-glucan content using the enzymatic data, was also detected using the NIRS method (Table 1). This marker is mapped at the position 79.7

cM position and represents the only QTL detected by both β-glucan quantification methods.

TABLE 1. Genomic regions controlling β-glucan content in elite oat germplasm adapted to subtropical conditions.

| Location | Phenotype | Marker | Mrg† | Position† (cM) | -log pvalue |
|---|---|---|---|---|---|
| LON 2017 | Enzymatic | avgbs2_50181.1.51 | **Mrg02** | **37.50** | 3.82 |
| LON 2017 | Enzymatic | avgbs_cluster_35845.1.61 | Mrg20 | 79.70 | 3.21 |
| LON 2017 | NIRS | avgbs_cluster_30914.1.60 | **Mrg02** | **85.20** | 3.29 |
| LON 2017 | NIRS | avgbs2_86064.1.51 | **Mrg11** | **3.70** | 4.35 |
| LON 2017 | NIRS | avgbs2_86064.1.6 | **Mrg11** | **3.70** | 4.20 |
| LON 2017 | NIRS | avgbs_200620.1.64 | **Mrg11** | **3.70** | 3.73 |
| LON 2017 | NIRS | avgbs2_44135.1.36 | **Mrg11** | **3.70** | 3.54 |
| LON 2017 | NIRS | avgbs_42869.1.46 | Mrg11 | 7.20 | 3.27 |
| LON 2017 | NIRS | avgbs_104028.1.28 | **Mrg11** | **8.80** | 3.11 |
| LON 2017 | NIRS | avgbs2_25762.1.10 | **Mrg11** | **8.80** | 3.07 |
| LON 2017 | NIRS | avgbs_cluster_32091.1.36 | Mrg19 | 76.30 | 3.55 |
| LON 2017 | NIRS | avgbs_cluster_35845.1.61 | Mrg20 | 79.70 | 3.08 |
| ELD 2017 | NIRS | avgbs_cluster_39442.1.33 | Mrg06 | 62.90 | 3.64 |
| ELD 2017 | NIRS | avgbs2_96603.1.42 | **Mrg11** | **52.80** | 3.02 |
| ELD 2017 | NIRS | avgbs_cluster_33876.1.64 | Mrg12 | 56.70 | 3.34 |
| ELD 2018 | NIRS | avgbs_cluster_39442.1.33 | Mrg06 | 62.90 | 4.03 |
| ELD 2018 | NIRS | avgbs_104028.1.28 | **Mrg11** | **8.80** | 3.88 |
| ELD 2018 | NIRS | avgbs_62906.1.35 | Mrg11 | 39.20 | 3.20 |
| ELD 2018 | NIRS | avgbs_225806.1.13 | Mrg11 | 45.60 | 3.18 |
| ELD 2018 | NIRS | avgbs_72741.1.47 | Mrg11 | 49.80 | 3.92 |
| ELD 2018 | NIRS | avgbs_cluster_33876.1.64 | Mrg12 | 56.70 | 4.27 |
| ELD 2018 | NIRS | avgbs_cluster_30172.1.9 | Mrg12 | 59.40 | 3.37 |
| ELD 2018 | NIRS | avgbs_cluster_32091.1.36 | Mrg19 | 76.30 | 3.15 |
| Multi-e | NIRS | avgbs2_50181.1.51 | **Mrg02** | **37.50** | 3.21 |
| Multi-e | NIRS | avgbs_cluster_39442.1.33 | Mrg06 | 62.90 | 4.01 |
| Multi-e | NIRS | avgbs_70879.1.19 | **Mrg06** | **72.10** | 3.30 |
| Multi-e | NIRS | avgbs_cluster_13360.1.22 | Mrg11 | 37.90 | 3.16 |
| Multi-e | NIRS | avgbs_72741.1.47 | Mrg11 | 49.80 | 3.05 |
| Multi-e | NIRS | avgbs2_96603.1.42 | **Mrg11** | **52.80** | 3.02 |
| Multi-e | NIRS | avgbs_cluster_33876.1.64 | Mrg12 | 56.70 | 3.23 |
| Multi-e | NIRS | avgbs_cluster_30172.1.9 | Mrg12 | 59.40 | 3.31 |
| Multi-e | NIRS | avgbs_cluster_39780.1.13 | Mrg19 | 70.10 | 3.48 |

†Based on Bekele *et al.* (2018). QTL were detected considering an arbitrary threshold (p-value ≤ 0.001; $-\log_{10}(p) \leq 3$). Multi-e, multi-environment mapping. Positions in bold were used for comparative mapping between oats and barley. LON, Londrina; ELD, Eldorado do Sul; Multi-e, multi-environment; NIRS, near infrared spectroscopy.

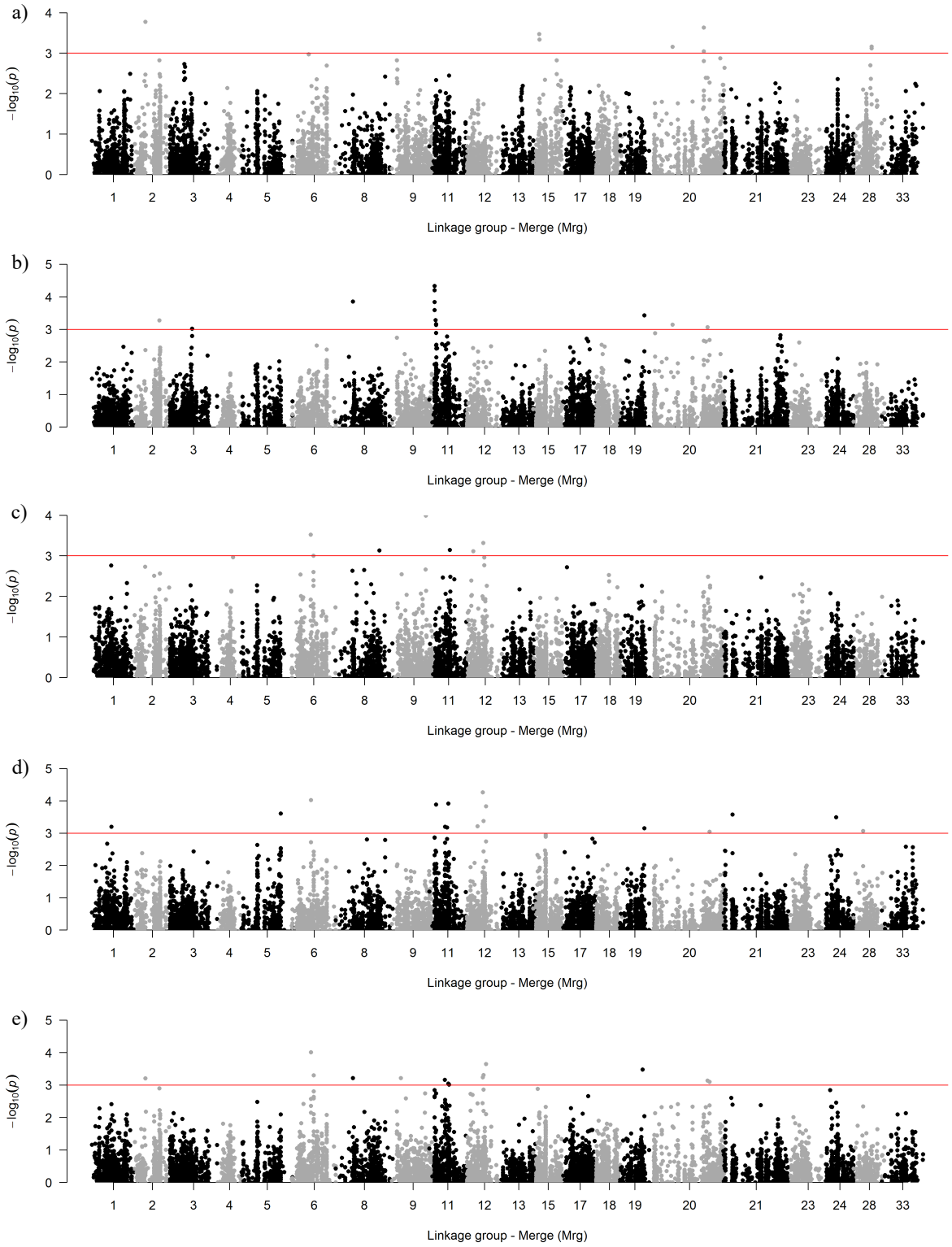



FIGURE 5. Manhattan plots for oat markers associated with β-glucan content. The red horizontal line represents a $-\log_{10}(p) \leq 3$ (p-value ≤ 0.001); a, Londrina 2017 (enzymatic); b, Londrina 2017 (NIRS); c, Eldorado do Sul 2017 (NIRS); d, Eldorado do Sul 2018 (NIRS); e, Multi-environment (NIRS).

GWA mapping for β-glucan content in Eldorado do Sul using the NIRS data from 2017 revealed three QTL associated with the trait. QTL associated with β-glucan content

were identified on the linkage groups Mrg06, Mrg11, and Mrg12 (Figure 5c). The QTL detected on Mrg06 is associated with the marker avgbs_cluster_39442.1.33, which is located at the 62.9 cM position (Table 1). The QTL detected on Mrg11 is associated with the marker avgbs2_96603.1.42, which is located at the 52.8 cM position. This region does not correspond to the QTL previously reported in Londrina. A final QTL was identified on Mrg12 associated with the marker avgbs_cluster_33876.1.64, which is located at the 56.7 cM position (Table 1).

Five QTL were detected using the NIRS method from Eldorado do Sul 2018. Two of these QTL were also significant in this same location in 2017 (Table 1). The QTL located on Mrg06, associated with the marker avgbs_cluster_39442.1.33, was also significant in 2018 (Table 1). Similarly, the QTL detected on Mrg12, associated with the marker avgbs_cluster_33876.1.64, was also signicant in 2018. The marker avgbs_cluster_30172.1.9, which is mapped at the 59.4 cM position in the same consensus group, was associated with β-glucan content in 2018. In this sense, this QTL is covering a region of 2.3 cM in Eldorado do Sul during 2018 (Table 1). Two QTL were identified on Mrg11, one located at the 8.8 cM position and the other one associated with multiple markers in different positions (Table 1). The first QTL associated with the marker avgbs_104028.1.28, which is mapped at the position 8.8 cM, was also detected for Londrina in 2017 (Table 1). The second QTL is associated with the markers avgbs_62906.1.35, avgbs_225806.1.13, and avgbs_72741.1.47, which are mapped at the positions 39.2 cM, 45.6 cM, and 49.8 cM, respectively. These markers span a region of 10.6 cM (Table 1). The fifth QTL was associated with the marker avgbs_cluster_32091.1.36 on Mrg19, which is mapped at the 76.3 cM position, and was also associated with β-glucan content in Londrina during 2017 (Table 1).

Similar β-glucan content QTL were detected using the multiple-environment GWA model compared to the individual environment analyses. A new region was detected on

Mrg06, associated with the marker avgbs_70879.1.19, which is mapped at the 72.1 cM position (Table 1). In addition, the QTL detected on Mrg19 in Londrina 2017 and Eldorado do Sul 2018 mappings was associated with the marker avgbs_cluster_39780.1.13, located at the 70.1 cM position (Table 1).

*3.3.5 Comparative mapping between barley and oats*

The comparative mapping analysis showed important similarities between oat and barley (Figure 6). Some QTL associated with β-glucan content identified in this study (Table 1) are close to important genes and QTL acting on this same trait in barley (Figure 6).

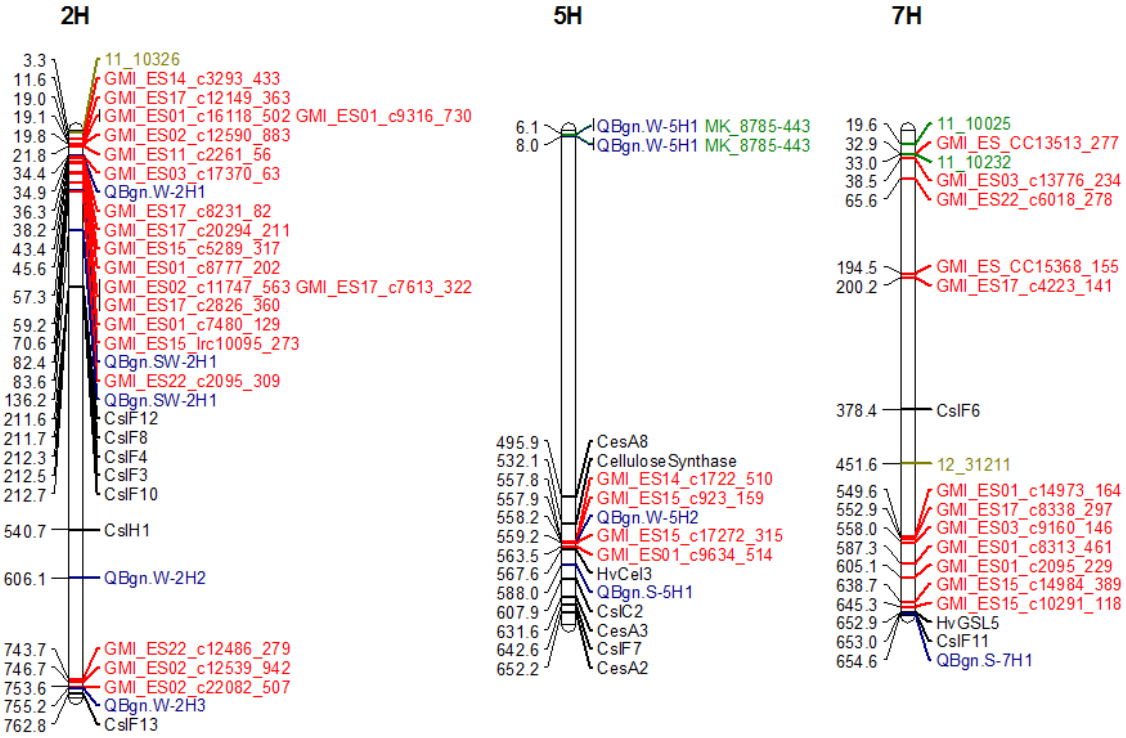



FIGURE 6. Comparative mapping between oat markers and barley chromosomes 2H, 5H, and 7H. Chromosomes 2H, 5H, and 7H showed similarities with Mrg11, Mrg06, and Mrg02, respectively. Red names are oat markers located on the same position of QTL associated with β-glucan content presented in this study. Green, blue, and golden names represent QTL associated with barley β-glucan content identified by Islamovic *et al.* (2013), Houston *et al.* (2014), and Mohammadi *et al.* (2014), respectively. Black names show candidate genes associated with β-glucan content in barley. Distances are show in Mbp.

The oat markers associated with β-glucan content on Mrg11 (Table 1) are syntenic with barley chromosome 2H (Figure 6). Several β-glucan content QTL have been

previously reported in this region of the barley genome including QBgn.W-2H1, QBgn.SW-2H1, and QBgn.W-2H3 (Houston *et al.*, 2014), and 11_10326 (Mohammadi *et al.,* 2014; Figure 6). In addition, a cluster of CslF genes is mapped on chromosome 2H (Figure 6). On barley chromosome 5H, nucleotide similarities were observed with oat markers mapped on Mrg06 (Table 1). The oat markers located on Mrg06 are close to the barley QTL QBgn.W-5H2 identified by Houston *et al.* (2014), which is associated with β-glucan content. Similarly, oat markers located on Mrg02 (Table 1) showed nucleotide similarities with barley chromosome 7H (Figure 6). These markers were located near to the QBgn.S-7H1 (Houston et al., 2014), 11_10025, and 11_10232 (Islamovic *et al.*, 2013) QTL (Figure 6), which are associated with β-glucan content in barley.

### 3.4 Discussion

The UFRGS Oat Panel showed a wide phenotypic variation for β-glucan content, ranging from 2.5 to 7.0 % according to the environmental condition/location, with an overall average around 4.3 % (Figure 1). Differences were revealed between the enzymatic and NIRS β-glucan content quantification methods. Comparing both methods, the NIRS method resulted in less variance than the enzymatic method (Figure 1). In this sense, the NIRS method is less desirable for phenotypic selection because of the decreased variance of the observations. The differences between the enzymatic and NIRS methods may imply in the identification of loci associated with β-glucan content. The new regions detected using NIRS method may be a consequence of false differences predicted for the β-glucan phenotype. In Londrina, for example, some differences were observed between enzymatic (Figure 5a) and NIRS results (Figure 5b). As consequence, these regions found exclusively using NIRS method may be useless for molecular breeding. In the present study, even using a NIRS curve built with more than 350 samples (unpublished results), additional filters were used to reduce the presence of false positive markers. The association between the enzymatic and NIRS methods is shown in the Supplementary Figure 2.

The phenotypic data also suggests an environmental influence on β-glucan content. This result is supported considering the contrast of Eldorado do Sul between years, as well as differences observed between Eldorado do Sul and Londrina in 2017 (Figure 1). The differences between Londrina and Eldorado do Sul environments are shown in the Supplementary Figure 3. Considering the environmental condition records, the results suggest that β-glucan content is higher under stressful conditions. In 2017, a stressful condition occurred after planting in Eldorado do Sul, with low and uneven rainfall during the vegetative stage, excess of rainfall during the reproductive stage, and high daily maximum temperatures during the reproductive stage. The environmental conditions were favorable in this same location in 2018, with well distributed rainfall throughout the year and lower temperature averages when compared to 2017, especially in the second half of the crop cycle. Many studies have been published indicating the environmental influence on β-glucan content in oats (Saastamoinen *et al*., 1992; Miller et al., 1993; Saastamoinen, 1995; Herrera *et al.*, 2016). The occurrence of successive rainfalls during grain filling tend to reduce the β-glucan content in oat kernels (Miller *et al*., 1993). On the other hand, in particular environments, including Canada, rainfalls during the vegetative stage are positively correlated with the β-glucan content in oats (Herrera *et al*., 2016). In Canada and Finland, high temperature conditions during the formation of kernels also tend to increase the oat β-glucan content (Saastamoinen *et al*., 1992; Miller *et al*., 1993; Saastamoinen, 1995). However, the results presented in this study are only suggestive, once statistical analyses were not performed on β-glucan content. Future studies involving the UFRGS Oat Panel may confirm the environmental influence on β-glucan content.

The PCA showed a weak population structure on the UFRGS Oat Panel. The first ten principal components account for only 32.65% of the genetic variance (Figure 2). In addition, no structure was visually observed by plotting the first two principal components (Figure 3). Groups from different oat breeding programs clustered relatively together,

while some of them overlapped each other (Figure 3). This result is explained because most genotypes included in this study are adapted to subtropical environmental conditions, including those ones from North America, which corresponds to the North American southern subpopulation (Esvelt Klos *et al.*, 2016; McNish *et al.*, 2019). Genotypes from the UFRGS Oat Panel were widely distributed and occupied all parts of the PCA plot (Figure 3), justifying the AMOVA results. This result suggests that: i) the UFRGS oat germplasm may provide genetically diverse germplasm for other oat breeding programs Brazil and abroad; and ii) the UFRGS oat germplasm and United States Southern subpopulation are genetically similar and reveal germplasm exchange between institutions. The absence of population structure in the UFRGS Oat Panel confirms that this panel is promising for future mapping studies involving other agronomic traits under subtropical conditions.

The LD analysis revealed a rapid LD decay to $r^2 = 0.2$ for marker pairs at least 2.2 cM apart on the 2018 oat consensus map. The linkage groups Mrg02, Mrg20, Mrg21, and Mrg28 showed some blocks in high LD, differing from the behavior observed for the other linkage groups (Figure 4 and Supplementary Figure 1). The identification of these blocks is useful to discriminate candidate QTL in the same linkage group. High degree of association among many unlinked and loosely-linked markers were observed by Esvelt Klos *et al.* (2016) on Mrg02, Mrg11, Mrg15, and Mrg28. Recently, McNish *et al.* (2019) also identified many blocks with high LD on Mrg02, Mrg20, Mrg21, and Mrg28.

GWA mapping identified seven candidate QTL controlling β-glucan content in oats, which are distributed on Mrg02, Mrg06, Mrg11, Mrg12, Mrg19, and Mrg20, under subtropical conditions (Table 1 and Figure 5). Some previous studies have been performed for β-glucan content in oats. However, they evaluated mostly North American oat germplasm, not involving genotypes adapted to subtropical conditions. In this sense, this study is the first genome-wide association mapping study to focus exclusively on oat

genotypes adapted to subtropical conditions. As consequence, besides identifying candidate regions acting on β-glucan content in subtropical environments, this study may also identify genomic regions in common for β-glucan content between contrasting germplasm.

The oat marker avgbs_cluster_30914.1.60, which was associated with β-glucan content in Londrina during 2017, is mapped at the 85.2 cM position on Mrg02. The marker GMI_ES15_c10291_118, which is also mapped at the 85.2 cM position on Mrg02, is flanking the *CslF6-D* gene in oats (Coon, 2012). A single sequence repeat marker identified on *CslF6-D* non-coding region was screened on the biparental 'HiFi' x 'SolFi' and included in the genetic map, allowing the identification of this homoelog on this oat biparental population (Coon, 2012). The markers oPt.0732 and oPt.4358, associated with β-glucan content in a diverse oat panel by Asoro *et al.* (2013), are also mapped at the 85.2 cM position on Mrg02. The oat marker avgbs_cluster_39442.1.33, mapped at the 62.9 cM position on Mrg06, was associated with β-glucan content in Eldorado do Sul in both years, as well as on the multi-environment approach (Table 1 and Figure 5). This marker is mapped 4 cM away from the Xcdo400 marker, which was associated with β-glucan content on the biparental populations 'Kanota' x 'Ogle' and 'Kanota' x 'Marion' (Kianian *et al.*, 2000).

The marker avgbs_cluster_33876.1.64, which is associated with the QTL identified on Mrg12, is located at the position 56.7 cM (Table 1 and Figure 5). This marker is close to the marker UMN136, mapped at the 63.5 cM position on Mrg12, that is near to the QBg1.jki-A1 QTL identified by Herrmann *et al.* (2014). The marker avgbs_cluster_32091.1.36, which is mapped at the 76.3 cM position on Mrg19, was associated with β-glucan content in Londrina 2017 and Eldorado do Sul 2018 (Table 1 and Figure 5). This marker is 2.2 cM away from the CDO1385F marker, which is a possible ortholog of the Xcdo1385H, associated with β-glucan content on the populations 'Kanota'

x ‘Ogle’ and ‘Kanota’ x ‘Marion’ (Kianian *et al.*, 2000). In this sense, these results illustrate that there are genomic regions acting on β-glucan content in common between northern versus southern oat germplasms, while other regions are specific to subtropical conditions or were not identified in spring oats under temperate conditions.

The comparative mapping showed genomic regions in common between oat and barley (Figure 6). The QTL-associated markers detected on oat consensus groups Mrg02, Mrg06, and Mrg11 showed nucleotide similarities with barley chromosome 7H, 5H, and 2H (Figure 6). These regions control barley β-glucan content and include many QTL and genes (Islamovic *et al.*, 2013; Houston *et al.*, 2014; Mohammadi *et al.*, 2014). This comparative mapping study reinforces the GWA findings, indicating that the genomic regions identified in oats are evolutionarily conserved in barley. In addition, the distances between oat QTL and barley genes, on barley chromosomes (Figure 6), may be reduced on the oat genome. In this sense, the release of an oat reference genome may identify which genes are close to the QTL identified in this study.

### 3.5 Conclusions

The UFRGS Oat Panel shows a weak population structure and is useful for future genome-wide association mappings in subtropical conditions. Seven QTL associated with β-glucan content were identified in oats. These QTL are located on Mrg02, Mrg06, Mrg11, Mrg12, Mrg19, and Mrg20. The QTL identified on Mrg02, Mrg06, and Mrg11 seem to be conserved in barley.

### 3.6 References

AACC - American Association of Cereal Chemists. Approved methods, 9 ed. Saint Paul: AACC, 1999.

ALTSCHUL, S. F. *et al*. Basic Local Alignment Search Tool. **Journal of Molecular Biology**, v. 215, p. 403-410, 1990.

ANDREWS, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available online at: http://www.bioinformatics.ba braham.ac.uk/projects/fastqc. Accessed on October 15th, 2018.

ASORO, F. G. *et al*. Genome-wide association study for beta-glucan concentration in elite North American oat. **Crop Science**, v. 53, p. 542-553, 2013.

AOAC - Association of Official Analytical Chemistry. Official methods of analysis of the Association of Official Analytical Chemistry, 16 ed. Washington: AOAC, 1997.

BRADBURY, P. J. *et al*. TASSEL: software for association mapping of complex traits in diverse samples. **Bioinformatics**, v. 23, p. 2633-2635, 2007.

BEKELE, W. A. *et al.* Haplotype-based genotyping-by-sequencing in oat genome research. **Plant Biotechnology Journal**, v. 16, p. 1452-1463, 2018.

BOLGER, A. M. *et al.* Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, p. 2114-2120, 2014.

CHO, K. C.; WHITE, P. J. Enzymatic analysis of beta-glucan content in different oat genotypes. **Cereal Chemistry**, v. 70, p. 539-542, 1993.

CHU, Y. *et al.* (Ed.). Oats nutrition and technology. John Wiley & Sons Incorporated, 2013.

COON, M. A. Characterization and variable expression of the *CslF6* homologs in oat (*Avena* sp.). **Thesis (Master)** - Department of Plant and Wildlife Sciences, Brigham Young University (BYU), Utah, f. 47, 2012.

DE KOEYER, D. L. *et al.* A molecular linkage map with associated QTLs from a hulless× covered spring oat population. **Theoretical and Applied Genetics**, v. 108, p. 1285-1298, 2004.

ENDELMAN, J. B. Ridge regression and other kernels for genomic selection with R package rrBLUP. **Plant Genome**, v. 4, p. 250-255, 2011.

ESVELT KLOS, K. *et al.* Population genomics related to adaptation in elite oat germplasm. **The Plant Genome**, v. 9, p. 1-12, 2016.

EXCOFFIER, L. *et al*. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. **Genetics**, v. 131, p. 479-491, 1992.

FINCHER, G. B.; STONE, B. A. Chemistry of nonstarch polysaccharides. In: WRIGLEY, C., CORKE, H., WALKER, C. E. (Eds.) Encyclopedia of Grain Science. Elsevier, Oxford, p. 206-223, 2004.

HERRERA, M. P. *et al.* β-Glucan content, viscosity, and solubility of Canadian grown oat as influenced by cultivar and growing location. **Canadian Journal of Plant Science**, v. 96, p. 183-196, 2016.

HERRMANN, M. H. *et al.* Quantitative trait loci for quality and agronomic traits in two advanced backcross populations in oat (*Avena sativa* L.). **Plant Breeding**, v. 133, p. 588-601, 2014.

HOUSTON, K. *et al.* 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**, v. 15, p. 907, 2014.

HUANG, Y. *et al.* Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. **Plos One**, v. 9, p. e102448, 2014.

ISLAMOVIC, E. *et al*. Genetic dissection of grain beta-glucan and amylose content in barley (*Hordeum vulgare* L.). **Molecular Breeding**, v. 31, p. 15-25, 2013.

KIANIAN, S. F. *et al.* Quantitative trait loci influencing β-glucan content in oat (Avena sativa, 2n= 6x= 42). **Theoretical and Applied Genetics**, v. 101, p. 1039-1048, 2000.

LU, F. *et al.* Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. **Plos Genetics** v. 9, p. e1003215, 2013.

MANGIN, B. *et al*. Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. **Heredity**, v. 108, p. 285-291, 2012.

MCCLEARY, B. V.; CODD, R. Measurement of (1→3),(1→4)-β-D-Glucan in Barley and Oats: A Streamlined Enzymic Procedure. **Journal of the Science of Food and Agriculture**, v. 55, p. 303-312, 1991.

MCNISH, I. G. *et al.* Mapping crown rust resistance at multiple time points in elite oat germplasm. **The Plant Genome**, 2019.

MILLER, S. S. *et al.* Oat β-glucans: an evaluation of eastern canadian cultivars and unregistered lines. **Canadian Journal of Plant Science**, v. 73, p. 429-436, 1993.

MOHAMMADI, M. *et al.* Association mapping of grain hardness, polyphenol oxidase, total phenolics, amylose content, and β-glucan in US barley breeding germplasm. **Molecular Breeding**, v. 34, p. 1229-1243, 2014.

NEWELL, M. A. *et al.* Genome-wide association study for oat (*Avena sativa* L.) beta-glucan concentration using germplasm of worldwide origin. **Theoretical and Applied Genetics**, v. 125, p. 1687-1696, 2012.

PARADIS, E. pegas: an R package for population genetics with an integrated-modular approach. **Bioinformatics**, v. 26, p. 419-420, 2010.

PETERSON, D. M. Genotype and environment effects on oat beta-glucan concentration. **Crop Science**, v. 31, p. 1517-1520, 1991.

POLAND, J. A. *et al*. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. **Plos One**, v. 7, p. e32253, 2012.

R DEVELOPMENT CORE TEAM. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

SAASTAMOINEN, M. *et al*. Genetic and environmental variation in β-glucan content of oats cultivated or tested in Finland. **Journal of Cereal Science**, v. 16, p. 279-290, 1992.

SAASTAMOINEN, M. Effects of environmental factors on the β-glucan content of two oat varieties. **Acta Agriculturae Scandinavica**, v. 45, p. 181-187, 1995.

TANHUANPÄÄ, P. *et al.* QTLs for important breeding characteristics in the doubled haploid oat progeny. **Genome**, v. 53, p. 482-493, 2010.

TANHUANPÄÄ, P. *et al*. An updated doubled haploid oat linkage map and QTL mapping of agronomic and grain quality traits from Canadian field trials. **Genome**, v. 55, p. 289-301, 2012.

TINKER, N. A. *et al.* Haplotag: Software for Haplotype-Based Genotyping-by-Sequencing Analysis. **Genes Genomes and Genetics**, v. 6, p. 857-863, 2016.

VOORRIPS, R. E. MapChart: Software for the graphical presentation of linkage maps and QTLs. **Journal of Heredity**, v. 93, p. 77-78, 2002.

WHITEHEAD, A. *et al.* Cholesterol-lowering effects of oat β-glucan: a meta-analysis of randomized controlled trials. **American Journal of Clinical Nutrition**, v. 100, p. 1413-1421, 2014.

YE, E. Q. *et al.* Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease and weight gain. **Journal of Nutrition**, v. 142, p. 1304-1313, 2012.

# 4 CHAPTER 2: GENOME-WIDE MAPPING FOR KERNEL SHAPE AND ITS ASSOCIATION WITH β-GLUCAN CONTENT IN OATS


## Abstract

Kernel shape and β-glucan content are decisive factors for oat millers. The identification of the genomic regions controlling these traits, as well as the phenotypic and genetic relationship between them, will be a key step to improve their genetic progress. The objectives of this study were to identify genomic regions associated with kernel shape and elucidate the relationship between kernel shape and β-glucan content in oats. An elite oat panel assembled by 407 varieties and breeding lines was evaluated in Londrina and Eldorado do Sul, Southern Brazil. Genome-wide association analyses were performed for each environment separately and also considering the multi-environment approach. In addition, phenotypic and genetic correlations coefficients were estimated between kernel shape-related traits and β-glucan content in oats. Genomic regions controlling kernel length were identified on Mrg06, Mrg21, and Mrg24. Kernel width and kernel thickness traits are genetically correlated and share a quantitative trait loci (QTL) in common, located on Mrg13 of the oat consensus map. Phenotypic and genetic correlations were found between kernel width and kernel thickness versus β-glucan content, indicating a negative effect of these traits on β-glucan content. The indirect selection for β-glucan content using kernel width seems to be effective. The selection of wider and thicker kernels may dilute the β-glucan content in oats. The results presented in this study are useful to understand the relationship between kernel shape and β-glucan content in oats.

### 4.1 Introduction

Oat (*Avena sativa* L.) kernels show unique chemical properties for a healthy life. The balanced amino acid composition of oat kernels, associated with high fiber content, turn this species a synonymous of healthy food. One of the most important dietary fibers found in oat kernels are β-glucans. β-glucans are soluble fibers located mainly in the external area of oat kernels, mostly in the aleurone layer, sub-aleurone layer, and in the starchy endosperm adjacent to the embryo. The richest region for β-glucan in oat kernels corresponds to the junction between aleurone layer and endosperm (Wood *et al*., 1991; Wood, 1993). The ingestion of β-glucan in proper quantities results in great benefits for human health, including improvement of the immune function and reduction of the glycemic index, cholesterol levels in the blood, obesity, and cardiovascular disease risk (Cho *et al.*, 2013; Whitehead *et al*., 2014; Ye *et al*., 2012).

Despite highly associated with many benefits for human health, selection for high β-glucan content is not performed in an effective way by many oat breeders. There are several factors limiting selection for β-glucan content, such as the quantitative nature of the trait, costs and complexity to obtain reliable phenotypes, as well as the influence exercised by environmental conditions on the trait. In addition, considering that oat breeding is still recent when compared to other small grains, the genetic progress for grain quality traits is lower when compared to the other ones. In subtropical environments, for example, the main selection criteria of oat breeders are related to adaptation, mainly focusing on flowering adequacy and crown rust resistance. Selections for kernel shape-related traits are performed considering only physical aspects, basically accounting for milling industry parameters. However, the relationship between kernel shape traits and quality traits, such as β-glucan content, was not well understood to date.

Genomic regions associated with grain shape were identified in oats and some of them are co-located with loci affecting β-glucan content (Groh *et al.*, 2001). Quantitative

trait loci (QTL) associated with kernel area, kernel length, and kernel width were identified in the biparental mapping populations ‘Kanota’ x ‘Ogle’ (KO) and ‘Kanota’ x ‘Marion’ (KM) using composite interval mapping, explaining from 9.2% to 60.7% of the phenotypic variation (Groh *et al.*, 2001). Six linkage groups carrying QTL for kernel shape identified by Groh *et al.* (2001) were also associated with β-glucan content. These results suggest that β-glucan content may be influenced by kernel shape. For example, in the linkage group KM 11–41, the alleles from ‘Kanota’ caused an increase in oil concentration, kernel area, and kernel length but decreased β-glucan concentration in mapping population KM (Groh *et al.*, 2001).

Loci controlling other traits were also associated with β-glucan content in oats. QTL affecting heading date were associated with β-glucan QTL in different genomic regions (Kianian *et al.*, 2000; Groh *et al.*, 2001; De Koeyer *et al.*, 2004). According to De Koeyer *et al.* (2004), the common location of heading date and β-glucan QTL could indicate a pleiotropic effect caused when later maturing lines fail to complete the optimum grain filling period, which results in shrunken kernels with less endosperm and enhanced relative β-glucan content. In this sense, heading date would affect directly kernel shape and indirectly β-glucan content. QTL associated with kernel size were also associated with other important traits in oats. Three QTL affecting kernel traits (area, length, and width) were detected in the mapping populations KO and KM (Groh *et al.*, 2001) in regions that correspond to grain yield QTL in the population ‘Terra’ x ‘Marion’ (De Koeyer *et al.*, 2004). The objectives of this study were to identify genomic regions associated with kernel shape in an oat panel adapted to subtropical conditions and clarify the relationship between kernel shape and β-glucan content in oats.

## 4.2 Materials and methods

### *4.2.1 Plant material*

An oat panel of 407 genotypes, including varieties and breeding lines, was assessed in this study. This panel, hereafter termed as UFRGS Oat Panel, was assembled prioritizing different pedigrees and shows a wide genetic variability. The UFRGS Oat Panel is originally composed by 413 genotypes, 384 hulled genotypes and 29 naked ones from different oat breeding programs. Nine of the 413 genotypes were not assessed and were not included in this study. A total of 358 genotypes of the assessed ones were developed by the Federal University of Rio Grande do Sul (UFRGS) Oat Breeding Program, from 1974 to 2015, 14 were developed by other Brazilian oat breeding programs, and 35 were developed by foreign oat breeding programs, mainly from North America. All lines are part of the UFRGS Oat Breeding Program germplasm.

*4.2.2 Field experiments*

The UFRGS Oat Panel was assessed in 2017 and 2018. In 2017, the experiment was carried out at the Agronomic Institute of Paraná (IAPAR) Experimental Station (23° 21' S and 51° 10' W), located in Londrina city, Paraná state, Southern Brazil. The experiment was sown in April 28$^{th}$ and conducted in a randomized complete block design, with two replications. Each experimental unit was composed by a single 1 m long row, spaced 0.25 m apart each other. Planting was carried out manually at a rate of 50 and 60 seeds per plot for hulled and naked genotypes, respectively. The base fertilization was 275 kg ha$^{-1}$ of a 5-20-20 N-P-K formula, and topdressing nitrogen, in the form of dry urea, was applied once, when plants showed four extended leaves, at a rate of 20 kg ha$^{-1}$ of N.

In 2018, the experiment was carried out at the UFRGS Agronomic Experimental Station (30° 07' S and 51° 41' W), located in Eldorado do Sul city, Rio Grande do Sul state, Southern Brazil. The experiment was sown in June 15$^{th}$ and conducted in a randomized complete block design, with two replications. Each experimental unit was composed by a two 2 m long row, spaced 0.2 m apart each other and 0.4 between plots. Planting was carried out mechanically at a rate of 60 and 80 seeds per linear m for hulled

and naked genotypes, respectively. The base fertilization was 300 kg $ha^{-1}$ of a 5-30-15 N-P-K formula, and topdressing nitrogen, in the form of dry urea, was applied twice, when plants showed four and six extended leaves, at a rate of 33 kg $ha^{-1}$ of N per application. Pests, fungal diseases, and weeds were chemically controlled in both experiments. At the field maturity stage, approximately 25 panicles per genotype were collected from a single plot for phenotypic analysis.

*4.2.3 Kernel shape assessment*

Kernel shape was assessed at the Rio Grandense Institute of Rice (IRGA), located in Cachoeirinha city, Rio Grande do Sul state, Southern Brazil, using a rice classer Image S (Selgron, Santa Catarina, Brazil). This instrument performs a three-dimensional screening of each grain. Three kernel-related traits were assessed using 100 dehulled kernels from a single replication: kernel length; kernel width; and kernel thickness, all in millimeters. Each kernel was assessed separately and a final report was created showing the number of assessed kernels per sample, the average of the assessed kernels for each trait, and the standard deviation. The phenotypic assessment for kernel shape-related traits is shown in Figure 1.


Grain sample *Image S* Individual screening

FIGURE 1. Phenotypic assessment for kernel shape-related traits in oats.

β-glucan content was predicted using a near-infrared spectroscopy (NIRS) model at the Center for Food Research of the University of Passo Fundo, Passo Fundo city, Rio Grande do Sul state, Southern Brazil. Initially, 15 g of dehulled kernels per genotype were ground in a Willey type mill (Marconi, Piracicaba, Brazil) with 0.5 mm sieve. After milled,

each whole flour sample was well homogenized using a set of sieves and stored at -20 ºC until β-glucan prediction. Samples were kept at room temperature (25 ºC) for 4 h before NIRS prediction to balance the moisture and temperature in order to avoid any influence of these factors on the reflectance and absorbance. The β-glucan prediction for each genotype was performed using whole flour, obtained from a single replication, in a NIRS™ DS2500 analyzer (Foss, Hillerod, Denmark).

*4.2.4 DNA isolation and genotyping by sequence*

DNA was isolated from dry-freeze lyophilized tissue collected from seedlings coming from a single panicle of each genotype at the University of Minnesota Genomics Center (UMGC). Genomic DNA was isolated using the DNeasy® Plant DNA isolation kit (Qiagen, Valencia, USA). After quantification, genotyping by sequencing (GBS) was performed at the UMGC, following a two-enzyme (PstI and MspI) protocol adapted from Poland *et al.* (2012). The libraries were sequenced using a NovaSeq™ 6000 sequencer (Illumina, San Diego, USA), producing individual 100-bp single-end reads. Sequence data was trimmed using the software Trimmomatic (Bolger *et al.*, 2014) and quality control of reads was verified using the FastQC software (Andrews, 2010).

*4.2.5 SNP calling and filtering*

Genotypes were called using a *de novo* pipeline without the requirement of a reference genome. The single-nucleotide polymorphisms (SNP) calling was performed using the Haplotag software (Tinker *et al.*, 2016). Initially, the Universal Network Enabled Analysis Kit (UNEAK) pipeline (Lu *et al.*, 2013), available on TASSEL 3 software (Bradbury *et al.*, 2007), was used for demultiplexing FastQ files into tag count files for each taxa. After, tag counts and a merged tag count files, all produced by UNEAK pipeline, were used in a discovery pipeline of the Haplotag software with default parameters (Tinker *et al.*, 2016). All SNP were aligned to the current oat consensus map (Bekele *et al.*, 2018) using a Basic Local Alignment Search Tool (BLAST; Altschul *et al.*,

1990). SNP markers were then filtered using the following criteria: i) biallelic; ii) missing data < 50%; iii) minor allele frequency (MAF) > 2%; and iv) heterozygosity < 10%. The Pearson's correlation coefficient (*r*) was calculated for all pairwise combinations of SNPs. Markers highly correlated ($r > 0.99$) were removed to avoid redundant information. After filtering steps, 10,182 SNPs were kept in the final matrix. Missing data points were imputed using the 'expectation-maximization' (EM) method, using the 'A.mat' function from the rrBLUP package (Endelman, 2011), available on R software (R Development Core Team, 2008).

### *4.2.6 Genome-wide association mapping*

Genome-wide association (GWA) mapping was conducted for kernel length, kernel width, and kernel thickness. Analyses were performed for each trial separately and also in a multi-environment model. GWA mapping was performed using the 'GWAS' function from the rrBLUP R package (Endelman, 2011). A mixed linear model (MLM) was employed accounting for population structure and kinship (Q + K model). Population structure was accounted for using the first three principal components, according to results presented previously in the Chapter 1. Kinship was accounted using a kinship matrix calculated using the 'A.mat' function from the rrBLUP R package (Endelman, 2011). Population structure, kinship, and trial (in the multi-environment model) were fit as fixed effect. Quantile-quantile (Q-Q) plots of the observed versus expected p-values under the null hypothesis were visualized for each analysis to verify signs of statistic inflation. Two statistical significance thresholds were established: i) a false discovery rate (FDR) corresponding to $\alpha = 0.05$, according to Benjamini and Hochberg (1995); and ii) an arbitrary p-value $\leq 0.001$. QTL were reported either if significant for the FDR threshold in one environment or significant in both environments for the arbitrary p-value threshold.

*4.2.7 Statistical analyses*

Boxplots were developed to illustrate the phenotypic variation observed for each trait in both assessed environments. Phenotypic and genetic correlations among traits were performed. Phenotypic correlations were estimated using Pearson's correlation coefficients using built-in R functions. Genetic correlations were estimated using a multi-trait Restricted Maximum Likelihood (REML) algorithm, considering the genomic relationship matrix with the NAM (Xavier *et al.*, 2015) R package. The genetic and residual covariance components were used to estimate the repeatability of each trait.

Once β-glucan content is a quantitative trait and highly influenced by the environmental conditions, an indirect selection approach using kernel width was considered, as follow:

$$\rho_{kw,bg} = \frac{CR}{R} = \frac{t_{kw}\, x\, r_{kw,bg}}{t_{bg}} \quad (1)$$

where CR is correlated response, R is direct response, $t_{kw}$ is kernel width repeatability, $r_{kw,bg}$ is genetic correlation between kernel width and β-glucan, and $t_{bg}$ is β-glucan repeatability. This approach was based on Xavier *et al.* (2017), who proposed indirect selection for grain yield in soybean using the average canopy coverage.

## 4.3 Results

*4.3.1 Phenotypic variation for kernel shape*

A detailed kernel shape characterization was performed for the UFRGS Oat Panel. All evaluated traits showed wide phenotypic variation in both locations, with emphasis for the kernel length trait (Figure 2). In Londrina, the kernel length trait varied from 4.99 to 9.10 mm, with overall average of 6.75 mm (Figure 2). The kernel width and kernel thickness traits showed lower phenotypic amplitude than kernel length in Londrina (Figure 2). The kernel width trait ranged from 1.87 to 2.84 mm, with overall average of 2.37 mm, while

the kernel thickness trait varied from 1.43 to 2.19 mm, with overall average of 1.83 mm (Figure 2).

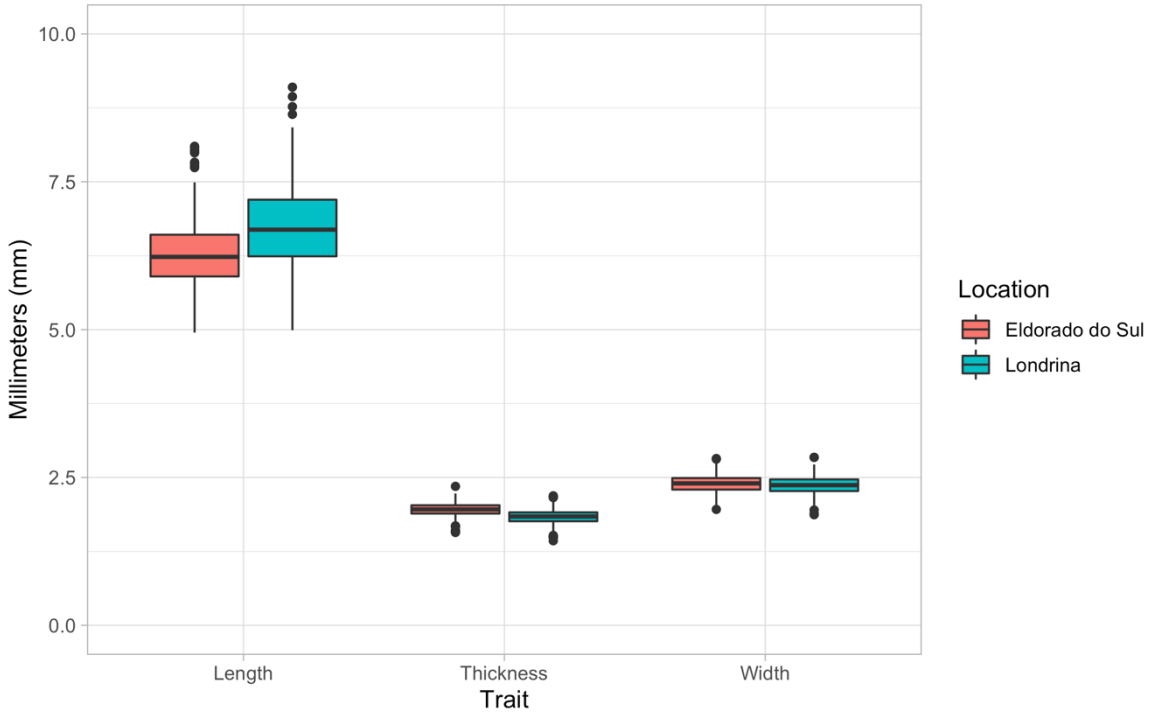



FIGURE 2. Phenotypic variation for kernel shape-related traits in two locations.

In Eldorado do Sul, the observed phenotypic variation for kernel length varied from 4.95 to 8.10 mm, with overall average of 6.27 mm (Figure 2). For the kernel width trait, the phenotypic variation ranged from 1.95 to 2.82 mm, with average of 2.39 mm (Figure 2). For the kernel thickness trait, values ranging from 1.43 to 2.19 were observed, with overall average of 1.83 mm. The kernel length amplitude was higher in Londrina than in Eldorado do Sul (Figure 2). Considering all assessed genotypes, kernels were on average 0.5 mm larger in Londrina (Figure 2). On the other hand, the kernel width and kernel thickness traits showed similar results in both locations (Figure 2).

*4.3.2 Association between β-glucan content and kernel shape*

Phenotypic and genetic correlations were estimated to identify the relationship between β-glucan content and kernel shape in oats. In general, similar results were observed between environments (Table 1). The phenotypic correlations in Londrina

indicated weak negative correlations between β-glucan content versus kernel width and kernel thickness, both estimated as -0.19 (Table 2). The same results were observed for phenotypic correlations in Eldorado do Sul, where β-glucan content showed negative correlations of -0.33 and -0.23 with kernel width and kernel thickness, respectively (Table 1). A strong phenotypic correlation was observed between kernel width and kernel thickness in both environments (Table 1).

TABLE 1. Phenotypic (upper-right diagonal) and genetic (lower-left) correlations between β-glucan and kernel shape-related traits in two locations.

| Trait | Kernel length | Kernel width | Kernel thickness | β-glucan† |
|---|---|---|---|---|
| | | Londrina, PR | | |
| Kernel length | 1.00 | 0.04 | 0.10 | 0.05 |
| Kernel width | -0.15 | 1.00 | 0.91 | -0.19 |
| Kernel thickness | -0.08 | 0.95 | 1.00 | -0.19 |
| β-glucan† | -0.07 | -0.44 | -0.37 | 1.00 |
| | | Eldorado do Sul, RS | | |
| Kernel length | 1.00 | 0.07 | 0.16 | 0.14 |
| Kernel width | 0.04 | 1.00 | 0.86 | -0.33 |
| Kernel thickness | 0.17 | 0.91 | 1.00 | -0.23 |
| β-glucan‡ | 0.23 | -0.42 | -0.25 | 1.00 |

†Quantified using enzymatic method; ‡Quantified using NIRS prediction.

The genetic correlation analysis also revealed negative correlations between β-glucan content versus kernel width and kernel thickness. The correlation between β-glucan content and kernel width was consistent in both environments, with coefficients of -0.44 and -0.42 for Londrina and Eldorado do Sul, respectively (Table 1). For the genetic correlation between β-glucan content and kernel thickness, a smooth change was observed between environments. In Londrina, a correlation coefficient of -0.37 was calculated, while -0.25 was observed in Eldorado do Sul (Table 1). As mentioned for phenotypic correlation, a strong genetic correlation was observed between kernel width and kernel thickness in both environments (Table 1).

Environmental variance, environment coefficient of variation, genetic variance, genetic coefficient of variation, and repeatability for all evaluated traits are shown in Table 2. In Londrina, the β-glucan content trait showed highest environmental variance, followed

by kernel length, while low variances were observed for kernel width and kernel thickness (Table 2). The β-glucan content was the only trait that showed higher environmental variance than genetic one in Londrina (Table 2). Once β-glucan content and kernel shape-related traits are assessed using different measure units, the comparison of genetic variances among these traits is not the best option. In this sense, considering the genetic coefficient of variation, the genetic variation for β-glucan content (6.78%) was slightly higher than for kernel length (6.29%) considering the enzymatic method. On the other hand, the genetic variation for β-glucan content (8.78%) was considerable higher when compared to kernel length (5.68%) considering the NIRS prediction (Table 2). For the environmental variances, kernel width and kernel thickness traits showed low genetic variation, as observed via genetic coefficient of determination (Table 2). The β-glucan content trait showed low repeatability, while all other traits presented moderate-high ones (Table 2).

TABLE 2. Environmental variance, genetic variance, and repeatability of β-glucan content, days from emergence to flowering, and kernel shape-related traits in two locations.

| Trait | Londrina† - PR | | | | | Eldorado do Sul‡ - RS | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Var(E) | CV(E) | Var(G) | CV(G) | *t* | Var(E) | CV(E) | Var(G) | CV(G) | *t* |
| KLen | 0.165 | 6.020 | 0.180 | 6.290 | 0.521 | 0.070 | 4.220 | 0.127 | 5.680 | 0.647 |
| KWid | 0.006 | 3.270 | 0.011 | 4.430 | 0.656 | 0.006 | 3.240 | 0.010 | 4.180 | 0.615 |
| KThi | 0.003 | 2.999 | 0.008 | 4.890 | 0.716 | 0.003 | 2.990 | 0.006 | 4.230 | 0.666 |
| β-glu | 0.252 | 11.670 | 0.085 | 6.780 | 0.252 | 0.076 | 6.560 | 0.136 | 8.780 | 0.642 |

KLen, kernel length; KWid, kernel width; KThi, kernel thickness; β-glu, β-glucan content; Var(E), environmental variance; CV(E), environment coefficient of variation; Var(G), genetic variance; CV(G), genetic coefficient of variation; *t*, repeatability; †β-glucan content quantified using enzymatic method; ‡β-glucan content quantified using NIRS prediction.

In Eldorado do Sul, genetic variances were higher than environmental ones for all traits (Table 2). β-glucan content and kernel length showed similar values, as well as kernel width and kernel thickness (Table 2). The β-glucan content repeatability coefficient was higher when compared to Londrina (Table 2). A slight change was also observed for the kernel length repeatability. On the other hand, kernel width and kernel thickness traits showed similar repeatability coefficients when compared to Londrina (Table 2).

*4.3.3 Indirect selection for β-glucan content using kernel width*

An indirect selection approach is proposed for β-glucan content due to the genetic complexity of this trait. In this sense, the main goal of this approach is to accelerate the genetic gain of β-glucan content using a correlated trait that is easily selected. In this study, kernel width was used because showed highest genetic correlation with β-glucan content (Table 1). Using the equation 1, indirect selection responses of 1.14 and 0.40 were calculated for Londrina and Eldorado do Sul, respectively. This result shows that indirect selection for β-glucan content using kernel width in Londrina is 14% more effective than using β-glucan content per se. However, the same response was not observed in Eldorado do Sul due to the inflation of the repeatability value caused by the NIRS method (Table 2).

*4.3.4 Genome-wide association mapping for kernel shape-related traits*

Genomic regions associated with oat kernel shape were identified using GWA mapping. The kernel length, kernel width, and kernel thickness traits were studied in two locations. The results are presented for each location separately and then considering the multi-environment approach. Q-Q plots of the observed versus expected p-values under the null hypothesis were visualized for each analysis to verify signs of statistic inflation and are shown in Supplementary Figure 1. The Q-Q plots revealed no inflation in the model for all assessed traits in both environments, as well as to the multi-environment approach (Supplementary Figure 1). Some deflation was observed for the kernel width in all analyses (Supplementary Figure1b, Supplementary Figure1e, and Supplementary Figure1h). Two factors could explain this behavior: i) low heritability coefficients for the trait; and ii) overcorrection of the model.

*4.3.4.1 Genome-wide association mapping for kernel length*

Two QTL associated with kernel length were detected in Londrina. These QTL are located on the consensus linkage groups Mrg21 and Mrg24 (Figure 3a). Although it has

not reached the statistical threshold, a third QTL is suggested on Mrg06 (Figure 3a). The QTL located on Mrg21 is associated with the marker avgbs_200007.1.34, mapped at the position 211.2 cM (Table 3). The QTL detected on Mrg24 is associated with the markers avgbs2_80887.1.63, avgbs_122997.1.31, and avgbs_244007.1.36, which are mapped at the position 41.8 cM (Table 3).

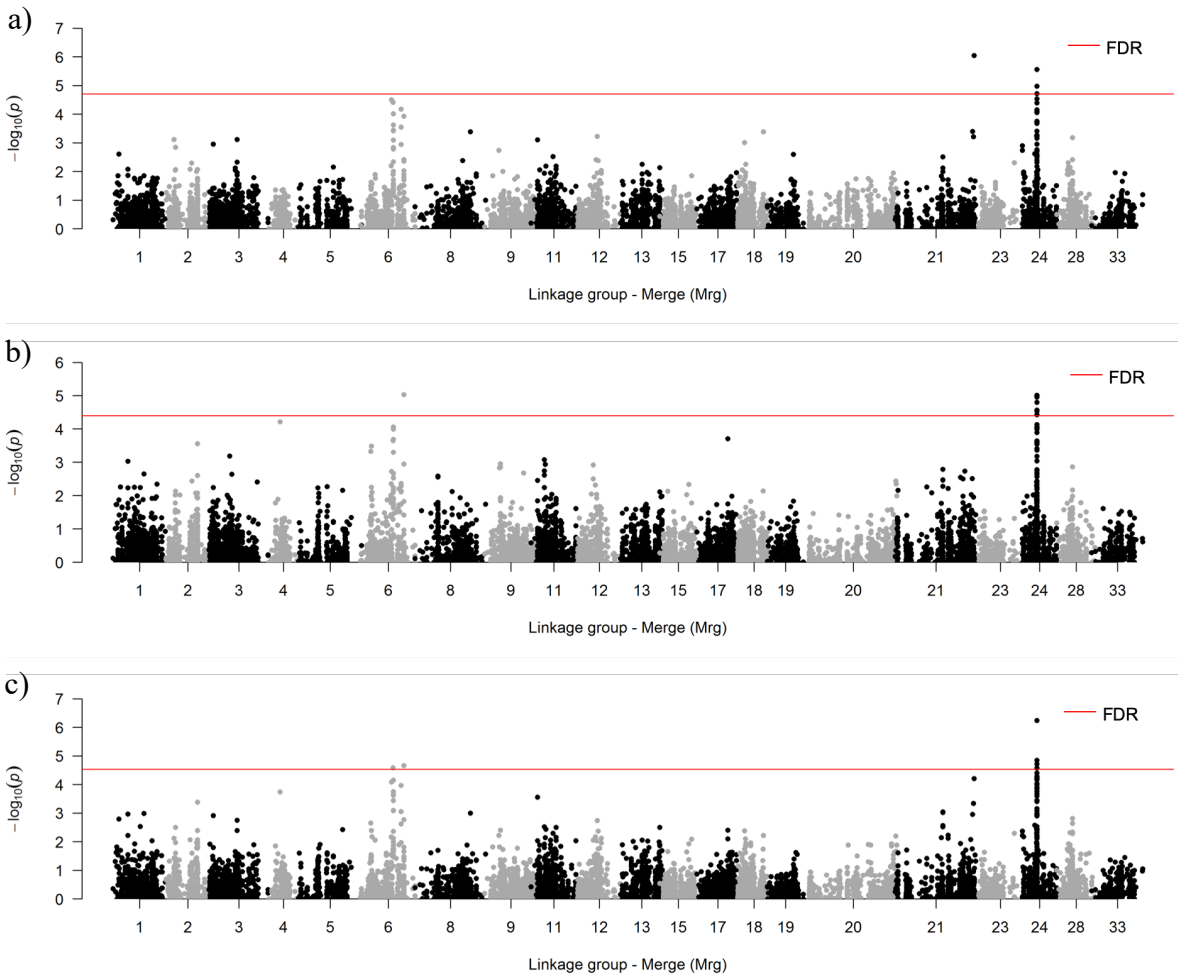



FIGURE 3. Manhattan plots for oat markers associated with kernel length. FDR, false discovery rate at 5%; a, Londrina; b, Eldorado do Sul; c, Multi-environment.

In Eldorado do Sul, two QTL associated with kernel length were also detected. These loci were located on Mrg06 and Mrg24 (Figure 3b), corresponding to the same genomic regions detected in Londrina (Figure 3a). The QTL located on Mrg06 is associated with the marker avgbs_cluster_11495.1.10, which is mapped at the position 116 cM. The QTL located on Mrg24 is associated with the markers avgbs_48068.1.53, avgbs2_80887.1.63, avgbs2_16989.1.15, avgbs_206876.1.18, avgbs_cluster_38861.1.50,

avgbs_cluster_38948.1.25, and avgbs_223170.1.42, which mapped at the position 41.8 cM (Table 3).

TABLE 3. Genomic regions controlling kernel length in elite oat germplasm adapted to subtropical conditions.

| Marker | Merge | Position† (cM) | -log p-value | Location |
|---|---|---|---|---|
| avgbs_200007.1.34 | Mrg21 | 211.2 | 6.0 | Londrina |
| avgbs2_80887.1.63 | Mrg24 | 41.8 | 5.6 | Londrina |
| avgbs_122997.1.31 | Mrg24 | 41.8 | 5.0 | Londrina |
| avgbs_244007.1.36 | Mrg24 | 41.8 | 4.7 | Londrina |
| avgbs_cluster_11495.1.10 | Mrg06 | 116.0 | 5.0 | Eldorado do Sul |
| avgbs_48068.1.53 | Mrg24 | 41.8 | 5.0 | Eldorado do Sul |
| avgbs2_80887.1.63 | Mrg24 | 41.8 | 5.0 | Eldorado do Sul |
| avgbs2_16989.1.15 | Mrg24 | 41.8 | 4.6 | Eldorado do Sul |
| avgbs_206876.1.18 | Mrg24 | 41.8 | 4.8 | Eldorado do Sul |
| avgbs_cluster_38861.1.50 | Mrg24 | 41.8 | 5.0 | Eldorado do Sul |
| avgbs_cluster_38948.1.25 | Mrg24 | 41.8 | 4.5 | Eldorado do Sul |
| avgbs_223170.1.42 | Mrg24 | 41.8 | 4.4 | Eldorado do Sul |
| avgbs2_71574.1.64 | Mrg06 | 86.5 | 4.6 | Multi-e |
| avgbs_cluster_11495.1.10 | Mrg06 | 116.0 | 4.7 | Multi-e |
| avgbs2_80887.1.63 | Mrg24 | 41.8 | 6.2 | Multi-e |
| avgbs_122997.1.31 | Mrg24 | 41.8 | 4.8 | Multi-e |
| avgbs_206876.1.18 | Mrg24 | 41.8 | 4.7 | Multi-e |
| avgbs_cluster_11191.1.9 | Mrg24 | 41.8 | 4.6 | Multi-e |

†Based on Bekele *et al.* (2018). QTL were detected considering a false discovery rate at 5% level. Multi-e, multi-environment mapping.

The multiple-environment GWA mapping revealed genomic regions controlling the kernel length trait in both locations. Three QTL associated with the trait were detected in the multi-environment approach (Figure 3c). Two of these QTL were located on Mrg06, while the third one was located on Mrg24. One of the two QTL located on Mrg06 was suggested and detected in Londrina and Eldorado do Sul, respectively, associated with the marker avgbs_cluster_11495.1.10 at position 116 cM (Table 3). The second QTL located on Mrg06 is associated with the marker avgbs2_71574.1.64 at the position 85.6 cM, 29.5 cM apart from the previous one (Table 3). The QTL detected on Mrg24 corresponds to the same region detected in both locations at the position 41.8 cM. This QTL is associated with the markers avgbs2_80887.1.63, which is detected on both locations, avgbs_122997.1.31, avgbs_206876.1.18, and avgbs_cluster_11191.1.9 (Table 3).

*4.3.4.2 Genome-wide association mapping for kernel width*

No QTL were detected for kernel width considering the FDR statistical threshold. In this sense, the arbitrary threshold (p-value ≤ 0.001) was used to identify genomic regions associated with the trait (Figure 4). In Londrina, markers located on the consensus linkage groups Mrg01, Mrg02, Mrg05, Mrg11, Mrg13, Mrg20, and Mrg23 reached the statistical threshold (Figure 4a). In Eldorado do Sul, significant markers were detected on Mrg05, Mrg13, and Mrg19 (Figure 4b). The locus located on Mrg13 has multiple markers above the arbitrary threshold, while other regions are represented by only one marker (Figure 4b).

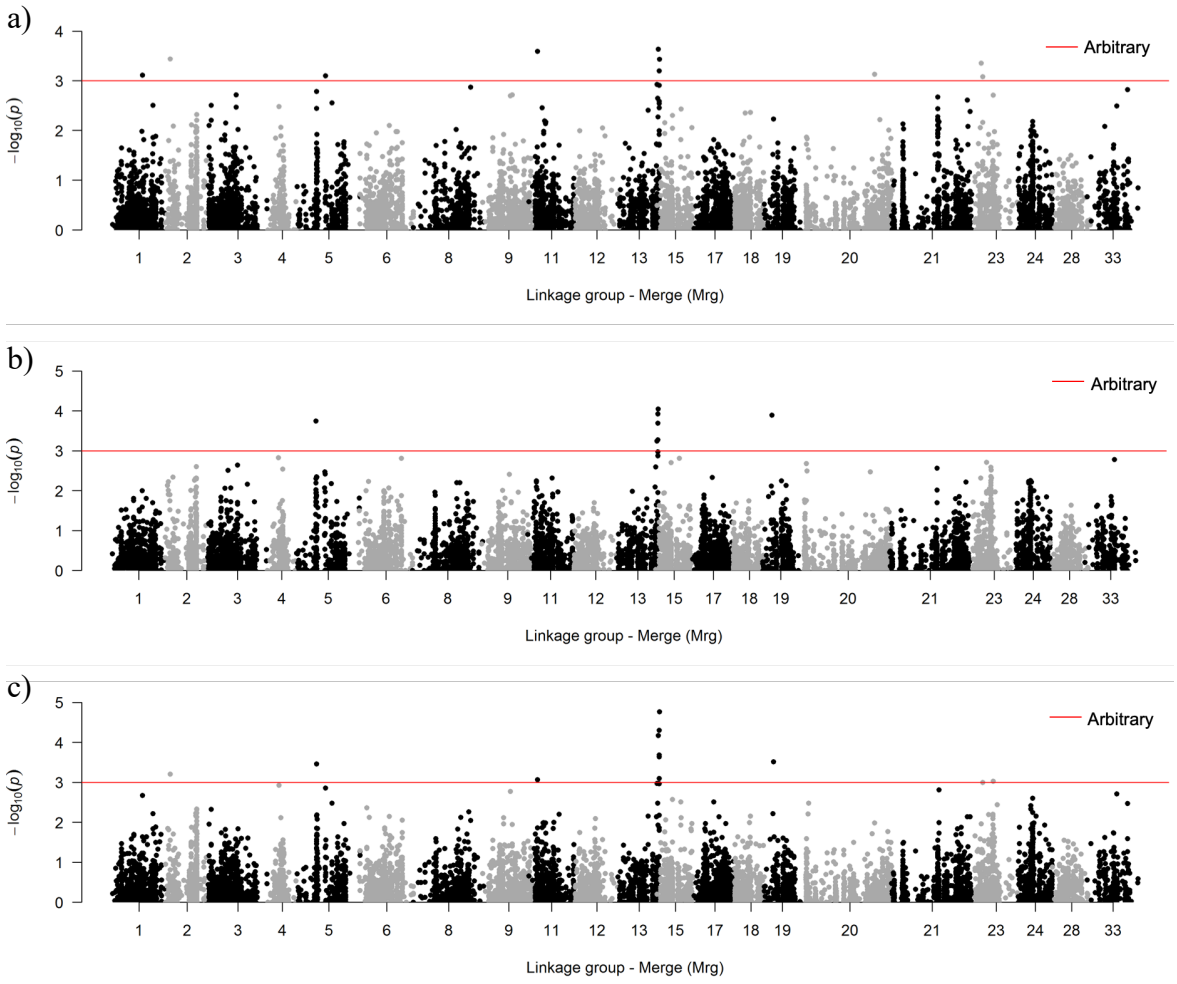



FIGURE 4. Manhattan plots for oat markers associated with kernel width. Arbitrary threshold, p-value ≤ 0.001; a, Londrina; b, Eldorado do Sul; c, Multi-environment.

A suggestive QTL associated with kernel width was detected on Mrg13 in the multi-environment mapping (Figure 4c). In addition, markers located on linkage groups

Mrg02, Mrg05, Mrg11, Mrg19, and Mrg23 also reached the arbitrary statistical threshold (Figure 4c). As mentioned in the Materials and Methods section, QTL were reported either if significant for the FDR threshold or significant in both environments for the arbitrary p-value threshold. In this sense, only the QTL located on Mrg13 was considered significant for kernel width. The number of markers associated with this QTL varies according to the location/analysis (Table 4).

TABLE 4. Genomic regions controlling kernel width in elite oat germplasm adapted to subtropical conditions.

| Marker | Merge | Position† (cM) | -log p-value | Location |
|---|---|---|---|---|
| avgbs2_38402.1.29 | Mrg13 | 111.7 | 3.6 | Londrina |
| avgbs_235979.1.30 | Mrg13 | 114.6 | 3.2 | Londrina |
| avgbs_cluster_22640.1.12 | Mrg13 | 115.4 | 3.4 | Londrina |
| avgbs2_38402.1.29 | Mrg13 | 111.7 | 3.2 | Eldorado do Sul |
| avgbs_12091.1.49 | Mrg13 | 114.6 | 3.3 | Eldorado do Sul |
| avgbs_21100.1.51 | Mrg13 | 114.6 | 3.7 | Eldorado do Sul |
| avgbs_235979.1.30 | Mrg13 | 114.6 | 3.9 | Eldorado do Sul |
| avgbs_cluster_22640.1.12 | Mrg13 | 115.4 | 4.0 | Eldorado do Sul |
| avgbs2_38402.1.29 | Mrg13 | 111.7 | 4.2 | Multi-e |
| avgbs_235979.1.30 | Mrg13 | 114.6 | 4.3 | Multi-e |
| avgbs_12091.1.49 | Mrg13 | 114.6 | 3.7 | Multi-e |
| avgbs_21100.1.51 | Mrg13 | 114.6 | 3.6 | Multi-e |
| avgbs_cluster_22640.1.12 | Mrg13 | 115.4 | 4.8 | Multi-e |

†Based on Bekele *et al.* (2018). QTL were detected considering an arbitrary threshold (p-value ≤ 0.001). Multi-e, multi-environment mapping.

In Londrina, the QTL is associated with the markers avgbs2_38402.1.29, avgbs_235979.1.30, and avgbs_cluster_22640.1.12, mapped at the positions 111.7 cM, 114.6 cM, and 115.4 cM, respectively, covering a genetic region of 3.7 cM (Table 4). These markers were also significant in Eldorado do Sul and in the multi-environment mapping (Table 4). The markers avgbs_12091.1.49 and avgbs_21100.1.51, which are mapped on Mrg13 at the position 114.6 cM, were also associated with the kernel width-related QTL in Eldorado do Sul and in the multi-environment mapping (Table 4).

*4.3.4.3 Genome-wide association mapping for kernel thickness*

For the kernel thickness trait, a suggestive QTL was detected on Mrg13 in all GWA mappings (Figure 5). The QTL is associated with the markers avgbs2_38402.1.29, avgbs_12091.1.49, avgbs_235979.1.30, avgbs_56322.1.38, avgbs_21100.1.51, avgbs_94638.1.46, and avgbs_cluster_22640.1.12, which are covering a genetic region of 3.7 cM, in both individual and multiple-environment approaches (Table 5). This QTL corresponds to the same region associated with kernel width (Figure 4).

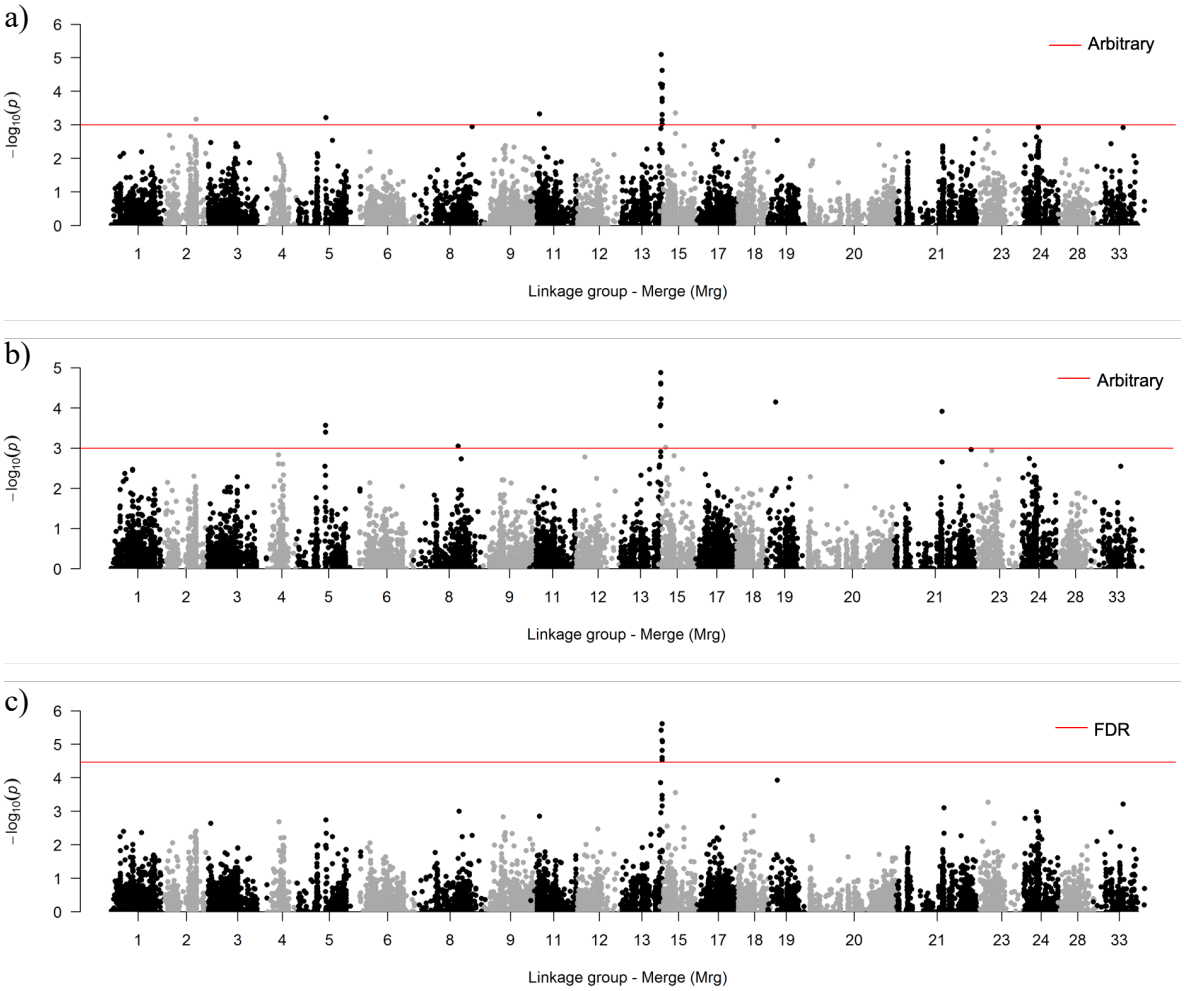



FIGURE 5. Manhattan plots for oat markers associated with kernel thickness. FDR, false discovery rate at 5%. Arbitrary, p-value ≤ 0.001; a, Londrina; b, Eldorado do Sul; c, Multi-environment.

In Londrina, markers located on Mrg02, Mrg05, Mrg11, and Mrg15 also trespassed the statistical threshold (Figure 5a). In Eldorado do Sul, besides the QTL on Mrg13, markers located on Mrg05, Mrg08, Mrg19, and Mrg21 also reached the arbitrary statistical

threshold (Figure 5b). For the multi-environment mapping, the QTL detected on Mrg13 was significant for kernel thickness considering the FDR statistical threshold (Figure 5c).

TABLE 5. Genomic regions controlling kernel thickness in elite oat germplasm adapted to subtropical conditions.

| Marker | Merge | Position† (cM) | -log p-value | Threshold | Location |
|---|---|---|---|---|---|
| avgbs2_38402.1.29 | Mrg13 | 111.7 | 5.1 | Arbitrary | Londrina |
| avgbs_12091.1.49 | Mrg13 | 114.6 | 4.6 | Arbitrary | Londrina |
| avgbs_235979.1.30 | Mrg13 | 114.6 | 4.6 | Arbitrary | Londrina |
| avgbs_56322.1.38 | Mrg13 | 114.6 | 4.1 | Arbitrary | Londrina |
| avgbs_21100.1.51 | Mrg13 | 114.6 | 3.8 | Arbitrary | Londrina |
| avgbs_94638.1.46 | Mrg13 | 114.6 | 3.1 | Arbitrary | Londrina |
| avgbs_cluster_22640.1.12 | Mrg13 | 115.4 | 4.2 | Arbitrary | Londrina |
| avgbs2_38402.1.29 | Mrg13 | 111.7 | 4.0 | Arbitrary | Eldorado do Sul |
| avgbs_56322.1.38 | Mrg13 | 114.6 | 3.6 | Arbitrary | Eldorado do Sul |
| avgbs_12091.1.49 | Mrg13 | 114.6 | 4.1 | Arbitrary | Eldorado do Sul |
| avgbs_94638.1.46 | Mrg13 | 114.6 | 4.6 | Arbitrary | Eldorado do Sul |
| avgbs_21100.1.51 | Mrg13 | 114.6 | 4.6 | Arbitrary | Eldorado do Sul |
| avgbs_235979.1.30 | Mrg13 | 114.6 | 4.9 | Arbitrary | Eldorado do Sul |
| avgbs_cluster_22640.1.12 | Mrg13 | 115.4 | 4.2 | Arbitrary | Eldorado do Sul |
| avgbs2_38402.1.29 | Mrg13 | 111.7 | 5.4 | FDR | Multi-e |
| avgbs_235979.1.30 | Mrg13 | 114.6 | 5.6 | FDR | Multi-e |
| avgbs_12091.1.49 | Mrg13 | 114.6 | 5.1 | FDR | Multi-e |
| avgbs_21100.1.51 | Mrg13 | 114.6 | 4.8 | FDR | Multi-e |
| avgbs_94638.1.46 | Mrg13 | 114.6 | 4.6 | FDR | Multi-e |
| avgbs_56322.1.38 | Mrg13 | 114.6 | 4.5 | FDR | Multi-e |
| avgbs_cluster_22640.1.12 | Mrg13 | 115.4 | 5.1 | FDR | Multi-e |

†Based on Bekele *et al.* (2018). QTL were detected considering an arbitrary threshold (p-value ≤ 0.001). Multi-e, multi-environment mapping.

### 4.4 Discussion

A wide phenotypic variation was observed for kernel-shape related traits assessed in this study (Figure 2). The kernel length trait showed higher phenotypic variation than other ones, with more amplitude and higher average in Londrina than in Eldorado do Sul (Figure 2). Kernel width and kernel thickness traits showed similar results in both locations. These results suggest that kernel width and kernel thickness are less influenced by the environmental conditions than kernel length. The environmental condition in

Londrina was favorable during flower formation, allowing the development of long kernels.

Kernel width and kernel thickness are phenotypically and genetically correlated with β-glucan content. Both traits showed negative correlations with β-glucan content, regardless of the environment (Table 1). The β-glucan content repeatability in Londrina, using the enzymatic approach, was 0.25. This low repeatability coefficient is explained by the genetic inheritance of the β-glucan content, which is a quantitative trait controlled by many loci. The use of NIRS predictions, which is represented by the Eldorado do Sul dataset, produced an underestimation of phenotypic variance and an overestimation of genetic variance for β-glucan content (Table 2). As consequence, the repeatability coefficient observed in Eldorado do Sul was more than twice as observed in Londrina, which used the enzymatic approach (Table 2). This inflation on β-glucan repeatability using NIRS values may justify why some studies using this approach detected many candidate regions associated with the trait.

The selection of genotypes with high β-glucan content is a major challenge for oat breeders due to phenotyping limitations. In this sense, this study proposes a new approach using indirect selection through kernel width, a genetically correlated trait with β-glucan content (Table 1). Considering the correlated response between traits, indirect selection for β-glucan content using kernel width is 14% more effective than using β-glucan content per se (Equation 1). On the other hand, this effectiveness was not observed for Eldorado do Sul because the β-glucan content repeatability was inflated by the NIRS method. Kernel shape-related traits were suggested to be used in the indirect selection for protein and lipid content traits in oats (Silva *et al*., 2008). Other studies also suggested indirect selection for oat kernel quality traits using kernel shape, such as Peterson and Wood (1997), who studied the composition and structure of oat kernels and observed higher oil contents in longer and thinner kernels.

The identification of genomic regions controlling kernel shape may represent a key step to understand genetic relationship with β-glucan content. Major quantitative trait loci (QTL) were associated with kernel shape-related traits in this study. Four different QTL were detected for kernel length, mapped on consensus linkage groups Mrg06, Mrg21, and Mrg24 (Figure 3 and Table 3). Mapping studies involving kernel shape-related traits are scarce in oats. QTL associated with kernel length, kernel area, and kernel width were identified on the biparental mapping populations 'Kanota' x 'Ogle' (KO) and 'Kanota' x 'Marion' (KM) using composite interval mapping (Groh *et al.*, 2001). Four QTL affecting kernel length were detected on linkage groups KO 3–38, KO 7–10–28, KO 17, and KO 27, explaining 44.1% of the phenotypic variation of the trait (Groh *et al.*, 2001). The linkage groups KO 3-38, KO 7-10-28, and KO 17 correspond to Mrg28, Mrg09, and Mrg02, respectively, while KO 27 is not anchored on the consensus map (Bekele *et al.,* 2018). A study performed at the University of Minnesota detected two QTL associated with oat kernel length on Mrg02 and Mrg28 (Ian G. McNish, unpublished results). Four QTL were associated with kernel length on the KM population, which were mapped on linkage groups KM 3, KM 7–28, KM 11–41, and KM x7, explaining 33.3% of phenotypic variation (Groh *et al.*, 2001). The linkage group 7-28 corresponds to Mrg33 on the consensus oat map (Bekele *et al.*, 2018), while all others are not anchored. In this sense, the studies performed by Groh *et al.* (2001) and McNish (unpublished results), using North America oat germplasm, share some findings. On the other hand, results of the present study are new candidate regions associated with kernel length in subtropical-adapted oat germplasm.

A QTL detected on Mrg13 was associated with kernel width (Figure 4 and Table 4) and kernel thickness (Figure 5 and Table 5) in oats. This result support the high phenotypic and genetic correlations observed between traits, confirming that both traits share a genetic region in common (Figure 4 and Figure 5). Four QTL controlling kernel width were found

on linkage groups KO 5, KO 9, KO 16, and KO 27, accounting for 30.5% of the observed phenotypic variation (Groh *et al.*, 2001). KO 5, KO 9, and KO 16 correspond to Mrg01, Mrg13, and Mrg24 on the oat consensus map (Bekele *et al.*, 2018). Only one QTL was detected for kernel width on KM population, located on KM 16 (Groh *et al.*, 2001)., corresponding to Mrg24 on the oat consensus map (Bekele *et al.*, 2018).

The QTL associated with kernel width identified by Groh *et al.* (2001) on KO 9 is associated with the marker Xcdo456, which is mapped at the position 114.6 cM on Mrg13, sharing the same position with QTL detected in the present study (Table 4). There are no other study involving characterization for kernel thickness in oats. In this context, the results of the present study are strong enough to suggest that the same region controlling kernel width in oats is also associated with kernel thickness.

Studies involving the identification of loci associated with kernel shape have been published in small grains (Ayoub *et al*., 2002; Sun *et al*., 2009). In oats, only Groh *et al.* (2001) identified candidate regions associated with kernel shape, but studying two biparental populations that are not adapted to subtropical conditions. In the present study, a diverse oat panel was studied in two contrasting environments in Southern Brazil. The results presented previously represent an important step to understand the genomic regions involved with oat kernel shape, allowing their future use for marker assisted selection (MAS) approaches. In addition, a consistent indirect selection purpose seems to improve the β-glucan content in oats using the correlated response with kernel width. Future studies involving β-glucan content and kernel shape-related traits may identify genomic regions in common between them and accelerate the genetic progress of these traits simultaneously using molecular tools.

### 4.5 Conclusions

Different genomic regions were identified for kernel shape-related traits in oats. Genomic regions controlling kernel length were identified on Mrg06, Mrg21, and Mrg24.

Kernel width and kernel thickness traits are genetically correlated and share a genomic region in common, located on Mrg13 of the oat consensus map. Negative correlations were found between kernel width and kernel thickness versus β-glucan content, suggesting that the selection of wider and thicker kernels dilute the β-glucan content in oats.

### 4.6 References

ALTSCHUL, S. F. *et al.* Basic Local Alignment Search Tool. **Journal of Molecular Biology**, v. 215, p. 403-410, 1990.

ANDREWS, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available online at: http://www.bioinformatics.ba braham.ac.uk/projects/fastqc. Accessed on October 15th, 2018.

AYOUB, M. *et al.* QTLs affecting kernel size and shape in a two-rowed by six-rowed barley cross. **Theoretical and Applied Genetics**, v. 105, p. 237-247, 2002.

BEKELE, W. A. *et al*. Haplotype-based genotyping-by-sequencing in oat genome research. **Plant Biotechnology Journal**, v. 16, p. 1452-1463, 2018.

BENJAMINI, Y., HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society**, v. 57, p. 289-300, 1995.

BRADBURY, P. J. *et al*. TASSEL: software for association mapping of complex traits in diverse samples. **Bioinformatics**, v. 23, p. 2633-2635, 2007.

BOLGER, A. M. *et al*. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, p. 2114-2120, 2014.

CHO, S. S. *et al.* Consumption of cereal fiber, mixtures of whole grains and bran, and whole grains and risk reduction in type 2 diabetes, obesity, and cardiovascular disease. **American Journal of Clinical Nutrition**, v. 98, p. 594-619, 2013.

DE KOEYER, D. L. *et al.* A molecular linkage map with associated QTLs from a hulless x covered spring oat population. **Theoretical and Applied Genetics**, v. 108, p. 1285-1298, 2004.

ENDELMAN, J. B. Ridge regression and other kernels for genomic selection with R package rrBLUP. **The Plant Genome**, v. 4, p. 250-255, 2011.

GROH, S. *et al*. Analysis of factors influencing milling yield and their association to other traits by QTL analysis in two hexaploid oat populations. **Theoretical and Applied Genetics**, v. 103, p. 9-18, 2001.

KIANIAN, S. F. *et al.* Quantitative trait loci influencing β-glucan content in oat (*Avena sativa*, 2n=6x=42). **Theoretical and Applied Genetics**, v. 101, p. 1039-1048, 2000.

LU, F. *et al.* Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. **Plos Genetics**, v. 9, p. e1003215, 2013.

PETERSON, D. M.; WOOD, D. F. Composition and structure of high-oil oat. **Journal of Cereal Science**, v. 26, p. 121-128, 1997.

POLAND, J. A. *et al.* Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. **Plos One**, v. 7, p. e32253, 2012.

R DEVELOPMENT CORE TEAM. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

SILVA, C. F. L. *et al.* Near infrared reflectance spectroscopy (NIRS) to assess protein and lipid contents in *Avena sativa* L. **Crop Breeding and Applied Biotechnology**, v. 8, p. 127-133, 2008.

SUN, X. *et al.* QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. **Euphytica**, v. 165, p. 615, 2009.

WHITEHEAD, A. *et al.* Cholesterol-lowering effects of oat β-glucan: a meta-analysis of randomized controlled trials. **American Journal of Clinical Nutrition**, v. 100, p. 1413-1421, 2014.

WOOD, P. *et al.* Potencial for β-glucan enrichment in brans derived from oat (*Avena sativa* L.) cultivars of different (1-3),(1-4)-β-D-glucan concentrations. **Cereal Chemistry**, v. 68, p. 48-51, 1991.

WOOD, P. J; BEER, M. U. Productos funcionales de avena. In: MAZZA, G. (Ed.). **Alimentos funcionales: aspectos bioquímicos y de procesado**. Zaragoza: Acribia, 2000. p. 1-38.

XAVIER, A. *et al.* NAM: association studies in multiple populations. **Bioinformatics**, v. 31, p. 3862-3864, 2015.

XAVIER, A. *et al.* Genetic architecture of phenomic-enabled canopy coverage in *Glycine max*. **Genetics**, v. 206, p. 1081-1089, 2017.

TINKER, N. A. *et al.* Haplotag: Software for Haplotype-Based Genotyping-by-Sequencing Analysis. **Genes Genomes and Genetics**, v. 6, p. 857-863, 2016.

YE, E. Q. *et al.* Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease and weight gain. **Journal of Nutrition**, v. 142, p. 1304-1313, 2012.

# 5 CHAPTER 3: CHARACTERIZATION OF β-GLUCAN CONTENT IN PRIMARY AND SECONDARY OAT KERNELS HARVESTED IN DIFFERENT PLANTING DATES AND YEARS


## Abstract

Selection for β-glucan content has been one of the main challenges for oat breeders worldwide. There are several factors limiting a complete elucidation of this trait, including his complex inheritance, the influence exercised by the environmental conditions, and the association between β-glucan content and other traits, such as kernel shape-related ones. In this sense, the objectives of this study were to characterize the β-glucan content in primary and secondary kernels, as well as to identify the influence of planting date and year on this trait. Six oat genotypes were evaluated for β-glucan content and kernel shape-related traits in two years and two planting dates. The experimental design was randomized blocks with two replications, with two meters long for each experimental unit. The phenotypic manifestation of all traits was associated with the interaction among genotype, planting date, and year. In addition, β-glucan content and kernel length traits were associated with the interaction between genotype, kernel type, and year. The β-glucan content varied between primary and secondary kernels, according to the genotype and environmental condition. Under unfavorable conditions, genotypes may present higher β-glucan content in secondary kernels than in primary ones.

## 5.1 Introduction

Oat consumption has gained great importance due to the benefits for human health. The ingestion of oat kernels in proper quantities provide many health benefits, including reduction on cholesterol levels in the blood (Whitehead *et al.*, 2014) and reduction on glycemic index in diabetics (Ye *et al.*, 2012). These benefits are associated with β-glucan consumption, which are soluble dietary fibers found in great proportions in oat kernels. The β-glucan content in oat kernels ranges from 2 to 7%, with overall average of 4.3% for the species. Along with barley, oats have the highest β-glucan content among grasses (Wood & Beer, 2000).

There are many factors influencing the β-glucan content in oats. The environmental conditions of each crop season are strictly associated with the oat β-glucan content (Saastamoinen *et al*., 1992; Doehlert *et al*., 2001; Beber *et al*., 2002; Saastamoinen *et al*., 2004). Regarding Brazilian oat varieties, Crestani *et al.* (2010) verified that the β-glucan content is influenced by the interaction among genotype, environment, and year. Similarly, studying four Canadian oat varieties, Humphreys *et al.* (1994) identified that different planting dates also impact the β-glucan content in oats. In addition, the occurrence of successive rainfalls during grain filling tend to reduce the β-glucan content in oat kernels (Miller *et al*., 1993), while rainfalls during the vegetative stage are positively correlated with the trait in temperate conditions (Herrera *et al*., 2016). In temperate environments, high temperature conditions during the kernel formation also tend to increase the β-glucan content in oats (Saastamoinen *et al*., 1992; Miller *et al*., 1993; Saastamoinen, 1995).

Most of the cultivated oat species show two or three florets per spikelet, rarely four (Valentine, 1995). Oat kernels of the same spikelet, starting from the bottom to the top, are termed as primary, secondary, and tertiary. Brazilian oat varieties usually have well-developed primary and secondary kernels in the same spikelet. Some Brazilian oat varieties rarely have tertiary kernels due to their genetic constitution and stressful

subtropical conditions where are cropped. In addition, Brazilian oat breeders have been selecting for kernel-related traits since 1980's, prioritizing full filling of primary and secondary kernels. In this sense, although selecting for kernel filling and uniformity, a detailed characterization of the β-glucan content for each kernel type separately was not performed to date.

Currently, β-glucan content analyses are performed using samples composed by both primary and secondary kernels without any distinction. In addition, the influence of planting date and year on the β-glucan content of primary and secondary kernels is completely unknow. A hypothesis to explain the phenotypic variation for β-glucan content suggest a dilution effect provided by the endosperm of each kernel, considering the grain filling period of each variety. This hypothesis was elaborated by De Koyer *et al.* (2004), when a common location of heading date and β-glucan QTL suggested a pleiotropic effect caused when later maturing lines fail to complete the optimum kernel filling period, which results in shrunken kernels with less endosperm and enhanced relative β-glucan content. In this sense, the objectives of this study were to characterize the β-glucan content in primary and secondary kernels, as well as to identify the influence of planting date and year on this trait.

## 5.2 Materials and methods

### *5.2.1 Plant material*

Six contrasting oat genotypes were selected for this study based on their phenotypic information available on 'Oat Global' platform (https://oatglobal.umn.edu/) and in personal information of our research group. Two of the six genotypes have high β-glucan content and the other ones have intermediate content. IAC 7 and ND 9508252-9 represent the genotypes with high β-glucan content and were developed by the Agronomic Institute of Campinas and by the North Dakota State University, respectively. IAC 7 is the Brazilian

variety with highest β-glucan content but is not cultivated nowadays. The breeding line ND 9508252-9 is a reselection within HiFi, an oat variety from North America which usually shows β-glucan content ranging from 6% to 7% (McMullen *et al.*, 2005). The genotypes IPR Artemis, URS Taura, UFRGS 8, and UFRGS 881971 have intermediate β-glucan content and were also included in this study. The variety IPR Artemis was developed by the Agronomic Institute of Paraná, while all other genotypes were developed by the Federal University of Rio Grande do Sul (UFRGS) Oat Breeding Program.

*5.2.2 Field experiments*

Genotypes were assessed in two planting dates at the UFRGS Agronomic Experimental Station, located in Eldorado do Sul city, Rio Grande do Sul state, Southern Brazil, in 2017 and 2018. The first planting date was June 12$^{th}$ in 2017 and June 15$^{th}$ in 2018, while the second one was July 7$^{th}$ in both years. The first planting date is within the best sowing window, while the second one is at the end of the sowing window. The experiments were conducted in a randomized complete block design, with two replications. Each experimental unit was composed by two 1 m long rows, spaced 0.25 m apart each other and 0.4 m between plots. Planting was carried out manually at a rate of 60 seeds per row. The base fertilization was 300 kg ha$^{-1}$ of a 5-30-15 N-P-K formula, and topdressing nitrogen, in the form of dry urea, was applied twice, when plants showed four and six extended leaves, at a rate of 33 kg ha$^{-1}$ of N per application. Pests, fungal diseases, and weeds were chemically controlled in both experiments. At the field maturity stage, approximately 50 panicles of each plot were collected for phenotypic analysis.

*5.2.3 Sample preparation*

Oat grains were separated manually from individual panicles originating two subsamples for each plot: primary and secondary grains. About 50 g of each subsample were dehulled using a Wintersteiger LD 180 laboratory thresher (Wintersteiger, Ried,

Austria). A total of 150 dehulled grains of each subsample were randomly selected for grain morphology characterization, while about 15 g of dehulled grains from each subsample were used for β-glucan quantification. For β-glucan quantification, samples were ground in a Willey type mill (Marconi, Piracicaba, Brazil) with 0.5 mm sieve. After milled, each whole flour sample was well homogenized using a set of sieves and stored at -20 ºC until β-glucan quantification.

*5.2.4 Phenotypic assessment*

Grain morphology was assessed at the Rio Grandense Institute of Rice (IRGA), located in Cachoeirinha city, Rio Grande do Sul state, Southern Brazil, using the rice classer Image S (Selgron, Santa Catarina, Brazil). Four morphology-related traits were assessed: grain length, grain width, grain thickness, and length-width ratio. Each grain was assessed separately and a final report was created showing the number of assessed grains per sample, the average of the assessed grains for each trait, and the standard deviation.

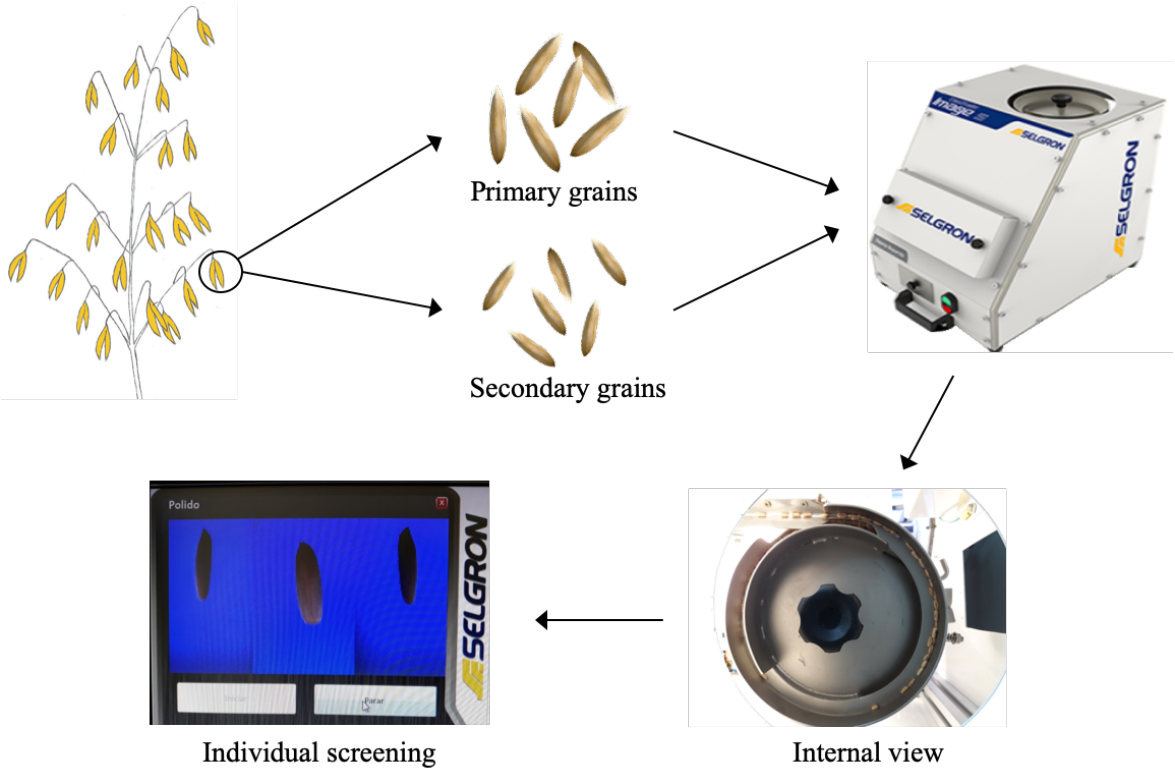



FIGURE 1. Phenotypic assessment for kernel shape in primary and secondary oat kernels.

Chemical analyses for β-glucan were carried out at the Center for Food Research of the University of Passo Fundo, Passo Fundo city, Rio Grande do Sul state, Southern Brazil. β-glucan content was quantified for all samples using the mixed-linkage β-glucan assay kit (Megazyme International Ireland Ltd, Wicklow, Ireland), in duplicates, based on McCleary and Codd (1991), according to AACC Method 32-23.01 and AOAC Method 995.16, with results expressed in dry basis.

### *5.2.5 Localization of β-glucan in primary and secondary kernels*

The β-glucan localization was verified using microscopic analyses at the University of Minnesota Imaging Center. Kernels from the IAC 7 and UFRGS 881971 genotypes, obtained from both planting dates in 2018, were assessed. These two genotypes were used because showed high and intermediate β-glucan content in previous analyses, respectively. The microscopic analyses were performed in primary and secondary kernels separately considering cross sections. Nine kernels were analyzed for each genotype, while only a representative one was imaged. Primary and secondary grains of each genotype were obtained from the same panicle, from a single replication. Briefly, each kernel was cut from both ends, fixed with 3% paraformaldehyde and 1% glutaraldehyde in 0.1 M Na-K phosphate buffer (pH 7.0) and dehydrated in a graded ethanol series. After, the kernels were embedded in hydroxyethyl methyl acrylate. Polymerized samples were sectioned 10 μm thick sections in a microtome. All sections were stained with aqueous 0.01% Calcofluor White M2R (Sigma-Aldrich, St. Louis, USA) for 1 min, as reported in Sikora *et al*. (2013). Stained sections were examined with a Nikon Eclipse 90i microscope (Nikon, Melville, USA) with excitation filter of 325-375 nm and emission filter of 435-485 nm. Micrographs were obtained using a Nikon D2-Fi2 color camera and Nikon Elements software.

*5.2.6 Statistical analyses*

Analyses of variance were carried out to identify the influence of years, planting dates, genotype, and kernel type on the evaluated traits. The factors year, planting date, genotype, and kernel type were included in a factorial design. Replication inside planting date was included as aleatory effect. The comparison of means was performed by the Tukey test at 5% level, using the 'slice' option to explore significant interactions. All statistical analyses were carried out using the SAS 9.4 software (SAS Institute Inc., Cary, USA).

## 5.3 Results

*5.3.1 Environmental influence on β-glucan content and kernel shape-related traits*

The analysis of variance showed significant effect of the interaction among year, planting date, and genotype on all assessed traits (Table 1). In addition, the triple interaction among year, genotype, and kernel type was also significant for β-glucan content and kernel length in oats (Table 1).

TABLE 1. Analysis of variance for β-glucan content and kernel shape-related traits in oats.

| Cause of variation | DF | Mean square | | | |
|---|---|---|---|---|---|
| | | β-glucan | Kernel length | Kernel width | Kernel thickness |
| Replication(PD) | 2 | 0.124 | 0.004 | 0.001 | 0.001 |
| Year (Y) | 1 | 13.769$^{**}$ | 14.348$^{**}$ | 0.243$^{**}$ | 0.031$^{**}$ |
| Planting date (PD) | 1 | 1.009$^{**}$ | 0.285$^{**}$ | 0.102$^{**}$ | 0.063$^{**}$ |
| Genotype (G) | 5 | 7.768$^{**}$ | 7.763$^{**}$ | 0.493$^{**}$ | 0.257$^{**}$ |
| Kernel type (KT) | 1 | 0.574$^{*}$ | 26.091$^{**}$ | 1.917$^{**}$ | 0.591$^{**}$ |
| Y*G | 5 | 0.348$^{*}$ | 0.300$^{**}$ | 0.005$^{*}$ | 0.007$^{**}$ |
| Y*PD | 1 | 0.012$^{ns}$ | 0.603$^{**}$ | 0.007$^{ns}$ | 0.042$^{**}$ |
| Y*KT | 1 | 0.142$^{ns}$ | 2.414$^{**}$ | 0.001$^{ns}$ | 0.009$^{*}$ |
| PD*KT | 1 | 0.141$^{ns}$ | 0.137$^{*}$ | 0.003$^{ns}$ | 0.001$^{ns}$ |
| G*PD | 5 | 0.269$^{ns}$ | 0.084$^{*}$ | 0.008$^{**}$ | 0.014$^{**}$ |
| G*KT | 5 | 0.241$^{ns}$ | 0.368$^{**}$ | 0.036$^{**}$ | 0.011$^{**}$ |
| Y*PD*KT | 1 | 0.051$^{ns}$ | 0.033$^{ns}$ | 0.001$^{ns}$ | 0.001$^{ns}$ |
| Y*G*PD | 5 | 0.307$^{*}$ | 0.225$^{**}$ | 0.007$^{**}$ | 0.008$^{**}$ |
| Y*G*KT | 5 | 0.390$^{*}$ | 0.175$^{**}$ | 0.001$^{ns}$ | 0.001$^{ns}$ |
| G*PD*KT | 5 | 0.045$^{ns}$ | 0.031$^{ns}$ | 0.001$^{ns}$ | 0.001$^{ns}$ |
| Y*PD*KT*G | 5 | 0.050$^{ns}$ | 0.036$^{ns}$ | 0.001$^{ns}$ | 0.001$^{ns}$ |
| Experimental error | 46 | 0.120 | 0.028 | 0.002 | 0.002 |
| ‡Y | - | 54.294 | 27.110 | 8.590 | 2.978 |
| ‡PD | - | 3.979 | 0.538 | 3.606 | 6.052 |
| ‡G | - | 30.631 | 14.668 | 17.427 | 24.688 |
| ‡KT | - | 2.263 | 49.298 | 67.762 | 56.772 |
| CV (%) | - | 6.990 | 2.430 | 1.880 | 2.350 |

$^{**}$ and $^{*}$ significant at 1% and 5% by the F test, respectively. $^{ns}$, non-significant. ‡, % of total mean square. BG, β-glucan content. DF, degrees of freedom.

The percentage of total mean square accounted for year, planting date, genotype, and kernel type are also shown in Table 1. For the β-glucan content trait, year and genotype accounted for 54% and 31% of the total variance, respectively, while planting date and kernel type together accounted for only 6% of the total variance (Table 1). Considering kernel length, kernel type (49%) and year (27%) played major impact on the total variation. Kernel type showed high influence on kernel width (68%) and kernel thickness (57%), followed by genotype with 17% and 25% on the respective traits (Table 1).

The factor year was isolated in all analyses, allowing the comparison among genotypes inside planting date/kernel type, inside genotype between planting date/kernel type, and inside genotype between year for each planting date/kernel type. In 2017, IAC7 and HiFi showed highest β-glucan content on the first planting date, while UFRGS 881971 and IPR Artemis were the lowest ones (Table 2). IAC 7 also showed the highest β-glucan content on the second planting date of 2017 (Table 2). As for the first planting date, UFRGS 881971 presented low β-glucan content in the second one, not differing from IPR Artemis (Table 2). In 2018, IAC7 and HiFi showed highest β-glucan content, regardless of the planting date. In the opposite hand, the genotypes IPR Artemis, UFRGS 8, and UFRGS 881971 presented lowest β-glucan content in both planting dates (Table 2). Variations of β-glucan content between planting dates, for each genotype and year, were not significant neither in 2017 nor in 2018 (Table 2). On the other hand, variations between year, for each genotype and planting date, were verified in both planting dates for β-glucan content. All genotypes showed higher β-glucan content in the first planting date of 2017 than in the 2018 one (Table 2). Similarly, IPR Artemis showed higher β-glucan content in the second planting date of 2017 than in the 2018 one, while all other genotypes did not differ in response to the year effect on this planting condition (Table 2).

TABLE 2. Averages for β-glucan content and kernel shape-related traits in oats. Comparison between years, planting dates, and genotypes.

| Year | Genotype | β-glucan content | | Kernel length | | Kernel width | | Kernel thickness | |
|---|---|---|---|---|---|---|---|---|---|
| | | 1st PD | 2nd PD | 1st PD | 2nd PD | 1st PD | 2nd PD | 1st PD | 2nd PD |
| 2017 | IPR Artemis | 4.7 Ac | 4.83 Acd | 8.79 Aa | 8.01 Aa | 2.44 Ac | 2.31 Ac | 2.04 Aa | 1.85 Ac |
| | IAC7 | 6.36 Aa | 6.69 Aa | 8.23 Ab | 7.99 Aa | 2.29 Ad | 2.27 Ac | 1.84 Ab | 1.79 Ac |
| | HiFi | 6.36 Aa | 5.43 Ab | 6.43 Ae | 6.8 Ac | 2.33 Ad | 2.12 Ad | 1.82 Ab | 1.58 Bd |
| | URS Taura | 5.41 Ab | 5.23 Abc | 6.64 Ae | 6.33 Ad | 2.62 Aa | 2.60 Aa | 2.06 Aa | 2.06 Aa |
| | UFRGS 8 | 5.34 Ab | 4.96 Abc | 7.51 Ac | 7.12 Ab | 2.53 Ab | 2.43 Ab | 2.03 Aa | 1.96 Bb |
| | UFRGS 881971 | 4.42 Ac | 4.35 Ad | 6.93 Ad | 6.67 Ac | 2.62 Aa | 2.61 Aa | 2.02 Aa | 2.00 Aab |
| 2018 | IPR Artemis | 3.96 Bc | 4.00 Bc | 7.06 Ba | 7.22 Ba | 2.32 Bc | 2.28 Ac | 1.88 Bb | 1.86 Ac |
| | IAC7 | 5.77 Ba | 5.08 Aa | 7.24 Ba | 7.32 Ba | 2.14 Bd | 2.15 Bd | 1.76 Bc | 1.76 Ad |
| | HiFi | 5.43 Ba | 5.11 Aa | 6.22 Ac | 6.1 Bc | 2.11 Bd | 2.04 Be | 1.76 Ac | 1.70 Ae |
| | URS Taura | 4.85 Bb | 4.52 Ab | 5.98 Bd | 5.95 Bc | 2.54 Ba | 2.48 Ba | 2.00 Aa | 1.96 Bb |
| | UFRGS 8 | 4.18 Bc | 3.96 Ac | 6.55 Bb | 6.7 Bb | 2.46 Bb | 2.40 Ab | 1.96 Ba | 2.04 Aa |
| | UFRGS 881971 | 3.96 Bc | 4.14 Abc | 6.07 Bc | 5.94 Bc | 2.56 Aa | 2.48 Ba | 1.98 Aa | 1.96 Ab |

PD, planting date. Capital letters are comparing year within genotype and planting date, while lowercase letters are comparing genotypes within planting date and year. Differences between planting date within year and genotype were not identified.

For kernel length, IPR Artemis showed longer kernels in the first planting date of 2017, while HiFi and URS Taura were the shortest ones in this same planting date (Table 2). For the second planting date of 2017, IPR Artemis and IAC 7 showed higher kernel length than the other ones, while URS Taura showed shortest kernel length (Table 2). Considering the same trait in 2018, the genotypes IPR Artemis and IAC 7 showed highest kernel length, regardless of the planting date (Table 2). URS Taura showed the shortest kernel length in the first planting date of 2018 (Table 2). In the second planting date of 2018, URS Taura, HiFi, and UFRGS 881971 showed shortest kernel length (Table 2). Variation on kernel length between planting date were not verified in both years. In general, kernel length was more pronounced in 2017 than in 2018 (Table 2). Except for the genotype HiFi on the first planting date, all other genotypes showed higher kernel length in 2017, regardless of the planting date (Table 2).

The genotypes URS Taura and UFRGS 881971 showed highest kernel width on both planting dates, regardless of the year (Table 2). On the other hand, within the first planting date, the genotypes IAC 7 and HiFi showed narrowest kernel widths, in both evaluated years. Within the second planting date of 2018, the genotype HiFi showed narrower kernels than all others. There was no significant variation for kernel width between planting dates inside each year (Table 2). Considering the year effect, IAC 7, HiFi, IPR Artemis, UFRGS 8, and URS Taura showed higher kernel width on the first planting date of 2017 than in the 2018 one, while no differences were observed for UFRGS 881971 (Table 2). For the second planting date, the genotypes IAC 7, HiFi, URS Taura, and UFRGS 881971 presented higher kernel width in 2017 than in 2018 (Table 2).

The genotypes IPR Artemis, URS Taura, UFRGS 8, and UFRGS 881971 showed thicker kernels in 2017, within the first planting date. In the opposite way, the genotypes IAC 7 and HiFi showed short kernel thickness on the same planting date in 2017 (Table 2). Within the second planting date of 2017, the genotype URS Taura kept its high

performance for the kernel thickness trait, not differing from UFRGS 881971, while the genotype HiFi showed the lowest performance for the trait (Table 2). In 2018, the genotypes URS Taura, UFRGS 8, and UFRGS 881971 kept their high averages for kernel thickness on the first planting date, while UFRGS 8 showed higher thickness average on the second one (Table 2). The genotypes IAC 7 and HiFi showed the smallest kernel thickness among genotypes, on the first planting date of 2018, while HiFi showed the narrower kernel thickness on the second planting date, exactly as verified in 2017 (Table 2). As mentioned before, variations for kernel thickness between planting dates were not observed in both years (Table 2). On the other hand, within genotypes and planting date, variations for kernel thickness were observed between years (Table 2). The genotypes IPR Artemis, IAC 7, and UFRGS 8 reduced their kernel thickness average on the first planting of 2018, when compared to 2017 (Table 2). On the second planting date, the genotypes HiFi and UFRGS 8 increased their kernel thickness in response to the environmental conditions of 2018, when compared to 2017, while URS Taura reduced the kernel thickness average in this same condition (Table 2).

### *5.3.2 β-glucan content and kernel length differences between primary and secondary kernels*

The interaction among year, planting date, and kernel type was sliced using mean comparisons (Table 3). The differences for β-glucan content among genotypes were more or less consistent when compared primary and secondary kernels separately. In 2017, when compared between kernel types, the genotypes IPR Artemis, HiFi, URS Taura, and UFRGS 881971 had primary kernels with higher β-glucan content. On the other hand, IAC 7 showed higher β-glucan content in secondary kernels, while UFRGS 8 did not vary between kernel type (Table 3). In 2018, there were no significant differences for β-glucan content between primary and secondary kernels, except for UFRGS 881971 that showed higher β-glucan content in primary kernels (Table 3).

Considering the performance of each genotype, inside kernel type and between years, significant differences were observed. Except for UFRGS 881971, all other genotypes reduced the β-glucan content of primary kernels in response to the environmental conditions of 2018, compared to 2017 (Table 3). As observed for primary kernels, the genotypes IAC 7 and UFRGS 8 also reduced the β-glucan content in secondary kernels in response to 2018 (Table 3).

Comparing kernel length averages between years, within genotype and kernel type, except URS Taura all other genotypes reduced their primary kernel length averages in 2018, when compared to 2017 (Table 3). For secondary kernels, all genotypes reduced their kernel length averages in 2018, when compared to 2017 (Table 3).

TABLE 3. Averages for β-glucan content and kernel length in oats. Comparison between years, kernel type, and genotypes.

| Year | Genotype | β-glucan content | | Kernel length | |
|---|---|---|---|---|---|
| | | Primary | Secondary | Primary | Secondary |
| 2017 | IPR Artemis | A 5.01 Ac | A 4.51 Bd | A 8.74 Aa | A 8.06 Ba |
| | IAC 7 | A 6.31 Ba | A 6.73 Aa | A 8.32 Ab | A 7.90 Ba |
| | HiFi | A 6.23 Aa | A 5.56 Bb | A 7.36 Ad | A 5.86 Bd |
| | URS Taura | A 5.57 Ab | A 5.07 Bc | A 6.52 Ae | A 6.46 Ac |
| | UFRGS 8 | A 4.99 Ac | A 5.31 Abc | A 7.73 Ac | A 6.90 Bb |
| | UFRGS 881971 | A 4.61 Ac | A 4.16 Bd | A 7.24 Ad | A 6.36 Bc |
| 2018 | IPR Artemis | B 3.85 Ac | A 4.12 Ab | B 7.80 Aa | B 6.48 Ba |
| | IAC 7 | B 5.61 Aa | B 5.24 Aa | B 7.94 Aa | B 6.63 Ba |
| | HiFi | B 5.37 Aa | A 5.16 Aa | B 6.97 Ac | B 5.36 Bcd |
| | URS Taura | B 4.56 Ab | A 4.79 Aa | A 6.51 Ad | B 5.22 Bd |
| | UFRGS 8 | B 4.06 Ac | B 4.08 Ab | B 7.40 Ab | B 5.84 Bb |
| | UFRGS 881971 | A 4.26 Abc | A 3.85 Bb | B 6.55 Ad | B 5.47 Bc |

Primary, primary kernel. Secondary, secondary kernel. Capital letters before values are comparing years within genotype and kernel type. Capital letters after values are comparing kernel types within year and genotype. Lowercase letters after values are comparing genotypes within kernel type and year.

In 2017, the highest kernel length of primary kernels was observed in IPR Artemis, while URS Taura showed the shortest primary kernels (Table 3). Considering the length of secondary kernels in 2017, the genotypes IPR Artemis and IAC 7 showed highest averages, while HiFi showed the lowest one (Table 3). In 2018, the genotypes IPR Artemis and IAC 7 showed highest kernel length, regardless of the kernel type. The genotypes URS

Taura and UFRGS 881971 showed shortest length for primary kernels, while URS Taura also revealed lowest average for this trait for secondary kernels (Table 3). Considering the length variation between kernel types in 2017, most of the genotypes showed primary kernels longer than secondary ones. Only the genotype URS Taura did not differ for the kernel length trait between primary and secondary kernels (Table 3). In 2018, all genotypes showed primary kernels longer when compared to secondary ones (Table 3).

### *5.3.3 Localization of β-glucans in primary and secondary kernels*

The image analysis of primary kernels from UFRGS 881971 and IAC 7 showed different β-glucan distribution pattern between them (Figure 2). The image analysis reinforces a contrasting for kernel shape between UFRGS 881971 and IAC 7, as presented previously, IAC 7 has reduced kernel width and kernel thickness when compared to UFRGS 881971 (Figure 2 and Table 3).

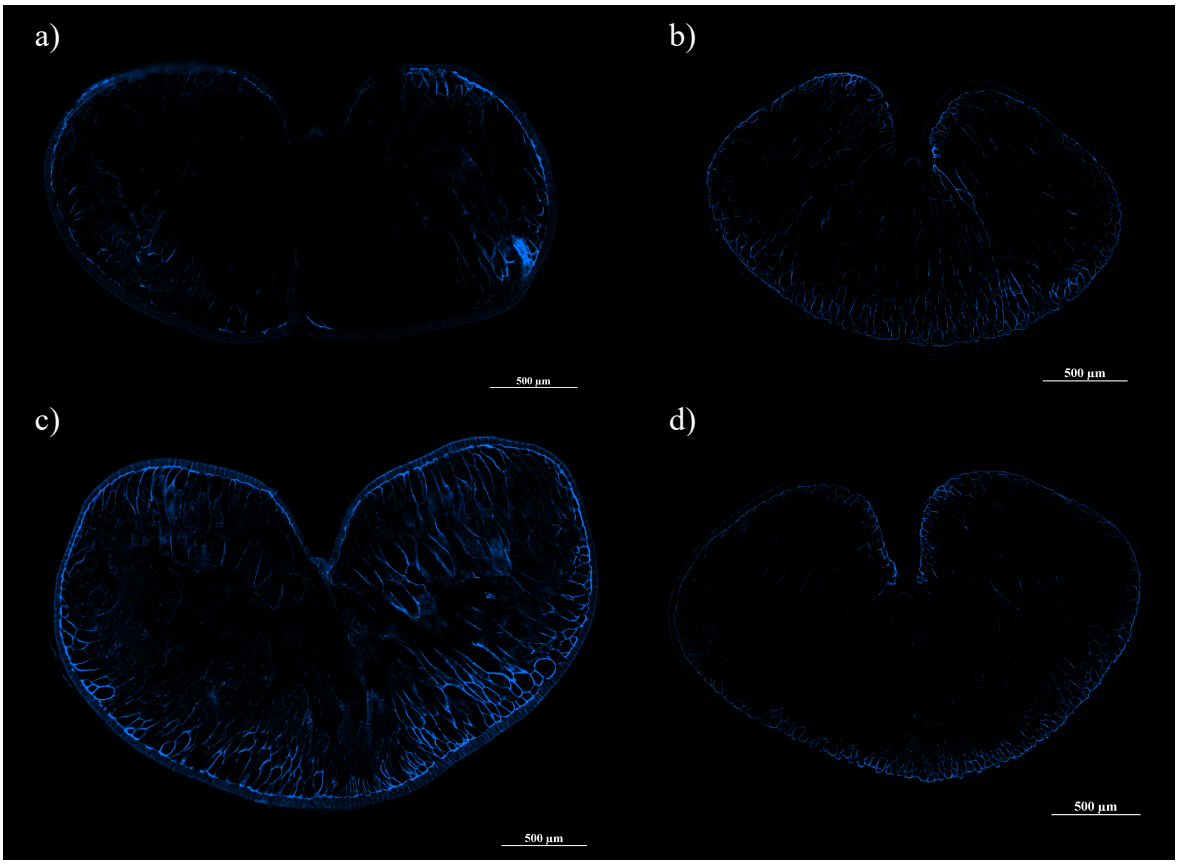



FIGURE 2. β-glucan (blue) formation on oat primary kernels in two planting dates. a) UFRGS 881971 and b) IAC 7, first planting date; c) UFRGS 881971 and d) IAC 7, second planting date.

On the first planting date, the genotype UFRGS 881971 showed β-glucans moderately located in the external area of the kernel, close to the aleurone layer and aleurone-endosperm junction (Figure 2a). The IAC 7's primary kernel also showed moderate β-glucan deposition on the first planting date, but with more pronounced distribution, reaching a wide part of the kernel when compared to UFRGS 881971 (Figure 2b). On the second planting date, UFRGS 881971's primary kernels showed higher β-glucan accumulation than IAC 7's ones in the same planting date (Figure 2c and Figure 2d), with pronounced β-glucan deposition in the aleurone layer and also in the starchy endosperm (Figure 2c). The primary kernel from IAC 7 showed β-glucan deposition in the external area of the kernel on the second planting date (Figure 2d), differing when compared to the first one (Figure 2b).

Differences on β-glucan deposition were also observed between secondary kernels from UFRGS 881971 and IAC 7 in both planting dates (Figure 3). On the first planting date, secondary kernels from UFRGS 881971 showed more pronounced β-glucan deposition in the external area of the kernel, including the aleurone layer and moving smoothly to the endosperm (Figure 3a). On the same planting date, secondary kernels from IAC 7 also showed β-glucan deposition in the aleurone layer, reaching more intensely the starchy endosperm (Figure 3b). In this sense, secondary kernels from IAC 7 showed higher β-glucan deposition than UFRGS 881971 ones on the first planting date. In addition, the number of cells in secondary kernels from IAC 7 seems to be higher when compared to UFRGS 881971 ones in the first planting date (Figure 3a and Figure 3b).

The formation and deposition of β-glucans reduced in secondary kernels from the second planting date, regardless of the genotype (Figure 3). On the second planting date, secondary kernels from UFRGS 881971 showed intense β-glucan deposition in the external area of the kernel, with low deposition in the endosperm (Figure 3c). In the same

planting date, secondary kernels from IAC 7 had modest β-glucan deposition, but reaching in a uniform way to the endosperm (Figure 3d).

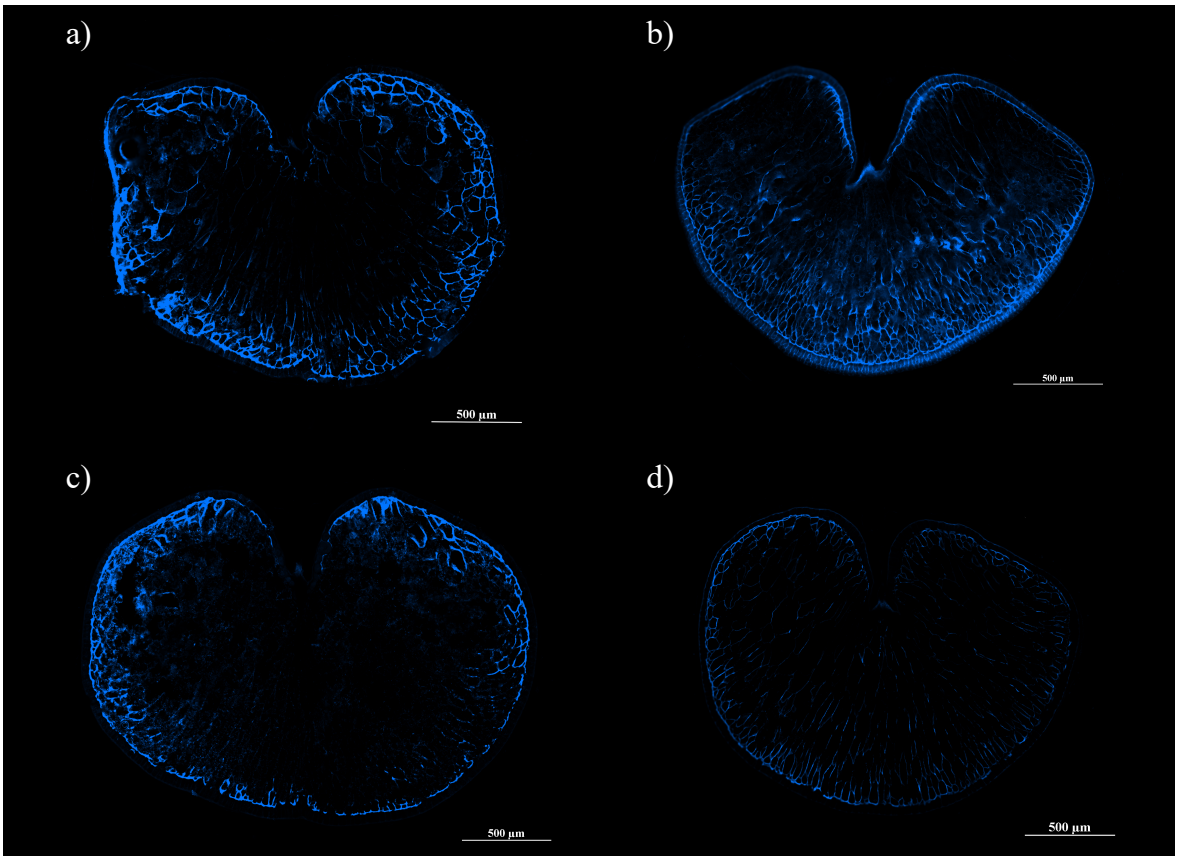



FIGURE 3. β-glucan (blue) formation on oat secondary kernels in two planting dates. a) UFRGS 881971 and b) IAC 7, first planting date; c) UFRGS 881971 and d) IAC 7, second planting date.

In general, the β-glucan deposition seems to be more pronounced in secondary kernels from UFRGS 881971 when compared to IAC 7 ones on the second planting date. Both genotypes showed secondary kernels with lower β-glucan deposition on the second planting date (Figure 3c and Figure 3d).

### 5.4 Discussion

β-glucan content and kernel shape-related traits are highly influenced by environmental conditions. The results indicated that all assessed traits depend of the triple interaction among year, planting date, and genotype (Table 1). The year accounted for 54% of the β-glucan content variation, followed by genotype with 31% of the total variation of the trait (Table 1). Planting date accounted for only 4% of the β-glucan content variation,

justifying why differences between planting dates were not detected for this trait (Table 2). Significant differences were observed between years, especially on the first planting date (Table 2). Similar results were observed by Herrera *et al.* (2016) that identified a major contribution of year on the oat β-glucan content, followed by genotype.

The year influence was also observed for kernel shape-related traits, especially kernel length and kernel width, regardless of the planting date (Table 2). The ANOVA revealed that year accounted for 27% and 9% of the total variation for kernel length and kernel width, respectively, while minor contribution (3%) was observed on kernel thickness (Table 1). In addition, kernel type was the major factor contributing for the observed variation, accounting for 49%, 68%, and 57% of the variation for kernel length, kernel width, and kernel thickness, respectively (Table 1). In this sense, the environmental conditions of each year were decisive for the phenotypic manifestation of the assessed traits.

Contrasting environmental conditions were observed in the evaluated years, which were classed as unfavorable and favorable. In 2017, a stressful condition occurred after planting, with low/uneven rainfall during the vegetative stage, excess during the reproductive one, and high temperature (Supplementary Figure 1A). On the other hand, the environmental conditions were favorable in 2018, with well distributed rainfall throughout the year and lower temperature averages than in 2017 (Supplementary Figure 1B). Other studies also support the findings of the present study (Crestani *et al.*, 2010; Humphreys *et al.* 1994). Studying the β-glucan content in 15 Brazilian oat genotypes, Crestani *et al.* (2010) identified a significant interaction among genotype, environment, and year influencing the trait. Similarly, Humphreys *et al.* (1994) identified the effect of planting date on β-glucan content in Canadian varieties.

The results from 2017 suggests a possible association between β-glucan content and kernel-shape related traits under unfavorable conditions. The genotypes IAC 7 and

HiFi, which showed highest β-glucan content in 2017, presented reduced kernel width and thickness (Table 2). On the other hand, the genotype UFRGS 881971, which showed low β-glucan content in 2017, showed large kernel width and thickness (Table 2). In 2018, even with favorable environmental conditions, the genotypes IAC 7 and HiFi showed high β-glucan content and smaller kernel width and thickness (Table 2).

Another result that supports the association between β-glucan content and kernel-shape related traits is the significant interaction among year, genotype, and kernel type (Table 3). Secondary kernels from the IAC 7 genotype showed higher β-glucan content than primary ones in 2017 (Table 3). This result indicates that some genotypes can modulate the β-glucan deposition according to the environment and kernel type. On the other hand, the IPR Artemis genotype, which has relatively narrow kernel width and kernel thickness, showed low performance for β-glucan content and long kernel length, regardless of the kernel type (Table 3). These results reinforce that the selection for narrow kernel width and thickness is not a guarantee of high beta-glucan content in oats. There are no other studies associating β-glucan content and kernel-shape related traits in Brazil. As consequence, the discussion of this chapter is limited. The association between β-glucan content and kernel-shape related traits used only mean analysis. Specific association analyses, such as correlation and regression coefficients could be used in this chapter. However, these analyses were not performed because a small dataset was produced, which could produce unreliable estimates.

In order to clarify the association between β-glucan content and kernel-shape related traits in oats, a detailed kernel image analysis was performed for primary and secondary kernels separately. The results illustrate a contrasting β-glucan deposition pattern between IAC 7 and UFRGS 881971 in both kernel types (Figure 2 and Figure 3). The deposition of β-glucans in the genotype IAC 7 reaches the endosperm, resulting in high β-glucan content, while in UFRGS 881971 is limited to the external area of the kernel

(Figure 2 and Figure 3). In addition, the images showed a higher deposition of β-glucans in secondary kernels than in primary ones (Figure 2 and Figure 3). The deposition of β-glucans in the secondary kernel from IAC 7 (Figure 3b) confirms the high β-glucan content average observed in secondary kernels in 2017 (Table 3). The images revealed a considerable β-glucan deposition for the UFRGS 881971 genotype, especially in secondary kernels. However, the absence of β-glucans in the starchy endosperm seems to be a decisive factor for smaller deposition. The β-glucan content of the UFRGS 881971 genotype is probably diluted by the endosperm, resulting in lower β-glucan content. Further studies involving more genotypes and planting dates may contribute for better understanding of the β-glucan formation in different oat kernel types.

### 5.5 Conclusions

The β-glucan content varies between primary and secondary kernels, according to the genotype and environmental condition. Under unfavorable conditions, some genotypes can present higher β-glucan content in secondary kernels than in primary ones, such as IAC 7 in 2017. The phenotypic manifestation of β-glucan content and kernel-shape related traits is associated with the interaction among genotype, year, and planting date, and/or genotype, year, and kernel type.

### 5.6 References

AACC - American Association of Cereal Chemists. Approved methods, 9 ed. Saint Paul: AACC, 1999.

AOAC - Association of Official Analytical Chemistry. Official methods of analysis of the Association of Official Analytical Chemistry, 16 ed. Washington: AOAC, 1997.

BEBER, R. C. *et al*. Caracterização química de genótipos brasileiros de aveia (*Avena sativa* L.). **Acta Científica Venezuelana**, v. 53, p. 202-209, 2002.

CRESTANI, M. *et al*. β-glucan content in white oat cultivars grown in different environments. **Pesquisa Agropecuária Brasileira**, v. 45, p. 261-268, 2010.

DE KOEYER, D. L. *et al.* A molecular linkage map with associated QTLs from a hulless x covered spring oat population. **Theoretical and Applied Genetics**, v. 108, p. 1285-1298, 2004.

DOEHLERT, D. C.; MCMULLEN, M. S.; HAMMOND, J. J. Genotypic and environmental effects on grain yield and quality of oat grown in North Dakota. **Crop Science**, v. 41, p. 1066-1072, 2001.

HERRERA, M. P. *et al.* β-Glucan content, viscosity, and solubility of Canadian grown oat as influenced by cultivar and growing location. **Canadian Journal of Plant Science**, v. 96, p. 183-196, 2016.

HUMPHREYS, D. G.; SMITH, D. L.; MATHER, D. E. Nitrogen fertilizer and seeding date induced charges in protein, oil and β-glucan contents of four oat cultivars. **Journal of Cereal Science**, v. 20, p. 283-290, 1994.

MCCLEARY, B. V.; CODD, R. Measurement of (1→3),(1→4)-β-D-Glucan in Barley and Oats: A Streamlined Enzymic Procedure. **Journal of the Science of Food and Agriculture**, v. 55, p. 303-312, 1991.

MCMULLEN M. S.; DOEHLERT, D. C.; MILLER, J. D. Registration of 'HiFi' oat. **Crop Science**, v. 45, p. 1664, 2005.

MILLER, S. S. *et al.* Oat β-glucans: an evaluation of eastern canadian cultivars and unregistered lines. **Canadian Journal of Plant Science**, v. 73, p. 429-436, 1993.

SAASTAMOINEN, M.; PLAAMI, S.; KUMPULAINEN, J. Genetic and environmental variation in β-glucan content of oats cultivated or tested in Finland. **Journal of Cereal Science**, v. 16, p. 279-290, 1992.

SAASTAMOINEN, M. Effects of environmental factors on the β-glucan content of two oat varieties. **Acta Agriculturae Scandinavica**, v. 45, p. 181-187, 1995.

SAASTAMOINEN, M. *et al.* β-glucan contents of groats of different oat cultivars in official variety, in organic cultivation, and nitrogen fertilization trials in Finland. **Agricultural and Food Science**, v. 13, p. 68-79, 2004.

SIKORA, P. *et al.* Identification of high β-glucan oat lines and localization and chemical characterization of their seed kernel β-glucans. **Food chemistry**, v. 137, p. 83-91, 2013.

VALENTINE, J. Naked Oats. In: WELCH, R. W. (Ed.). **The oat crop: production and utilization**. London: Chapman & Hall, 1995. p. 505-532.

WHITEHEAD A. *et al.* Cholesterol-lowering effects of oat β-glucan: a meta-analysis of randomized controlled trials. **American Journal of Clinical Nutrition**, v. 100, p. 1413-1421, 2014.

WOOD, P. J.; BEER, M. U. Productos funcionales de avena. In: MAZZA, G. (Ed.). **Alimentos funcionales: aspectos bioquímicos y de procesado.** Zaragoza: Acribia, 2000. p. 1-38.

YE, E. Q. *et al.* Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease and weight gain. **Journal of Nutrition**, v. 142, p. 1304-1313, 2012.

# 6 CHAPTER 4: CHARACTERIZATION AND ABSOLUTE QUANTIFICATION OF THE *Cellulose synthase-like F6* GENE IN OATS


## Abstract

The *Cellulose synthase-like F6* (*CslF6*) gene has been reported as a key factor for the β-glucan synthesis in grasses. However, the action of this gene on oat β-glucan content is poorly understood to date. In this sense, the objectives of this study were: i) to quantify the absolute expression of the *CslF6* transcripts in oats under field conditions; and ii) to characterize the *CslF6* homoeologs in contrasting oat genotypes for β-glucan content. Developing kernels of five oat genotypes were collected under field conditions, in two planting dates. The *CslF6* absolute quantification was performed using droplet digital PCR in two different approaches: i) global expression; and ii) subgenome specific quantifications. The third exon of the *CslF6* gene, which corresponds to approximately 62.2% of the coding region, was sequenced in each subgenome using PacBio sequencing. The *CslF6* expression was highly influenced by the planting date, kernel developmental stage, and genotype. High *CslF6* transcript number was not associated with high β-glucan content in oats, indicating the action of a complex biosynthesis pathway controlling this trait. Under stressful conditions, the *CslF6* expression was pronounced and associated with β-glucan content. The CslF6 motif has 1767 bases pair and is highly conserved between A and D subgenomes. Single nucleotide polymorphisms were not identified among the assessed genotypes, regardless of the subgenome. The results presented in this study may contribute to the elucidation of the CslF6 influence on oat β-glucan content.

## 6.1 Introduction

Selection for β-glucan content has been one of the main objectives of oat breeding programs worldwide. β-glucans, which are soluble dietary fibers found in oat kernels, are associated with many benefits for human health, including improvement in the immune function and reductions of the glycemic index, cholesterol levels in the blood, obesity, and cardiovascular disease risk (Cho *et al.*, 2013; Whitehead *et al*., 2014; Ye *et al*., 2012). Although β-glucan content has been targeted of oat breeders in the last decades, selection for this trait is still difficult due to limitations on phenotyping methods, as well as the environmental influence on the trait. Most importantly, the genetic mechanisms controlling oat β-glucan content are not fully understood to date.

The genetic mechanisms involved on oat β-glucan synthesis are scarce in the literature. β-glucan content in cereals is a quantitative trait controlled by genes with additive effects (Holthaus *et al*., 1996; Crestani *et al*., 2012). There are many classes of enzymes involved on the synthesis and deposition of polysaccharides into the cell wall. One of these classes are the glycosyltransferases, which act modifying carbohydrates (Coutinho and Henrissat, 1999) and are divided in two main groups: i) polysaccharides synthases (type I); and ii) glycosyltransferases (type II). Type I enzymes are coded by Cellulose Synthase gene family members, which are divided in different gene subfamilies, including Cellulose synthase (CesA) and Cellulose synthase-like (Csl) (Richmond and Somerville, 2000). Csl subfamily genes are classed from ‘A’ to ‘H’ and were identified in different plant species, including 30 and 37 members in Arabidopsis and rice, respectively (Hazen *et al*., 2002; Somerville *et al*., 2004).

A comparative mapping study showed that a cluster of the CslF gene subfamily from rice is located in a syntenic region corresponding to a quantitative trait loci (QTL) with major effects on β-glucan content in barley (Burton *et al*., 2006). In addition, the insertion and expression of rice CslF genes into Arabidopsis allowed the detection of β-

glucan in the cell walls (Burton *et al*., 2006), showing that these genes are a key component in the β-glucan synthesis pathway. In wheat, RNAi silencing of the *Cellulose synthase-like F6* (*CslF6*) gene showed an average reduction of 42.4% of the β-glucan content in the kernels (Nemeth *et al*., 2010). In barley, *CslF6* overexpression under an oat endosperm-specific promoter increased more than 80% of the β-glucan content in the kernels (Burton *et al*., 2011). In addition, *betaglucanless* (*bgl*) mutants in barley have a point mutation that changes a highly conserved amino acid in the encoded protein (Taketa *et al.,* 2012).

Studying the expression of seven CslF subfamily members in the barley varieties 'Himalaya' and 'Sloop', which are contrasting for β-glucan content in their kernels, Burton *et al.* (2008) observed different expression profile for the *CslF6* and *Cellulose synthase-like F9* (*CslF9*) genes. These genes were the only ones CslF genes expressed during formation and filling of barley kernels. The *CslF9* gene showed a peak of expression eight days after anthesis (DAA) in the 'Sloop' variety, which has reduced β-glucan content in his kernels. On the other hand, the *CslF6* gene increased his expression between 12 and 20 DAA, reaching a peak of expression at 20 DAA in the 'Himalaya' variety, which has high β-glucan content (Burton *et al.*, 2008). These results suggested a distinct and opposite expression of the *CslF6* and *CslF9* genes, being the first one acting positively on the β-glucan content.

The studies presented previously indicate that Csl gene family members are highly associated with β-glucan content in small grains. These genes encoded synthases that act together with other enzymes on the complex β-glucan biosynthesis pathway, as well as on its accumulation in the cell walls. Among many genes, *CslF6* has been described as a key component of the β-glucan synthesis in cereals because knockout mutants of this gene have essentially no β-glucan (Taketa *et al.*, 2012; Vega-Sánchez *et al*., 2012). *CslF6* may acts on β-glucan content trait in different ways, due to his expression (increasing on

transcript level) or by mutations in its coding region that could produce glycosyltransferases with different functional properties. The *CslF6* gene have been identified and cloned in oats (Chawade *et al*., 2010; Jobling, 2015). However, studies identifying polymorphisms on the *CslF6* gene between contrasting oat varieties, possibly associated with the β-glucan content trait, as well as the expression of this gene, are scarce in the literature. In this sense, the objectives of this study were: i) to quantify the absolute expression of the *CslF6* transcripts in oats under field conditions; and ii) to characterize the *CslF6* gene homoeologs in contrasting oat genotypes for β-glucan content.

## 6.2 Materials and methods

### *6.2.1 Plant material*

Five oat genotypes were tested in this study, one of them with high β-glucan content and four intermediate ones. The variety IAC 7, which was developed by the Agronomic Institute of Campinas, represents the genotype with higher β-glucan content. The genotypes IPR Artemis, URS Taura, UFRGS 8, and UFRGS 881971, which have intermediate β-glucan content, were also assessed. The variety IPR Artemis was developed by the Agronomic Institute of Paraná, while all other genotypes were developed by the Federal University of Rio Grande do Sul (UFRGS) Oat Breeding Program. A reselection of the HiFi oat variety, which has high β-glucan content (McMullen *et al*., 2005), was included in the *CslF6* gene characterization.

### *6.2.2 Field experiment*

Genotypes were assessed in two planting dates at the UFRGS Agronomic Experimental Station, located in Eldorado do Sul city, Rio Grande do Sul state, Southern Brazil, in 2017. The first planting date was June 12$^{th}$, while the second one was July 7$^{th}$. The experiment was conducted in a randomized complete block design, with two replications. Each experimental unit was composed by two 1 m long rows, spaced 0.2 m

apart each other and 0.4 m between plots. Planting was carried out manually at a rate of 60 seeds per row. The base fertilization was 300 kg $ha^{-1}$ of a 5-30-15 N-P-K formula, and topdressing nitrogen, in the form of dry urea, was applied twice, when plants showed four and six extended leaves, at a rate of 33 kg $ha^{-1}$ of N per application. Pests, fungal diseases, and weeds were chemically controlled. At the field maturity stage, approximately 50 panicles of each plot were collected for phenotypic analysis.

*6.2.3 β-glucan content analysis*

β-glucan content was quantified at the Center for Food Research of the University of Passo Fundo, Passo Fundo city, Rio Grande do Sul state, Southern Brazil. About 15 g of dehulled grains from each plot were ground in a Willey type mill (Marconi, Piracicaba, Brazil) with 0.5 mm sieve. After milled, each whole flour sample was well homogenized using a set of sieves and stored at -20 ºC until quantification. β-glucan content was quantified using the mixed-linkage β-glucan assay kit (Megazyme International Ireland Ltd, Wicklow, Ireland), in duplicates, based on McCleary and Codd (1991), according to AACC Method 32-23.01 and AOAC Method 995.16, with results expressed in dry basis.

*6.2.4 DNA isolation*

Fifteen seeds of each genotype, obtained from a single panicle collected randomly, were germinated in wet germination paper at the Department of Crop Science of the UFRGS, located in Porto Alegre city, Rio Grande do Sul state, Southern Brazil. Samples were allocated in a BOD (Biochemical Oxygen Demand) chamber at 25 ºC, without photoperiod. After seven days, 100 mg of fresh tissue were collected of each genotype, immediately dry-freeze lyophilized for 48 h and stored at -20 ºC until DNA isolation. Tissue disruption was performed using a TissueLyser II (Qiagen, Maryland, USA) and DNA was extracted using a modified cetyltrimethylammonium bromide (CTAB) method. Sodium chloride and polyvinylpyrrolidone were used to remove polysaccharides and

polyphenols, respectively, following the procedures described by Lodhi *et al.* (1994) and Lefort and Douglas (1999). DNA quality was verified in 1% agarose gel. Total DNA samples were shipped to the University of Minnesota Genomics Center (UMGC), United States, where molecular cloning and sequencing were carried out.

*6.2.5 Tissue collection, RNA isolation, and cDNA synthesis*

Developing grains were collected under field conditions at 7, 14, and 21 days after anthesis (DAA), always between 11 am and 12 pm, with three biological replicates. These three developing stages were determined considering the *CslF6* gene expression in barley (Burton *et al.,* 2008). After collected, samples were immediately kept in liquid nitrogen and, later, stored at -80 ºC. Tissues were manually disrupted in liquid nitrogen and, right after, total messenger RNA (mRNA) was isolated using the reagent Concert$^{TM}$ (Invitrogen, California, USA). The mRNA was quantified using spectrophotometer and diluted in RNAse-free water for a concentration of 125 ng $\mu L^{-1}$. Aiming to remove any DNA residues, 8 µL of concentrated mRNA (1 µg of mRNA) was purified using DNAse I. The complementary DNA (cDNA) synthesis was performed using the reverse transcriptase Moloney Murine Leukemia Virus (MMLV), provided by the SuperScript II kit (Invitrogen, California, USA) for real time PCR. Individual tubes containing 1 µg of purified mRNA received 1 µL of 50 µM oligo-dT mix, 1 µL of 10 mM dNTPs mix, and ultrapure water to complete a final volume of 13 µL. Each tube was incubated for 5 min at 65 ºC in a C1000 Touch™ Thermal Cycler (Bio-Rad, California, USA). After, 4 µL of 5X First-Strand Buffer, 2 µL of 0.1 M DTT, and 0.4 µL of MMLV (200 units $\mu L^{-1}$) were added in each tube. Tubes were incubated for 50 min at 37 ºC, followed by an inactivation step for 15 min at 70 ºC. Finally, mRNA and cDNA quality were verified in 1% agarose gel. Samples of cDNA were shipped to the UMGC for absolute quantification of *CslF6* transcripts.

*6.2.6 Primer design and absolute quantification using droplet digital PCR*

Droplet digital PCR (ddPCR) was performed to determine *CslF6* transcript copy numbers in two different approaches: i) a global expression analysis, amplifying a conserved region among subgenomes; and ii) specific expression analyses amplifying for *CslF6-AA* and *CslF6-CC*. The *CslF6* gene structure and all regions amplified for transcripts quantification are shown in the Figure 1. The specific polymorphisms for the subgenome DD, which would be used for primer development, were found at least 400 bp apart each other. In this sense, considering that ddPCR recommendation is for amplicons ranging from 100 to 200 bp, the expression of the *CslF6-DD* was not performed.

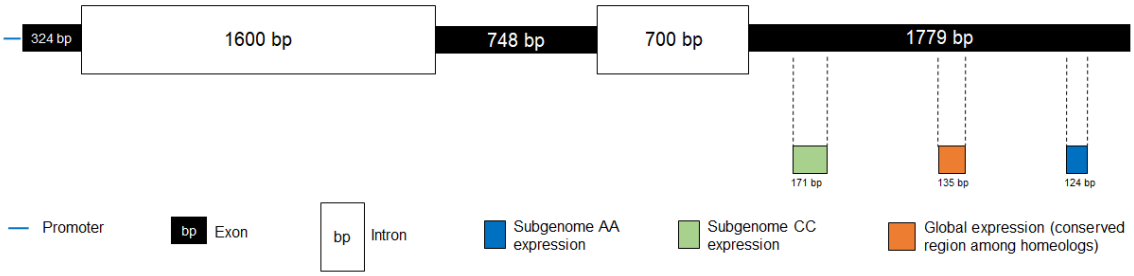



FIGURE 1. *CslF6* gene structure and location of amplified regions for gene expression.

Oligonucleotide sequences used for global expression analysis were developed using the Primer3Plus software (Untergasser *et al.*, 2007), while subgenome-specific oligonucleotide sequences were developed using the GSP software (Wang *et al.*, 2016). Subgenome-specific primers were selected according to the number of variant sites among subgenomes, length (preference for shorter primers) and differences at the 3' end. All oligonucleotide sequences are provided in Table 1.

TABLE 1. Oligonucleotide sequences used for *CslF6* expression analyses.

| Primer pair | Orientation | Sequence (5' - 3') | Amplicon (bp) |
|---|---|---|---|
| *CslF6*-global | Forward | TCACGCTACTGCTCCATCTA | 134 |
| *CslF6*-global | Reverse | AAGAGCGGGTTGTTCTTGG | |
| *CslF6*-AA | Forward | GCTCATCAAGGTGATATTCCA**A** | 124 |
| *CslF6*-AA | Reverse | GATCATCAGCGGTGTCCA**T** | |
| *CslF6*-CC | Forward | CATCTTCTTCGACGGCTC**C** | 171 |
| *CslF6*-CC | Reverse | TACTTGGTCTTGGCGAACA**AT** | |

bp, bases pair. Red nucleotides are subgenome-specific.

Initially, a ddPCR gradient was performed using genomic DNA in order to set the best annealing temperature for each primer pair. After, once all mRNA samples were normalized before cDNA synthesis, a dilution curve was performed to select the best cDNA concentration for absolute quantification. Each ddPCR reaction mix included 10 µL of QX200 ddPCR EvaGreen Supermix (Bio-Rad, California, USA), 3 µL of diluted cDNA (1:15; cDNA:ultrapure water), 0.2 µL of 10 mM forward primer, 0.2 µL of 10 mM reverse primer, and 6.6 µL of ultrapure water, in a total of 20 µL. Droplets were generated in eight-well cartridges using a QX200 droplet generator (Bio-Rad, California, USA). Water-in-oil emulsions were transferred into 96-well plates and amplified in a C1000 Touch Thermal Cycler (Bio-Rad, California, USA). Thermal cycling conditions were 5 min at 95 °C (enzyme activation); 40 cycles of a two-step thermal profile comprising of 30 s at 95 °C (denaturation) and 1 min at 60 °C or 62 °C (annealing/extension, depending of the primer pair). After cycling, samples were incubated for 5 min at 4 °C and for 5 min at 90 °C (stabilization), followed by an 'infinite' hold at 4 °C. Ramp rate of 2 °C per second was used during amplification. Plates were then transferred to the QX200 droplet reader (Bio-Rad, California, USA). Data acquisition and analysis were performed using the QuantaSoft software (Bio-Rad, California, USA) to quantify numbers of positive and negative droplets, with or without the amplicons, respectively. Considering that *CslF6* global expression is the sum of expressions of subgenomes AA, CC, and DD, subgenome DD expression was estimate by difference. Once it is an indirect measure, subgenome DD results are speculative.

### *6.2.7 Sequencing of CslF6 homoeologs*

The *CslF6* gene was partially sequenced using a PacBio sequencer at the UMGC. A primer pair was anchored in a conserved region of *CslF6* homoeologs, covering the third exon of the gene (Figure 1). This part of the gene corresponds to the CslF6 motif, the most important region for the CslF6 coded protein. A pool of six barcodes libraries, one for each

genotype, was created and sequenced for 10 h on a SMRT cell. This sequencing produced more than 9 Gb of sequences with the expected length. The files were demultiplexed and adapters/barcodes were removed for future analyses.

*6.2.8 In silico analyses*

Initially, nucleotide sequences of each genotype were kept apart each other in specific FASTA files. A total of 50 nucleotide sequences of each subgenome were randomly collected for alignments and to create a consensus sequence for each genotype. The identification of each subgenome was performed using specific polymorphisms among them described by Coon (2012) and Erick Jellen (personal communication). Sequences were aligned using Muscle (Edgar, 2004) available on Mega X software (Kumar *et al*., 2018). Later, consensus sequences were generated using BioEdit software (Hall *et al*., 1999). Finally, a complete FASTA file was created with nucleotide sequences of the *CslF6-AA*, *CslF6-CC*, and *CslF6-DD* for each genotype. A new alignment was performed to illustrate nucleotide differences among genotypes and also among subgenomes. Sequences were translated to compare the amino acid sequence of each subgenome. Images were produced using the GeneDoc software (Nicholas, 1997).

*6.2.9 Statistical analyses*

Analyses of variance (ANOVA) were performed separately for phenotypic and expression data. In the first analysis, using phenotypic data, ANOVA was used to identify the effect of genotype and/or planting date on β-glucan content. In the second one, using ddPCR expression data, the objective was to detect the effect of genotype, planting date, and/or developmental stage on *CslF6* transcript levels. Both analyses were complemented by the Tukey test at 5% level, slicing significant interactions if necessary. All statistical analyses were carried out using the SAS 9.4 software (SAS Institute Inc., Cary, NC, USA).

**6.3 Results**

*6.3.1 Phenotypic variation for β-glucan content*

β-glucan content was assessed in five genotypes, using exclusively primary kernels, over two planting dates. The ANOVA showed that β-glucan content varied according to genotype, not being associated neither with planting date nor with genotype by planting date interaction (Table 2). The observed average of β-glucan content among genotypes was 5.30%, considering both planting dates (Table 2).

TABLE 2. Analysis of variance for β-glucan content in five oat genotypes assessed in two planting dates.

| Cause of variation | DF | Sum of squares | Mean square | F value | Pr > F |
|---|---|---|---|---|---|
| Rep(Planting date) | 2 | 0.11 | 0.06 | 0.32 | 0.73 |
| Planting date | 1 | 0.05 | 0.05 | 0.28 | 0.61 |
| Genotype | 4 | 7.00 | 1.75 | 9.96 | 0.01 |
| Genotype*Planting date | 4 | 0.44 | 0.11 | 0.63 | 0.65 |
| Experimental error | 8 | 1.41 | 0.18 | - | - |
| β-glucan mean | 5.3 | - | - | - | - |
| CV (%) | 7.9 | - | - | - | - |

DF, degrees of freedom; Pr, probability; CV, coefficient of variation.

The mean comparison test showed that IAC 7 had higher β-glucan content than IPR Artemis, UFRGS 8, and UFRGS 881971, considering the average of both planting dates, reaching 6.31% (Figure 2). On the other hand, IAC 7 genotype did not differ from URS Taura, which showed 5.57% of β-glucan (Figure 2).

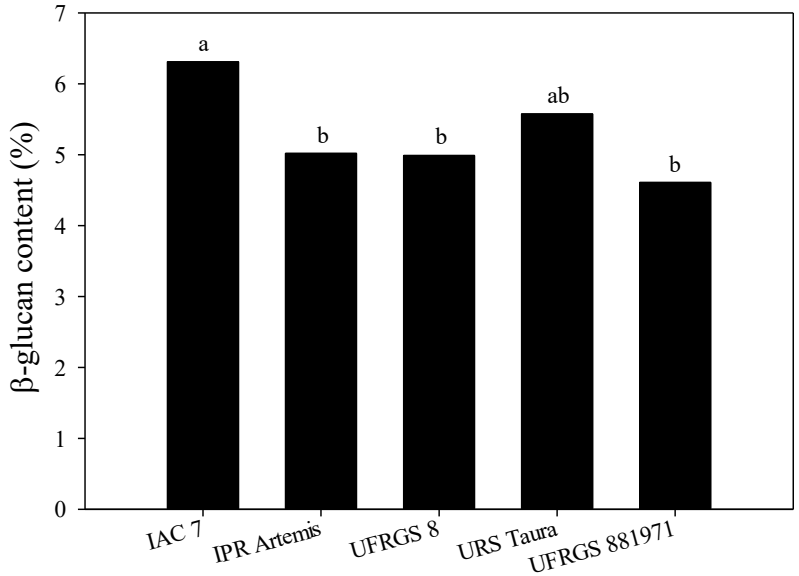



FIGURE 2. Mean comparison for β-glucan content in five oat genotypes using the Tukey test. Same letter between varieties do not differ statistically.

The variety URS Taura did not show any statistical difference for β-glucan content when compared to other genotypes. The genotypes IPR Artemis, UFRGS 8, and UFRGS 881971 presented β-glucan content of 5.02%, 4.99%, and 4.61%, respectively (Figure 2).

*6.3.2 Expression of the CslF6 homoeologs under field conditions*

The ANOVA for *CslF6* transcript copies revealed a significant interaction among genotype, planting date, and kernel developmental stage on *CslF6* expression (Table 3). The interaction among genotype, planting date, and kernel developmental stage was observed in both approaches, either on global or subgenome-specific expression analysis (Table 3).

TABLE 3. Analysis of variance for *CslF6* transcripts copy number under field conditions.

| Cause of variation | DF | Mean square | | | |
|---|---|---|---|---|---|
| | | Global | AA | CC | DD† |
| Rep(Planting date) | 4 | 104.24$^{ns}$ | 20.63$^{ns}$ | 0.61$^{ns}$ | 33.91$^{ns}$ |
| Planting date (PD) | 1 | 5795.08$^{**}$ | 304.96$^{**}$ | 5.85$^{**}$ | 3162.80$^{**}$ |
| Kernel stage (KS) | 2 | 6523.78$^{**}$ | 484.04$^{**}$ | 6.81$^{**}$ | 3388.11$^{**}$ |
| Genotype (G) | 4 | 1261.73$^{**}$ | 51.42$^{**}$ | 3.76$^{**}$ | 766.62$^{**}$ |
| PD*KS | 2 | 4427.18$^{**}$ | 380.37$^{**}$ | 4.22$^{**}$ | 2024.38$^{**}$ |
| G*PD | 4 | 1513.38$^{**}$ | 56.57$^{**}$ | 0.74$^{ns}$ | 1026.49$^{**}$ |
| G*KS | 8 | 651.46$^{**}$ | 17.03$^{ns}$ | 1.18$^{*}$ | 439.55$^{**}$ |
| G*PD*KS | 8 | 1382.70$^{**}$ | 62.89$^{**}$ | 1.52$^{**}$ | 858.85$^{**}$ |
| Experimental error | 40 | 107.94 | 11.41 | 0.50 | 53.61 |
| ‡G | - | 5.80 | 3.70 | 14.93 | 6.52 |
| ‡PD + KS + (PD*KS) | - | 76.93 | 84.17 | 67.01 | 72.95 |

DF, degrees of freedom; †, number of copies obtained by difference (DD copies = Global – AA – CC); ‡, % of total mean square; $^{*}$, significant at 5% level by the F test; $^{**}$, significant at 1% level by the F test; ns, non-significant.

The percentage of total mean square value was calculated for genotype and planting date, kernel stage, and their interaction (Table 3). This result is important to understand the environmental influence on *CslF6* expression under field conditions. The sum of PD, KS, and PD*KS mean squares contributed with 77%, 84%, 67%, and 73% of the total variance for global, *CslF6-AA*, *CslF6-CC*, and *CslF6-DD* expressions, respectively (Table 3). The genotype contributed with only 6%, 4%, and 6.5% of the total variance for global, *CslF6-*

*AA*, and *CslF6-DD* expressions, respectively (Table 3). The influence of genotype on *CslF6-CC* was higher (15%) when compared to the other expression approaches (Table 3).

The results of the *CslF6* gene expression are presented for each planting date separately, where gene expression is compared: i) among genotypes inside each kernel developmental stage; and ii) among kernel developmental stages inside the same genotype (Figures 3 and 4). After, *CslF6* expression is compared between planting dates for each kernel developmental stage of the same genotype (Figure 5). Results will be presented initially for the *CslF6* global expression and then for subgenome-specific expression.

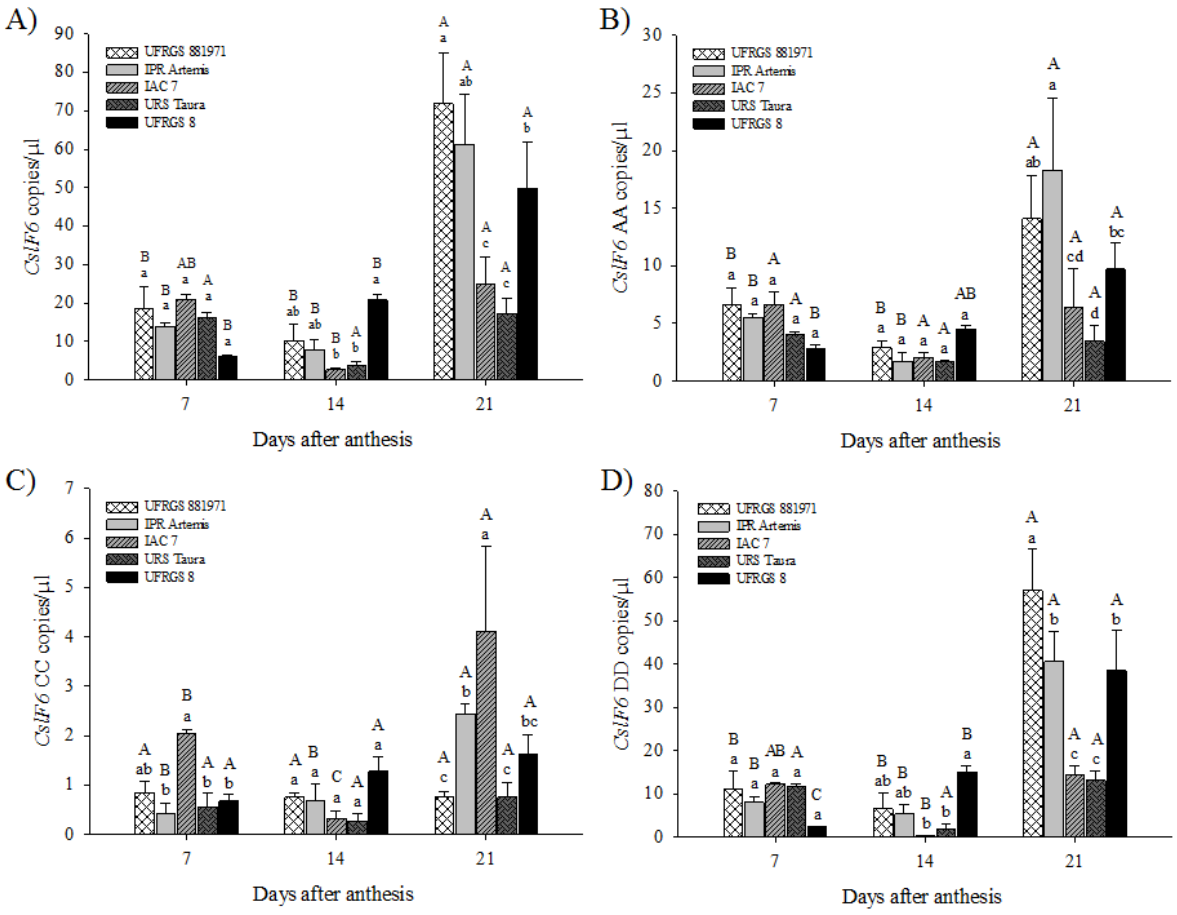



FIGURE 3. Expression of the *CslF6* gene on the first planting date. A, global expression; B, subgenome AA expression; C, subgenome CC expression; and D, subgenome DD expression (estimate). Genotypes that share a lowercase letter inside the same kernel developmental stage do not differ statistically. Kernel developmental stages that share a capital letter within genotype do not differ statistically. Error bars represent the standard error.

Considering the *CslF6* global expression on the first planting date, significant differences among genotypes were identified only at 14 and 21 DAA (Figure 3A). The

genotype UFRGS 8 showed higher number of *CslF6* transcripts than IAC 7 and URS Taura at 14 DAA but did not differ from UFRGS 881971 and IPR Artemis (Figure 3A). The genotype UFRGS 881971 showed higher *CslF6* expression than UFRGS 8, IAC 7, and URS Taura at 21 DAA, while no differences were detected when compared to IPR Artemis (Figura 3A). The genotype IPR Artemis also did not show any statistical difference when compared to UFRGS 8, but both genotypes presented higher *CslF6* transcripts than IAC 7 and URS Taura at 21 DAA (Figure 3A).

Three global expression patterns were identified comparing the *CslF6* expression among kernel developmental stages inside genotype. The most common expression pattern was observed for the UFRGS 881971, IPR Artemis, and UFRGS 8 genotypes, which showed highest *CslF6* expression at 21 DAA (Figure 3A). For the IAC 7 genotype, *CslF6* expression at 21 DAA was higher than 14 DAA but did not vary when compared to 7 DAA. The *CslF6* gene presented the same expression during kernel filling on the URS Taura genotype (Figure 3A).

Differences for the *CslF6* AA homoeologs (*CslF6-AA*) expression on the first planting date were detected only at 21 DAA (Figure 3B). The genotype IPR Artemis presented higher number of transcripts than IAC 7, URS Taura, and UFRGS at 21 DAA, while no differences were detected when compared to UFRGS 881971 (Figure 3B). The genotype UFRGS 881971 showed higher *CslF6-AA* expression than IAC 7 and URS Taura, which presented the lowest expression and did not vary between them, while no differences were detected when compared to UFRGS 8 (Figure 3B). The genotype UFRGS 8, which has intermediate *CslF6-AA* expression at 21 DAA, showed higher expression than URS Taura, while no differences were detected when compared to IAC 7 (Figure 3B). Considering the *CslF6-AA* expression among the assessed developmental stages within each genotype, three expression profiles were identified. The genotypes UFRGS 881971 and IPR Artemis showed highest *CslF6-AA* expression at 21 DAA, as presented previously

for the global expression (Figure 3B). The genotypes IAC 7 and URS Taura showed a similar expression during kernel filling, not varying among the assessed kernel developmental stages (Figure 3B). The *CslF6-AA* expression on the UFRGS 8 genotype at 21 DAA was higher than at 7 DAA, while no differences were identified when compared to 14 DAA. This result indicates a progressive increasing of the *CslF6-AA* expression during kernel filling on the UFRGS 8 genotype, reaching a peak at 21 DAA (Figure 3B).

The *CslF6* CC homoeologs (*CslF6-CC*) expression was reduced when compared to other subgenomes (Figure 3C). Differences among genotypes for *CslF6-CC* expression were identified at 7 and 21 DAA (Figure 3C). At 7 DAA, the genotype IAC 7 showed higher *CslF6-CC* expression than IPR Artemis, URS Taura, and UFRGS 8. Also, at 7 DAA, the genotype UFRGS 881971, which presented intermediate expression for *CslF6-CC*, did not vary neither from IAC 7 nor all other genotypes (Figure 3C). Considering the *CslF6-CC* expression among kernel developmental stages, two expression patterns were observed. The genotypes UFRGS 881971, URS Taura, and UFRGS 8 showed the same expression level among all kernel developmental stages, while IPR Artemis and IAC 7 showed highest expression at 21 DAA (Figure 3C).

The *CslF6* DD homoeologs (*CslF6-DD*) expression, which was estimated based on the differences between the global expression and *CslF6-AA* and *CslF6-CC* subgenomes expression is presented in the Figure 3D. Significant differences among genotypes were identified only at 14 and 21 DAA (Figure 3D), as observed for the global expression (Figure 3A). The genotype UFRGS 8 showed higher *CslF6-DD* expression than URS Taura and IAC 7 at 14 DAA, while no differences were verified when compared to IPR Artemis and UFRGS 881971 in this same kernel developmental stage (Figure 3D). The genotypes IPR Artemis and UFRGS 881971 did not show any statistical difference neither with UFRGS 8 nor URS Taura and IAC 7 at 14 DAA (Figure 3D). The genotype UFRGS 881971 showed higher *CslF6-DD* expression at 21 DAA than all other ones (Figure 3D).

The genotypes UFRGS 881971 and IPR Artemis showed higher expression than URS Taura and IAC 7, which presented the lowest *CslF6-DD* expression at 21 DAA (Figure 3D). Different expression patterns were observed for the *CslF6-DD* expression among kernel developmental stages. The genotypes UFRGS 881971 and IPR Artemis showed highest expression at 21 DAA and did not present significant differences between 7 and 14 DAA (Figure 3D). The genotype UFRGS 8 presented a similar behavior, reaching a peak of expression at 21 DAA, but with differences between 7 and 14 DAA (Figure 3D). This genotype showed a progressive increasing of *CslF6-DD* expression from 7 to 14 DAA. Similarly, the genotype IAC 7 presented highest expression at 21 DAA, being that no differences were observed when compared to 7 DAA (Figure 3D).

The environmental condition verified on the second planting date changed the *CslF6* expression (Table 3). Significant differences among genotypes were observed in all kernel developmental stages assessed on the second planting date, either for global or subgenome specific expression analyses (Figure 4). Considering the *CslF6* global expression on the second planting date, the genotype UFRGS 8 showed higher expression at 7 DAA than all other ones (Figure 4A). The genotype URS Taura presented higher global expression than UFRGS 881971 at 7 DAA, while no differences were observed when compared to IPR Artemis and IAC 7 in this same kernel developmental stage. The genotypes UFRGS 881971, IPR Artemis, and IAC 7 did not differ among them at 7 DAA (Figure 4A). For kernels collected at 14 DAA on the second planting date, the genotypes IPR Artemis and UFRGS 8 presented higher *CslF6* global expression than UFRGS 881971 and URS Taura, which showed low number of transcripts. However, differences were not observed when IPR Artemis and UFRGS 8 were compared to IAC 7 at 14 DAA (Figure 4A). The genotype IAC 7 showed higher *CslF6* global expression than all other ones at 21 DAA (Figure 4A). The genotypes showed contrasting expression pattern among kernel developmental stages for the *CslF6* global expression on the second planting date. The

genotype UFRGS 881971 showed highest *CslF6* expression at 7 and 21 DAA, while IPR Artemis presented highest expression at 7 and 14 DAA (Figure 4A). The genotypes URS Taura and UFRGS 8 showed highest expression at 7 DAA on the second planting date (Figure 4A). Conversely, only IAC 7 showed highest *CslF6* exclusively at 21 DAA on the second planting date (Figure 4A).

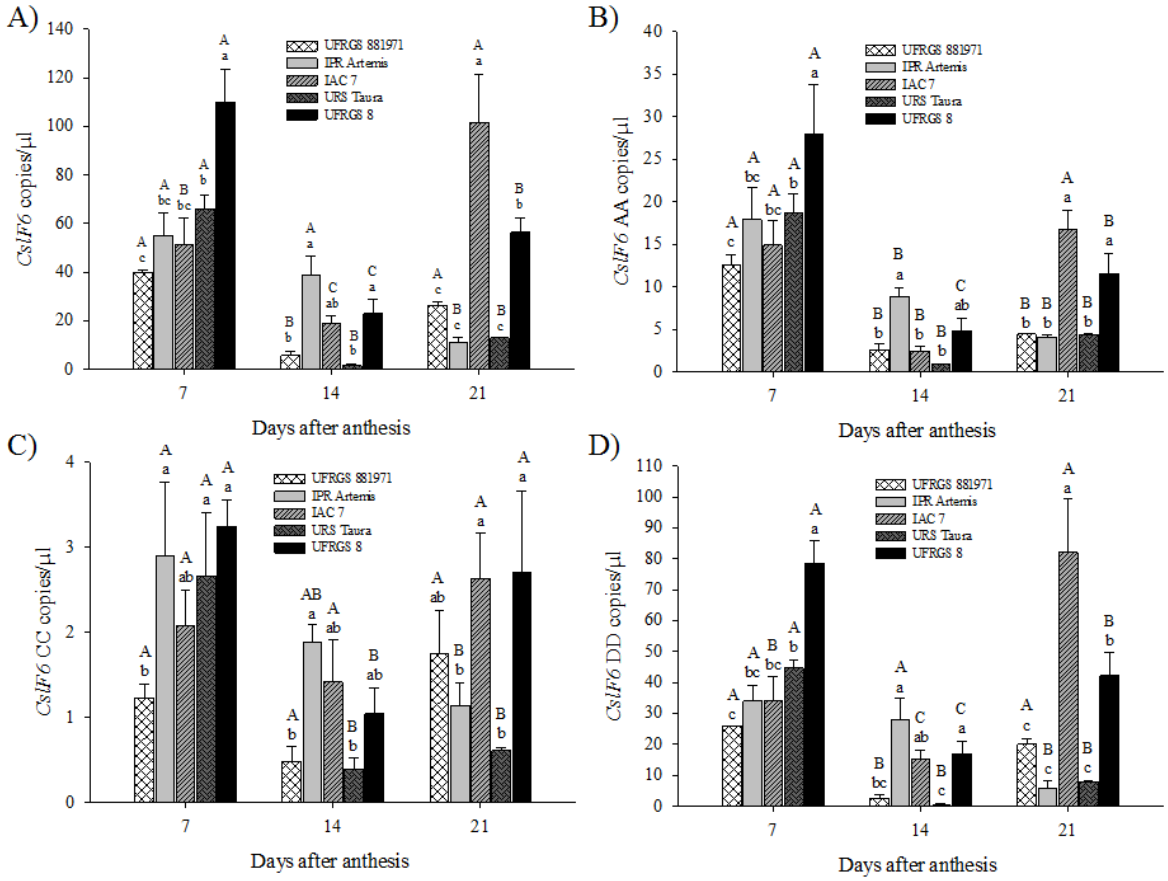



FIGURE 4. Expression of the *CslF6* gene on the second planting date. A, global expression; B, subgenome AA expression; C, subgenome CC expression; and D, subgenome DD expression (estimate). Genotypes that share a lowercase letter inside the same kernel developmental stage did not differ statistically. Kernel developmental stages that share a capital letter within genotype did not differ statistically. Error bars represent the standard error.

The *CslF6-AA* expression (Figure 4B) was highly similar when compared to global expression (Figure 4A), especially at 7 and 14 DAA. Considering the CslF6-AA expression at 7 DAA, on the second planting date, the genotype UFRGS 8 showed higher expression than all other ones, which presented intermediate expression (Figure 4B). Considering the *CslF6-AA* expression at 14 DAA on the second planting date, the genotype IPR Artemis showed higher expression than UFRGS 881971, IAC 7, and URS

Taura, while statistical differences were not verified when compared to UFRGS 8 (Figure 4B). UFRGS 8, which showed an intermediate *CslF6-AA* expression at 14 DAA on the second planting date, also did not differ from UFRGS 881971, IAC 7, and URS Taura (Figure 4B). The genotypes IAC 7 and UFRGS 8 showed higher *CslF6-AA* expression than all other genotypes at 21 DAA on the second planting date and did not differ between them (Figure 4B). The genotypes UFRGS 881971, IPR Artemis, and URS Taura presented basically the same *CslF6-AA* expression level at 21 DAA, not differing statistically among them (Figure 4B).

The *CslF6-CC* expression was reduced on the second planting date, as well as observed on the first one, when compared to other genomes. Considering the *CslF6-CC* expression on the second planting date, the genotypes IPR Artemis, URS Taura, and UFRGS 8 showed higher transcript number than UFRGS 881971 at 7 DAA (Figure 4C). However, differences between these genotypes and IAC 7 were not significant in this same kernel developmental stage (Figure 4C). At 14 DAA, the genotype IPR Artemis presented higher *CslF6-CC* expression than UFRGS 881971 and URS Taura, while no differences were observed in comparison to IAC 7 and UFRGS 8 (Figure 4C). The genotypes IAC 7 and UFRGS 8, which did not differ between them for *CslF6-CC* expression at 21 DAA, showed higher expression than IPR Artemis and URS Taura (Figure 4C). In this same kernel developmental stage, the genotype UFRGS 881971 showed intermediate *CslF6-CC* expression, differing neither from IAC 7 and UFRGS 8 nor from IPR Artemis and URS Taura (Figure 4C). The genotypes presented different patterns for the *CslF6-CC* expression among kernel developmental stages on the second planting date. Differences of expression of the *CslF6-CC* gene were not observed in UFRGS 881971 and IAC 7 among the assessed kernel developmental stages, indicating that these varieties presented a constant expression during kernel filling (Figure 4C). The genotype URS Taura showed highest *CslF6-CC* expression at 7 DAA on the second planting date (Figure 4C). The

genotype IPR Artemis also showed higher *CslF6-CC* expression at 7 DAA, being that no differences were observed between the expression on 7 and 14 DAA in this genotype (Figure 4C). The genotype UFRGS 8 showed high expression of *CslF6-CC* at 7 and 21 DAA on the second planting date (Figure 4C).

The expression results for the *CslF6-DD* on the second planting date were similar to the global expression in this same planting date (Figure 4A and 4D). At 7 DAA, the genotype UFRGS 8 showed higher *CslF6-DD* expression than all other genotypes on the second planting date (Figure 4D). On the second planting date, at 14 DAA, the genotypes IPR Artemis and UFRGS 8 presented higher *CslF6-DD* expression than UFRGS 881971 and URS Taura (Figure 4D). On the other hand, no differences where observed when IPR Artemis and UFRGS 8 were compared to IAC 7 in this same kernel developmental stage (Figure 4D). The genotype IAC 7, which has great β-glucan content, showed higher *CslF6-DD* expression than all other genotypes at 21 DAA (Figure 4D). Considering the *CslF6-DD* expression among kernel developmental stages on the second planting date, four of the five genotypes showed highest expression at 7 DAA (Figure 4D). The genotypes URS Taura and UFRGS 8 showed highest *CslF6-DD* expression at 7 DAA, while the expression of the gene at 7 DAA did not differ when compared to 14 and 21 DAA for IPR Artemis and UFRGS 881971, respectively (Figure 4D).

One of the main objectives of this study was to detect the environmental influence on *CslF6* expression in oats. In this sense, the expression of the *CslF6* gene was assessed and compared on two planting dates, which are characterizing two different environmental conditions (Figure 5). In a general way, the *CslF6* global expression varied according to the genotype on the first planting date. On the second planting date, an important increasing on *CslF6* transcripts was observed at 7 DAA in both global and genome specific approaches (Figure 5).

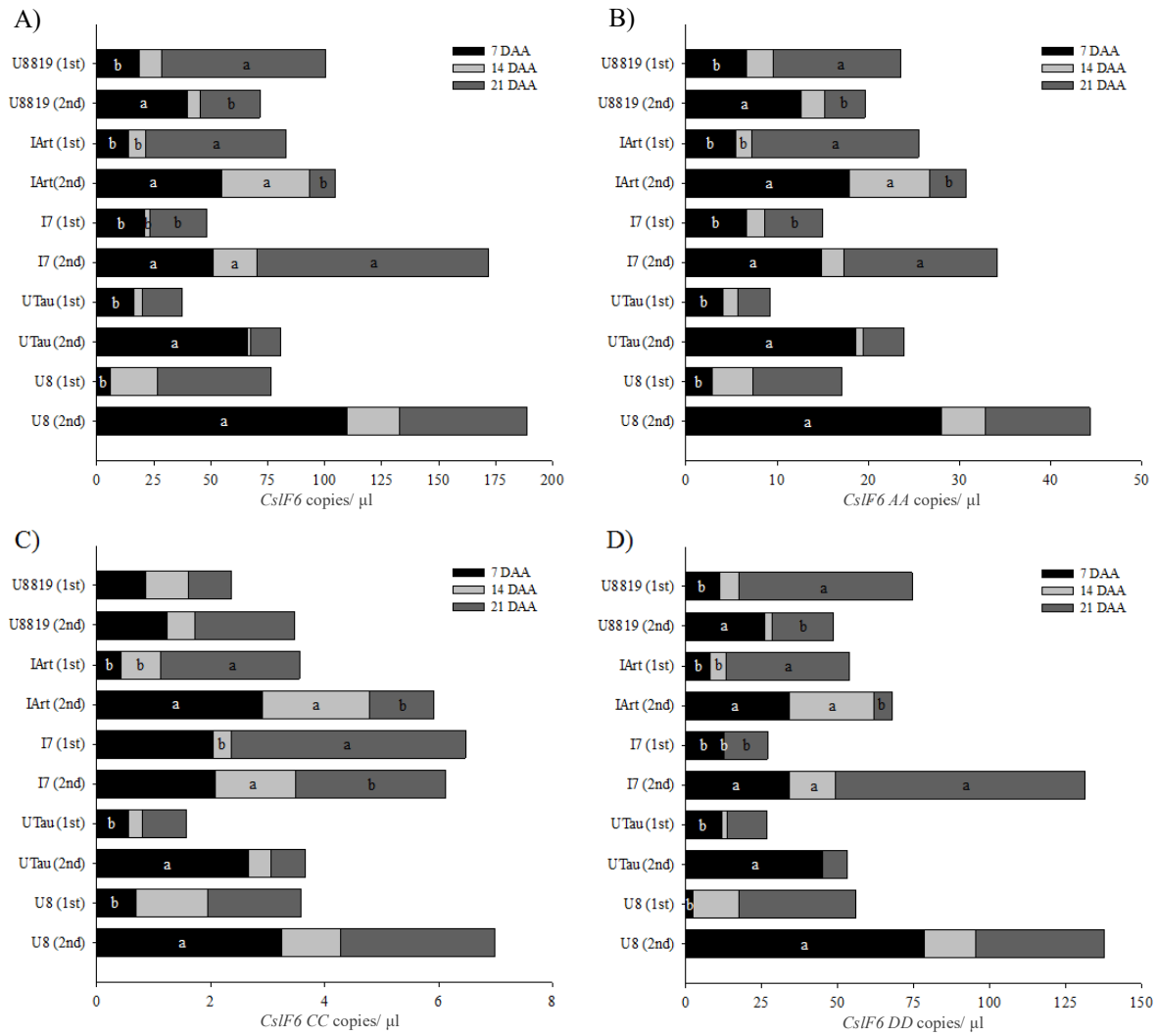



FIGURE 5. Expression of the *CslF6* gene between planting dates. A, global expression; B, subgenome AA expression; C, subgenome CC expression; and D, subgenome DD expression (estimate). Lowercase letters between planting dates for the same kernel developmental stage and genotype did not differ statistically. U8819, UFRGS 881971; IArt, IPR Artemis; I7, IAC 7; UTau, URS Taura; U8, UFRGS 8. 1st, first planting date; and 2nd, second planting date.

Considering the CslF6 global expression in each genotype, inside each kernel developmental stage and between planting dates, all assessed genotypes increased the transcript number at 7 DAA in response to the second planting date (Figure 5A). The genotypes IPR Artemis and IAC 7 showed significant differences between planting dates at 14 DAA, increasing the CslF6 expression on the second planting date, while all other ones presented the same expression level at this kernel developmental stage (Figure 5A). Three expression patterns were observed for the CslF6 global expression between planting dates at 21 DAA. The genotype UFRGS 881971 and IPR Artemis reduced the CslF6 transcript number at 21 DAA in response to the second planting date, while URS Taura

and UFRGS showed the same transcript level between planting dates (Figure 5A). On the other hand, the genotype IAC 7, which showed the highest β-glucan content, was the only one that increased the transcript number at 21 DAA in response to the second planting date (Figure 5A).

As expected, subgenome expression changes in response to different planting dates were similar when compared to the *CslF6* global expression. For the *CslF6-AA* expression, with the exception of IAC 7 at 14 DAA, all observed expression changes due to planting date were same ones verified on the global expression (Figure 5A and Figure 5B). For the *CslF6-CC*, some genotypes showed contrasting behavior when compared to *CslF6-AA* and global expression (Figure 5C). For example, the genotype UFRGS 881971 did not show any difference on the transcript number in response to planting date change, regardless of the kernel developmental stage (Figure 5C). For the IAC 7 genotype, the *CslF6-CC* expression increased at 14 DAA in response to the second planting date (Figure 5C), as verified on the global expression (Figure 5A), while the expression at 21 DAA was reduced from the first planting date to the second one, differing from both global and *CslF6-AA* expressions (Figure 5C). The genotypes IPR Artemis, URS Taura, and UFRGS 8 showed the same behavior as observed for global and *CslF6-AA* expressions (Figure 5C).

*6.3.3 Molecular characterization of the CslF6 homoeologs*

Nucleotide sequences of the *CslF6* homoeologs are presented as supplementary material. The CslF6 motif, which corresponds to the third exon of the *CslF6* gene, has 1767 bases pair (bp) in oats. This number did not vary among *CslF6-AA*, *CslF6-CC*, and *CslF6-DD* (Supplementary Figure 2, 3, and 4). The *CslF6-AA* and *CslF6-DD* showed high nucleotide similarity, with only 4 SNP between them in the third exon (Supplementary Figure 2, 3, and 4). These SNP are located at the positions 102, 1110, 1341, and 1425 (Supplementary Figure 2, 3, and 4). The *CslF6-CC* showed many SNP when compared to *CslF6-AA* and *CslF6-DD* ones (Supplementary Figure 2, 3, and 4). The comparison

between *CslF6-AA* and *CslF6-CC* revealed 49 SNP, which are located at the positions 7, 36, 63, 102, 109, 195, 207, 246, 297, 309, 333, 387, 498, 519, 520, 541, 565, 583, 585, 642, 654, 674, 690, 697, 698, 709, 710, 712, 724, 729, 741, 749, 756, 765, 939, 963, 1065, 1110, 1125, 1260, 1341, 1401, 1425, 1437, 1614, 1726, 1728, and 1763 (Supplementary Figure 2, 3, and 4). A total of 45 SNP was observed comparing *CslF6-CC* and *CslF6-DD* subgenomes. These SNP correspond to the same ones mentioned previously in the *CslF6-AA* versus *CslF6-CC* alignment, except those ones at the positions 102, 1110, 1341, and 1425. The alignment of nucleotide sequences did not show any SNP among genotypes, regardless of the subgenome (Supplementary Figure 2, 3, and 4).

The nucleotide sequences of each subgenome were translated to amino acid sequences and compared among them (Figure 6). The *CslF6-AA* and *CslF6-DD* showed the same amino acid sequences (Figure 6), revealing that the SNP identified between them are synonymous. On the other hand, the amino acid sequence from the *CslF6-CC* showed 13 amino acid changes when compared to *CslF6-AA* and *CslF6-DD* ones (Figure 6). These amino acid changes are located at the positions 3, 37, 174, 181, 189, 195, 225, 233, 237, 238, 242, 250, and 576 (Figure 6). Three of these amino acid changes, located at the positions 3, 37, and 242, have the same functional property, showing the same color in the Figure 6. The other ones changed the functional property of the amino acid, as observed by different colors in the Figure 6.

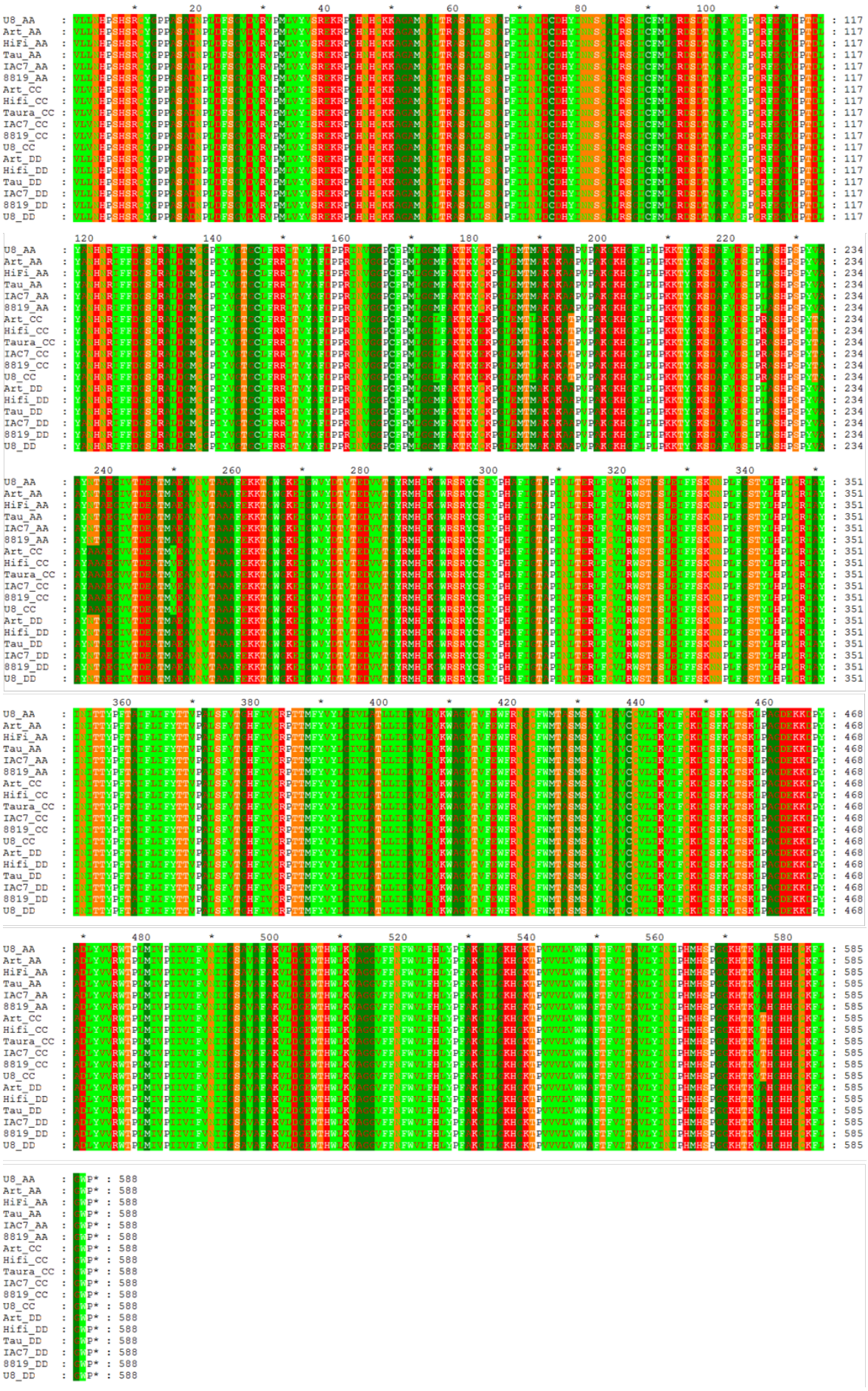



FIGURE 6. Amino acid composition of the CslF6 motif in oats.

### 6.4 Discussion

The ANOVA revealed that β-glucan content was influenced by the genotype, but not by planting date or planting date by genotype interaction (Table 2). This result is contrasting with some studies that detected a strong influence of planting date on oat β-glucan content, such as Humphreys *et al.* (1994). It can be justified by the similar conditions observed for both planting dates (Supplementary Figure 1). A mildly limiting condition was observed on the first planting date, with low rainfall during the vegetative stage (Supplementary Figure 1A). In principle, this planting date should represent a favorable environment to contrast with the second one, less favorable, planted one month later. As consequence, both planting dates could be classed as unfavorable condition, with water restriction during the vegetative stage. In addition, higher daily maximum temperatures were observed after anthesis on the first planting date.

The variety IAC 7 showed higher β-glucan content than UFRGS 8, UFRGS 881971, and IPR Artemis (Figure 2). However, differences between IAC 7 and URS Taura were not identified (Figure 2). In this sense, these results confirm that IAC 7 has higher β-glucan content, reaching 6.31%. This variety has been reported as one of the best sources to increase β-glucan content in subtropical conditions (Crestani *et al*., 2012). URS Taura is a commercial oat variety, highly appreciated by Brazilian farmers due to its agronomic type, which include reduced plant height, short cycle, lodging resistance, high grain yield, and high milling yield. In the present study, a high β-glucan content was observed for URS Taura (5.57%), approximately 1.2% more than the oat average, indicating that this variety is also a great donor for this trait.

The *CslF6* gene expression was highly influenced by the environmental conditions, as observed by the ANOVA of the transcript number (Table 3). The expression of the *CslF6* gene depends of the triple interaction among genotype, planting date, and kernel developmental stage (Table 3). The results revealed that genotype has a minor influence on

the *CslF6* gene expression. The genotype accounted between 4 and 6% of the total variance for global, *CslF6-AA*, and *CslF6-DD* expression, and 15% of the *CslF6-CC* total variance (Table 3). On the other hand, the sum of planting date, kernel stage and their interaction accounted for 77%, 84%, 67%, and 73% of the total variance for global, *CslF6-AA*, *CslF6-CC*, and *CslF6-DD* expressions, respectively (Table 3). The environmental influence on *CslF6* expression is easily detected observing the gene expression on the first planting date. Except UFRGS 8, all genotypes reduced the expression of the *CslF6* gene from 7 to 14 DAA. During these days, high rainfall and low temperatures were observed (Supplementary Figure 1A). In this sense, the UFRGS 8 *CslF6* expression was not influenced by these conditions due to its early flowering, not coinciding with the other genotypes (Supplementary Figure 1A).

The *CslF6* oat gene reached an expression peak at 21 DAA on the first planting date, regardless of the subgenome (Figure 3). Similar results were observed for the *CslF6* gene in barley and durum wheat (Burton *et al*., 2008; Ilaria *et al*., 2018). In the barley variety 'Himalaya', which has high β-glucan content, the *CslF6* gene increased his expression between 12 and 20 DAA, reaching a peak of expression at 20 DAA (Burton *et al.*, 2008). Similarly, a peak of *CslF6* transcription was observed at 21 DAA in durum wheat genotypes (Ilaria *et al*., 2018). The genotypes UFRGS 881971, IPR Artemis, and UFRGS 8 showed higher *CslF6* transcripts than IAC 7 and URS Taura, considering the global expression, *CslF6-AA*, and *CslF6-DD* subgenomes (Figure 3). However, the higher expression of the *CslF6* gene was not associated with high β-glucan content averages on the first planting date. These genotypes showed lower β-glucan content than IAC 7, which presented lower expression on the global, *CslF6-AA*, and *CslF6-DD* approaches at 21 DAA (Figure 2 and Figure 3). In this sense, the high *CslF6* expression is not a guarantee of high β-glucan content in oats. Similar results were published in wheat, a species that has low β-glucan content, indicating that the abundance of *CslF6* transcripts was not

necessarily be a good indicator of enzyme activity and/or β-glucan deposition in cell wall (Ilaria *et al*., 2018). The molecular mechanisms of β-glucan biosynthesis are complex, requiring the involvement of different genes/enzymes that act coordinately for their deposition and accumulation (Houston *et al.,* 2014).

The IAC 7 genotype showed higher *CslF6-CC* expression than the other ones on the first planting date (Figure 3), suggesting a possible involvement of this subgenome on the β-glucan content. The subgenome C showed lower expression than the A and D ones, regardless of the planting date (Figure 3 and Figure 4). Similar results were verified by Coon (2012) studying a set of diploid and hexaploid oat genotypes.

The *CslF6* gene expression was completely different on the second planting date showing and increasing at 7 DAA in most of the assessed genotypes (Figure 4). These increasing was observed for the global, subgenome A, and subgenome D quantifications, while the subgenome C showed similar transcript level at 7 and 21 DAA for some genotypes (Figure 4C). The IAC 7 genotype, which has high β-glucan content, showed high *CslF6* expression for the global, *CslF6-AA*, and *CslF6-DD* approaches (Figure 4), suggesting the involvement of this gene for the β-glucan formation in this variety on the second planting date. On the other hand, the UFRGS 8 variety showed high and regular *CslF6* expression at 7 and 21 DAA, respectively, on the second planting date, and produced intermediate β-glucan content. In this context, the action of the *CslF6 per se* is not enough to increase the β-glucan content level in oats.

The *CslF6* gene expression was compared between planting dates to understand the environmental influence on each genotype and kernel developmental stage. Except UFRGS 881971 and IAC 7 for *CslF6-CC* expression, the second planting date increased the *CslF6* gene expression at 7 DAA (Figure 5). The genotypes IAC 7 and UFRGS 8 showed similar number of *CslF6* transcripts on the second planting date and contrasting β-glucan content (Figure 5A). However, the expression on each kernel developmental stage

was different. The IAC 7 increased the *CslF6* global expression in all kernel developmental stages on the second planting date compared to the first one, while UFRGS showed a huge increase at 7 DAA and did not vary at 14 and 21 DAA (Figure 5A). These results suggest that the CslF6 gene expression is not associated with high β-glucan formation when expressed at 7 DAA. The IAC 7 variety also increased the *CslF6-AA* and *CslF6-DD* at 21 DAA on the second planting date, indicating that the expression of the gene is important during the last third of the kernel filling in genotypes with high β-glucan content.

The partial characterization of the *CslF6* homoeologs did not reveal any SNP among the assessed genotypes. In this sense, all genotypes share the same CslF6 motif for the A, C, and D subgenomes. Although the *CslF6* gene was not completely sequenced, the high β-glucan content in some genotypes does not seem to be associated with any amino acid change of the CslF6 enzymes. The characterization of the *CslF6* gene was performed in oats by Coon (2012). However, this study focused exclusively on diploid ancestral species and hexaploid lines not adapted to subtropical conditions. In addition, the nucleotide sequences obtained by Coon (2012) are not available to date.

Studies involving the characterization of the *CslF6* gene in oats are scarce in the literature. Scientific papers were not published to date. In the present study, the *CslF6* gene expression was characterized under field conditions, in two planting dates, to identify the environmental influence on this gene. In addition, the most important region of the *CslF6* gene, corresponding to the CslF6 motif, was sequenced and characterized for each subgenome. The CslF6 motifs encoded by the *CslF6-AA* and *CslF6-DD* subgenomes share the same amino acid sequence, while 13 amino acid changes were observed between *CslF6-CC* and the other ones. Similar results were observed by Coon (2012). Further studies focusing on the promoter region of the *CslF6* homoeologs may elucidate the

expression profiles observed in the present study and clarify the association between *CslF6* expression and β-glucan content in oats.

### 6.5 Conclusions

The high *CslF6* gene expression is not a guarantee of high β-glucan content in oats. The *CslF6* gene expression is highly influenced by the environmental conditions and kernel developmental stage, and in minor degree by the genotype. The CslF6 motif has 1767 bp in oats, regardless of the subgenome. The third exon of the *CslF6-AA* and *CslF6-DD* subgenomes encodes the same amino acid sequence. The *CslF6-CC* subgenome has 13 amino acid changes in the CslF6 motif when compared to the other ones. SNP were not identified for the *CslF6* homoeologs among the assessed genotypes.

### 6.6 References

AACC - American Association of Cereal Chemists. Approved methods, 9 ed. Saint Paul: AACC, 1999.

AOAC - Association of Official Analytical Chemistry. Official methods of analysis of the Association of Official Analytical Chemistry, 16 ed. Washington: AOAC, 1997.

BURTON, R. A. *et al.* Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)-b-D-glucans. **Science**, v. 311, p. 1940-1942, 2006.

BURTON, R. A. *et al.* The genetics and transcriptional profiles of the Cellulose Synthase-Like HvCslF gene family in barley. **Plant Physiology**, v. 146, p. 1821-1833, 2008.

BURTON, R. A. *et al.* Overexpression of specific HvCslF cellulose synthase-like genes in transgenic barley increases the levels of cell wall (1,3;1,4)-β-D-glucans and alters their fine structure. **Plant Biotechnology Journal**, v. 9, p. 117-135, 2011.

CHAWADE, A. *et al.* Development and characterization of an oat TILLING-population and identification of mutations in lignin and beta-glucan biosynthesis genes. **BMC Plant Biology**, v. 10, p. 1-13, 2010.

CHO, S. S. *et al*. Consumption of cereal fiber, mixtures of whole grains and bran, and whole grains and risk reduction in type 2 diabetes, obesity, and cardiovascular disease. **American Journal of Clinical Nutrition**, v. 98, p. 594-619, 2013.

CRESTANI, M. *et al.* Combining ability for grain chemistry quality traits in a white oat diallelic cross. **Euphytica**, v. 185, p. 139-156, 2012.

COON, M. A. Characterization and variable expression of the *CslF6* homologs in oat (*Avena* sp.). **Thesis (Master)** - Department of Plant and Wildlife Sciences, Brigham Young University (BYU), Utah, f. 47, 2012.

COUTINHO, P. M.; HENRISSAT, B. Carbohydrate-active enzymes: an integrated database approach. In: GILBERT, H. J.; DAVIES, G. J.; HENRISSAT, B.; SVENSSON, B. (Eds.). Recent advances in carbohydrate bioengineering: Carbohydrate-active enzymes: an integrated database approach. The Royal Society of Chemistry, Cambridge, 1999. p. 3-12.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32, p. 1792-1797, 2004.

HALL, T. A. *et al.* BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: Nucleic acids symposium series. London: Information Retrieval Ltd., c1979-c2000., 1999. p. 95-98.

HAZEN, S. *et al.* Cellulose synthase-like genes of rice. **Plant Physiology**, v. 128, p. 336-340, 2002.

HOLTHAUS, J. F. *et al.* Inheritance of β-glucan content of oat grain. **Crop Science**, v.36, p. 567-572, 1996.

HOUSTON, K. *et al.* 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**, v. 15, p. 907, 2014.

HUMPHREYS, D. G. *et al*. Nitrogen fertilizer and seeding date induced charges in protein, oil and β-glucan contents of four oat cultivars. **Journal of Cereal Science**, v. 20, p. 283-290, 1994.

ILARIA, M. *et al.* Expression analysis of cellulose synthase-like genes in durum wheat. **Scientific Reports**, v. 8, p. 15675, 2018.

KUMAR, S. *et al.* MEGA X: molecular evolutionary genetics analysis across computing platforms. **Molecular Biology and Evolution**, v. 35, p. 1547-1549, 2018.

JOBLING, S. A. Membrane pore architecture of the CslF6 protein controls (1-3,1-4)-β-glucan structure. **Science Advances**, v.1, p.e.e1500069, p. 1-9, 2015.

LODHI, M. A. *et al.* A simple and efficient method for DNA extraction from grapevine cultivars and Vitis species. **Plant Molecular Biology Reporter**, v. 12, p. 6-13, 1994.

LEFORT, F.; DOUGLAS, C. An efficient micro-method of DNA isolation from mature leaves of four hardwood tree species Acer, Fraxinus, Prunus and Quercus. **Annals of Forest Science**, v. 56, p. 259-263, 1999.

MCCLEARY, B. V.; CODD, R. Measurement of (1→3),(1→4)-β-D-Glucan in Barley and Oats: A Streamlined Enzymic Procedure. **Journal of the Science of Food and Agriculture**, v. 55, p. 303-312, 1991.

MCMULLEN M. S. *et al.* Registration of 'HiFi' oat. **Crop Science**, v. 45, p. 1664, 2005.

NEMETH, C. *et al.* Down-regulation of the *CslF6* gene results in decreased (1,3;1,4)-β-D-glucan in endosperm of wheat. **Plant Physiology**, v. 152, p. 1209-1218, 2010.

NICHOLAS, K. B. Genedoc: a tool for editing and annoting multiple sequence alignments. Available online at http://wwwpscedu/biomed/genedoc, 1997.

RICHMOND, T.; SOMERVILLE, C. The cellulose synthase superfamily. **Plant Physiology**, v. 124, p. 495-498, 2000.

SOMERVILLE, C. *et al.* Toward a systems approach to understanding plant cell walls. **Science**, v. 306, p. 2206-2211, 2004.

TAKETA, S. *et al.* Functional characterization of barley betaglucanless mutants demonstrates a unique role for *CslF6* in (1,3;1,4)-β-D-glucan biosynthesis. **Journal of Experimental Botany**, v. 63, p. 381-392, 2012.

UNTERGASSER, A. *et al*. Primer3Plus, an enhanced web interface to Primer3. **Nucleic Acids Research**, v. 35, p. W71-W74, 2007.

VEGA-SÁNCHEZ, M. E. Y. *et al.* Loss of cellulose synthase-like F6 function affects mixed-linkage glucan deposition, cell wall mechanical properties, and defense responses in vegetative tissues of rice. **Plant Physiology**, v. 159, p. 56-69, 2012.

WANG, Y. *et al*. GSP: a web-based platform for designing genome-specific primers in polyploids. **Bioinformatics**, v. 32, p. 2382-2383, 2016.

WHITEHEAD A. *et al.* Cholesterol-lowering effects of oat β-glucan: a meta-analysis of randomized controlled trials. **American Journal of Clinical Nutrition**, v. 100, p. 1413-1421, 2014.

YE, E. Q. *et al.* Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease and weight gain. **Journal of Nutrition**, v. 142, p. 1304-1313, 2012.

# 7 GENERAL CONCLUSIONS

The oat β-glucan content is a complex trait and highly influenced by the environmental conditions, as showed by the low repeatability. Seven QTL associated with β-glucan content were identified in oats. These QTL are located on Mrg02, Mrg06, Mrg11, Mrg12, Mrg19, and Mrg20. The QTL on Mrg02, Mrg06, and Mrg11 seem to be conserved in barley. The UFRGS oat germplasm is important for other Brazilian and foreign oat breeding programs, showing a wide phenotypic variation for β-glucan content and other agronomic traits. The UFRGS Oat Panel shows weak population structure and is useful for future genome-wide association mappings in subtropical environments.

Genomic regions associated with kernel shape-related traits were also identified in oats. Three QTL controlling kernel length were identified on Mrg06, Mrg21, and Mrg24. Kernel width and kernel thickness traits, which are phenotypically and genetically correlated, share a genomic region in common on Mrg13. Negative correlations were found between these last two traits and β-glucan content, suggesting that the selection of wider and thicker kernels dilute the β-glucan content in oats. The indirect selection for β-glucan content using kernel width is effective in some environmental conditions.

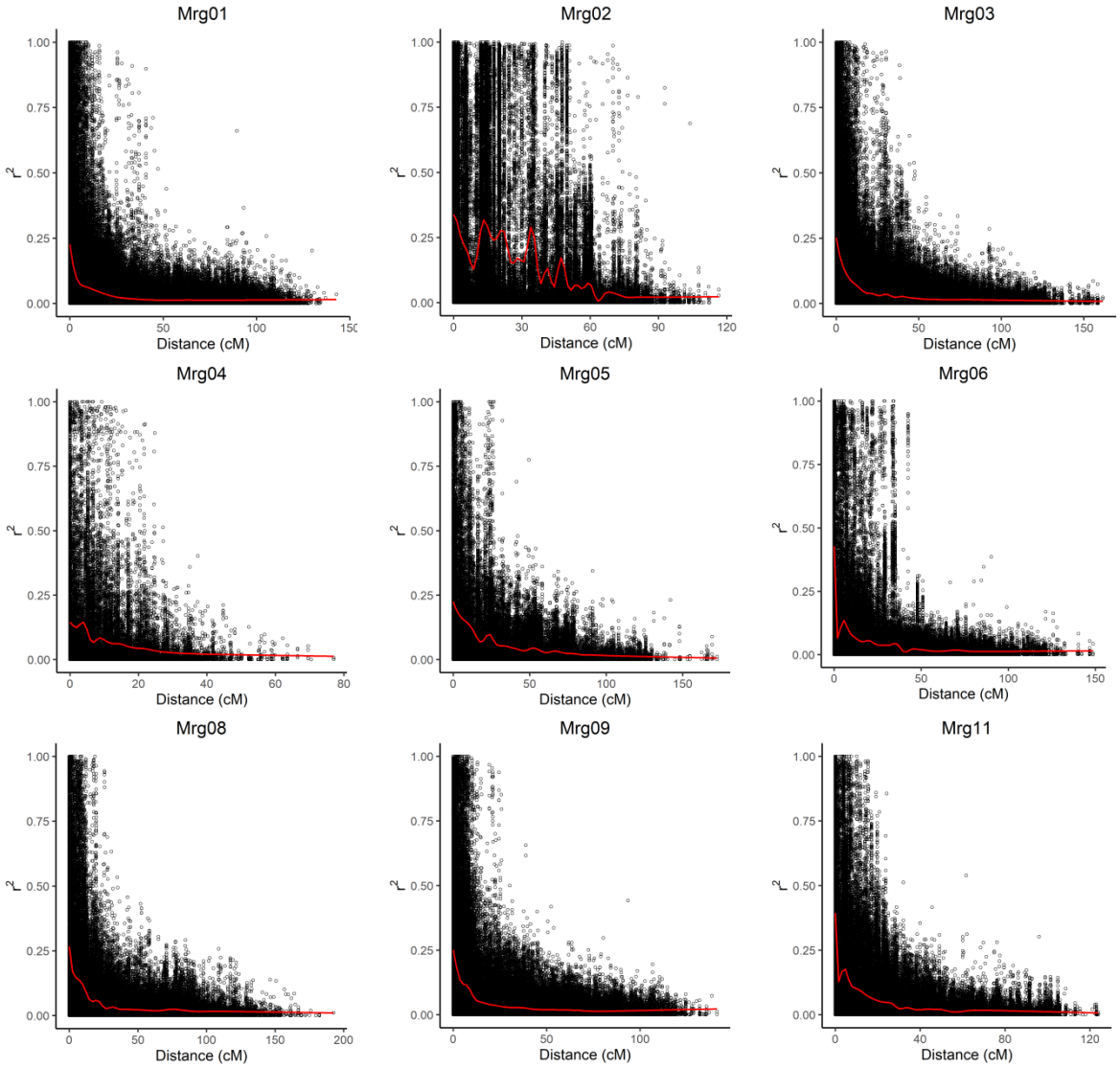
The oat β-glucan content varies between primary and secondary kernels, according to the genotype and environmental conditions. Under unfavorable conditions, some genotypes can present higher β-glucan content in secondary kernels than in primary ones. The phenotypic variation of β-glucan content and kernel-shape related traits in oats is

associated with the interaction among genotype, year, and planting date, and/or genotype, year, and kernel type.

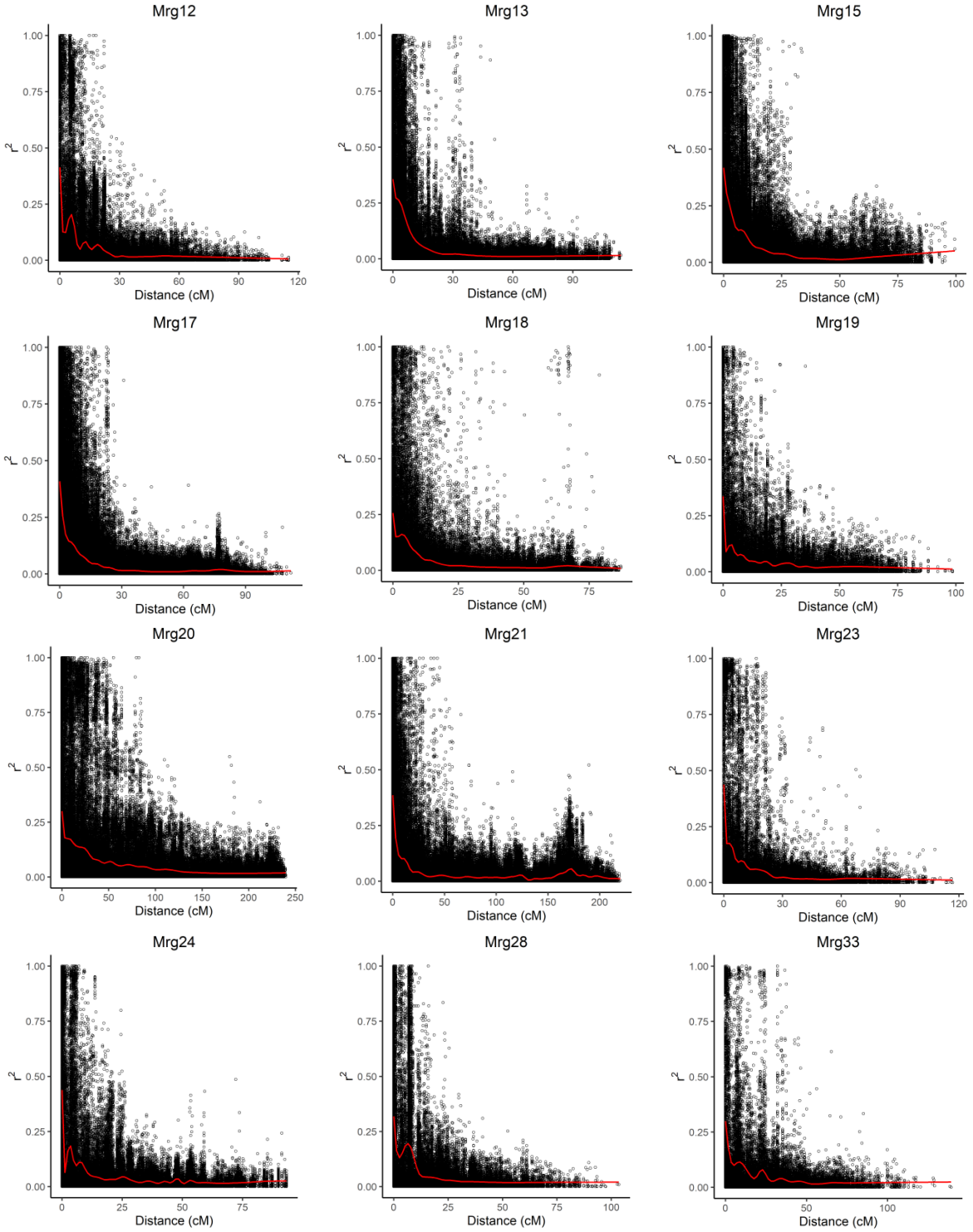
The high *CslF6* gene expression is not a guarantee of high β-glucan content in oats. The *CslF6* gene expression is mainly influenced by the environmental conditions and kernel developmental stage, and in a minor degree by the genotype. The CslF6 motif has 1767 bp in oats, regardless of the subgenome. The third exon of the *CslF6-AA* and *CslF6-DD* subgenomes encodes the same amino acid sequence. The *CslF6-CC* subgenome has 13 amino acid changes in the CslF6 motif when compared to the other ones. SNP were not identified for the *CslF6* homoeologs among the assessed genotypes, indicating that the CslF6 proteins with the same functional properties are encoded in these genotypes. The results presented in this study contribute to improve the understanding of the β-glucan content in oats under subtropical conditions.

# 8 SUPPLEMENTARY MATERIAL

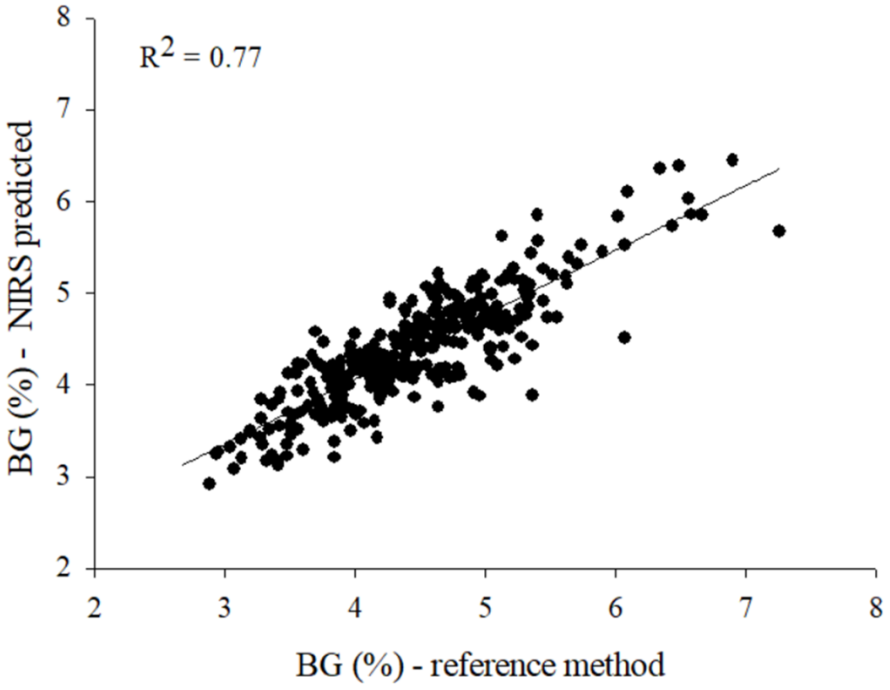
## 8.1 Chapter 1



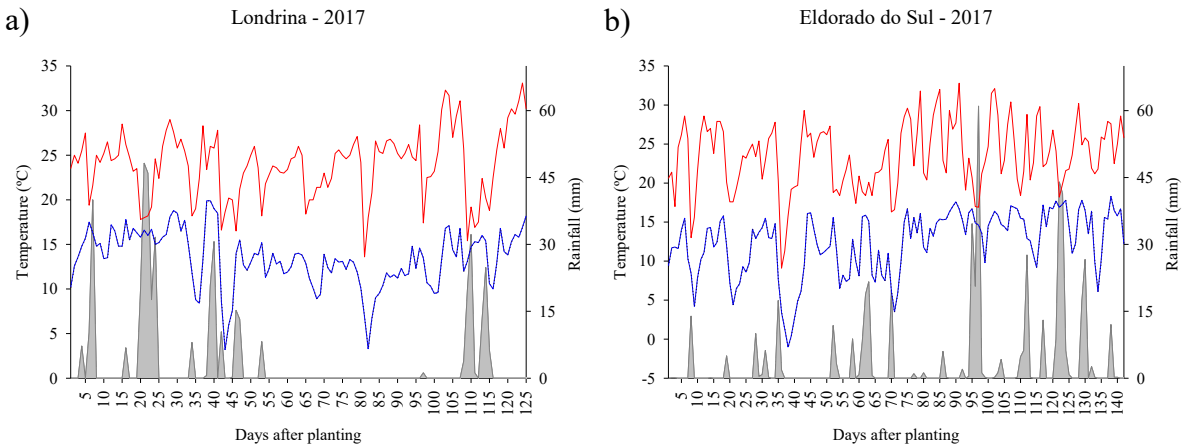

SUPPLEMENTARY FIGURE 1. Linkage disequilibrium (LD) scatter plots showing correlations ($r^2$) between marker pairs inside each consensus group as a function of genetic position; the red line is a smoothing spline.



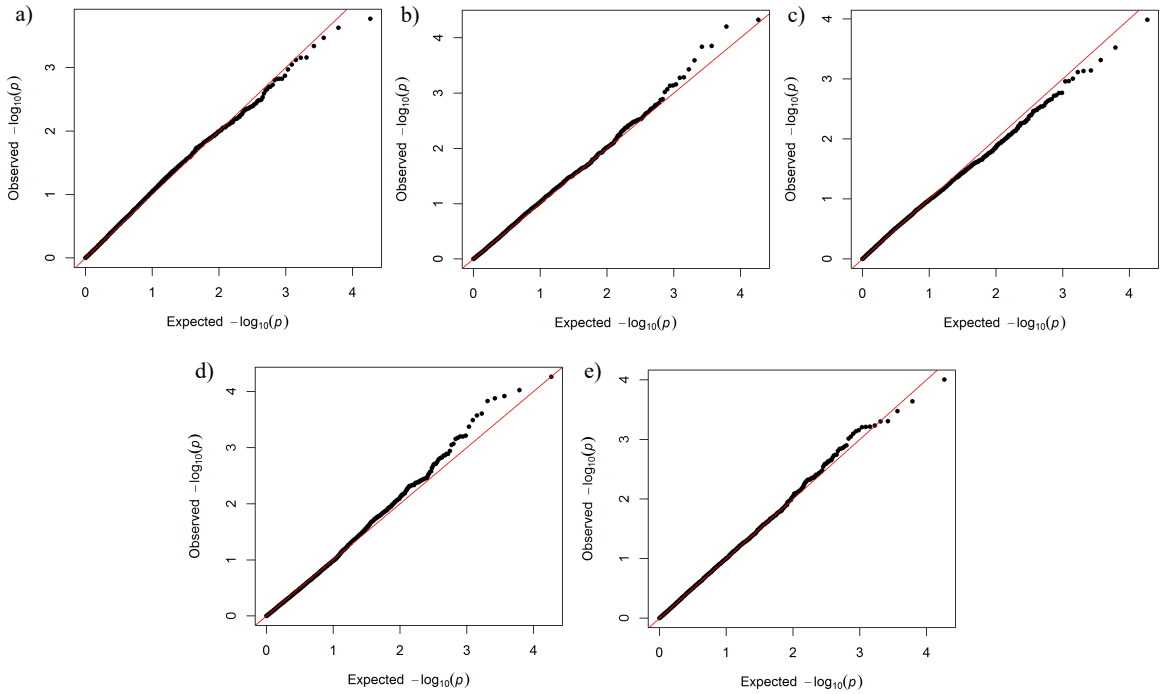

(CONTINUING) SUPPLEMENTARY FIGURE 1. Linkage disequilibrium (LD) scatter plots showing correlations ($r^2$) between marker pairs inside each consensus group as a function of genetic position; the red line is a smoothing spline.




SUPPLEMENTARY FIGURE 2. Association between enzymatic (reference) and NIRS methods for β-glucan content quantification in oats.

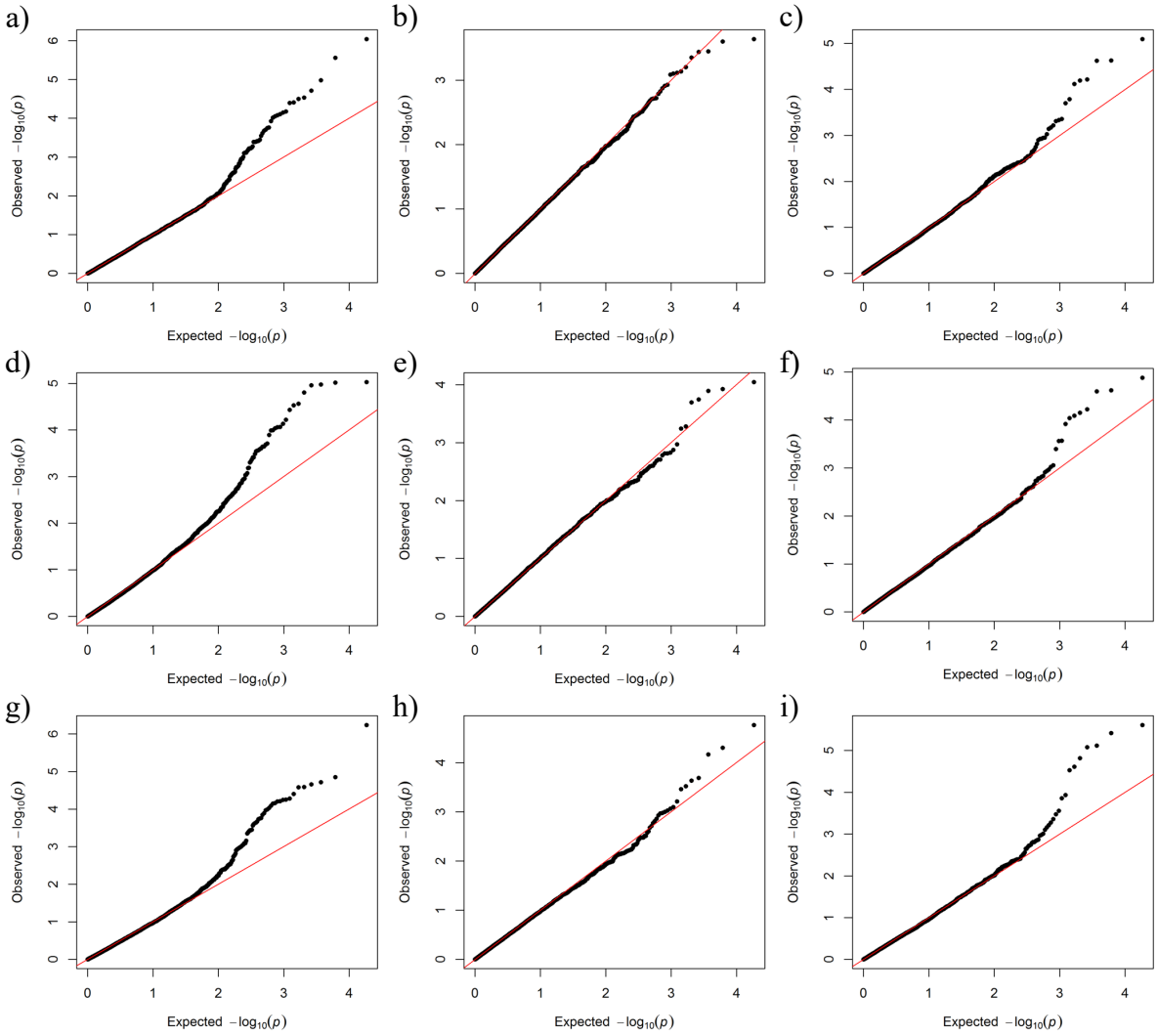



SUPPLEMENTARY FIGURE 3. Rainfall (grey), mean daily maximum (red), and minimum (blue) temperatures recorded from planting to maturity in two locations during 2017: Londrina (a); and Eldorado do Sul (b) Two irrigations of 20 mm were performed in Londrina between 60 and 95 days after planting. Sources: IAPAR and UFRGS local data.



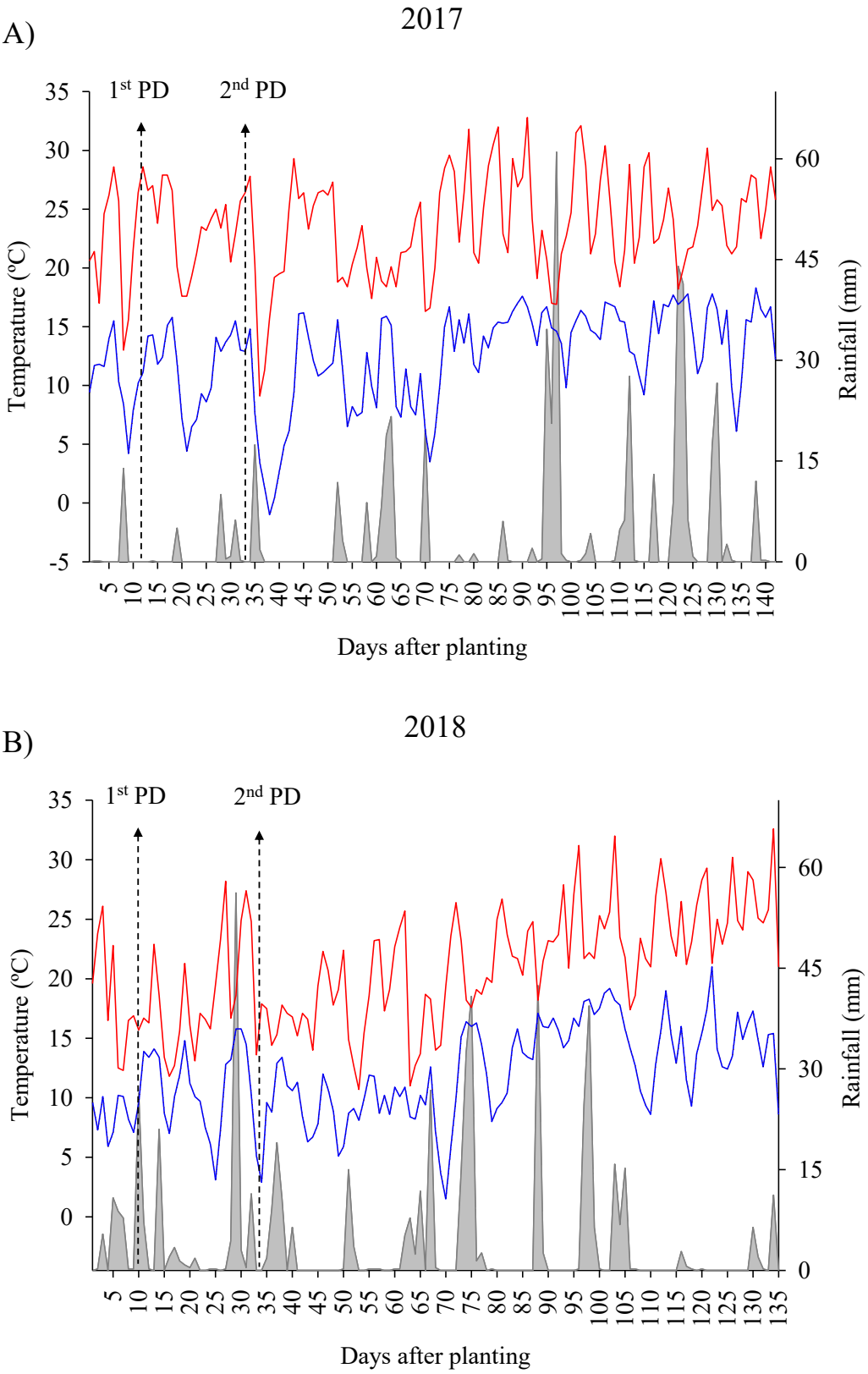

SUPPLEMENTARY FIGURE 4. Quantile-quantile plots of the observed versus expected p-values under the null hypothesis for β-glucan content in oats. a, Londrina 2017 (enzymatic); b, Londrina 2017 (NIRS); c, Eldorado do Sul 2017 (NIRS); d, Eldorado do Sul 2018 (NIRS); e, Multi-environment (NIRS).

## 8.2 Chapter 2



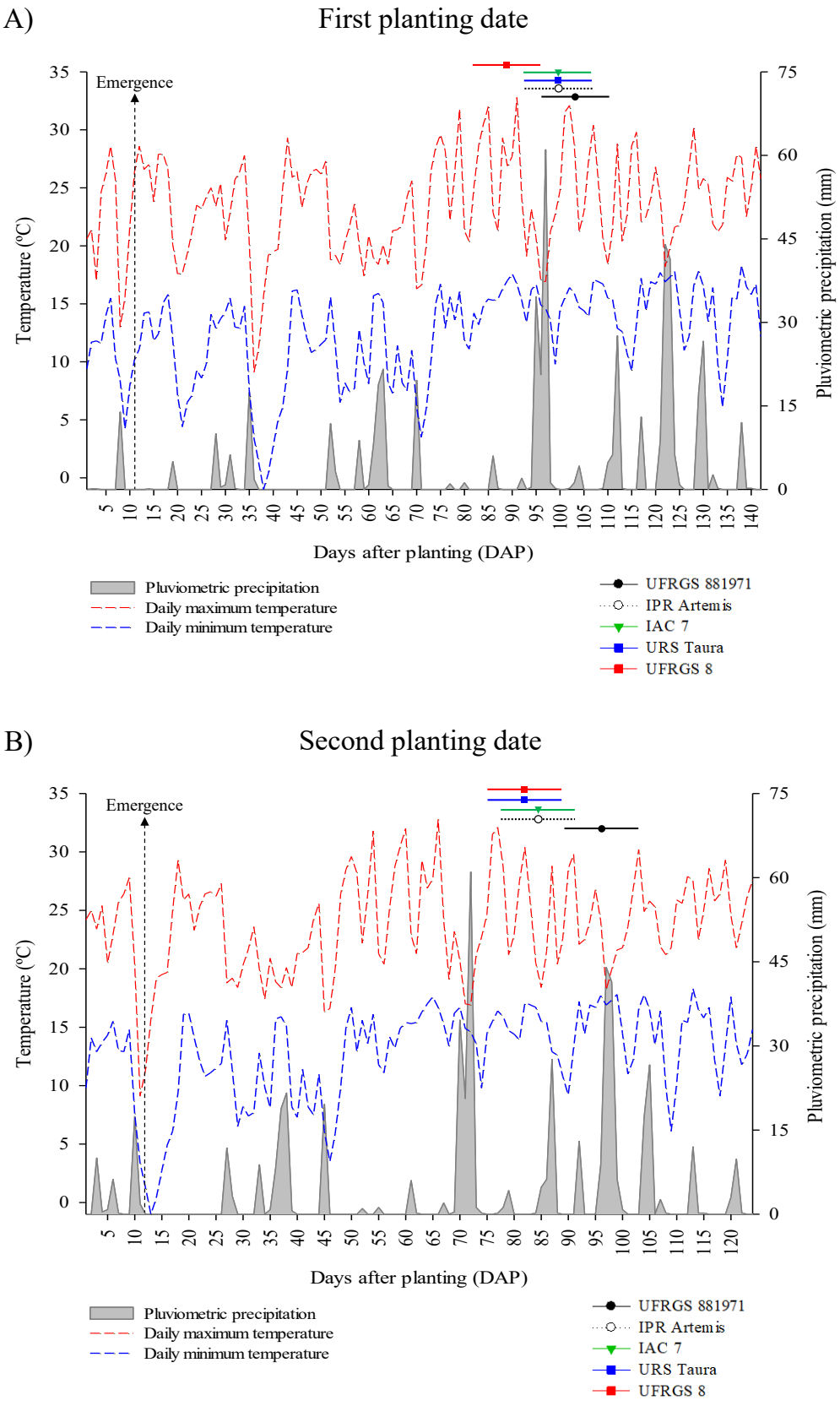

SUPPLEMENTARY FIGURE 1. Quantile-quantile plots of the observed versus expected p-values under the null hypothesis for kernel shape-related traits in oats. a, b, and c = kernel length, kernel width, and kernel thickness in Londrina. d, e, and f = kernel length, kernel width, and kernel thickness in Eldorado do Sul. g, h, and i = kernel length, kernel width, and kernel thickness in the multi-environment analysis.

**8.3 Chapter 3**

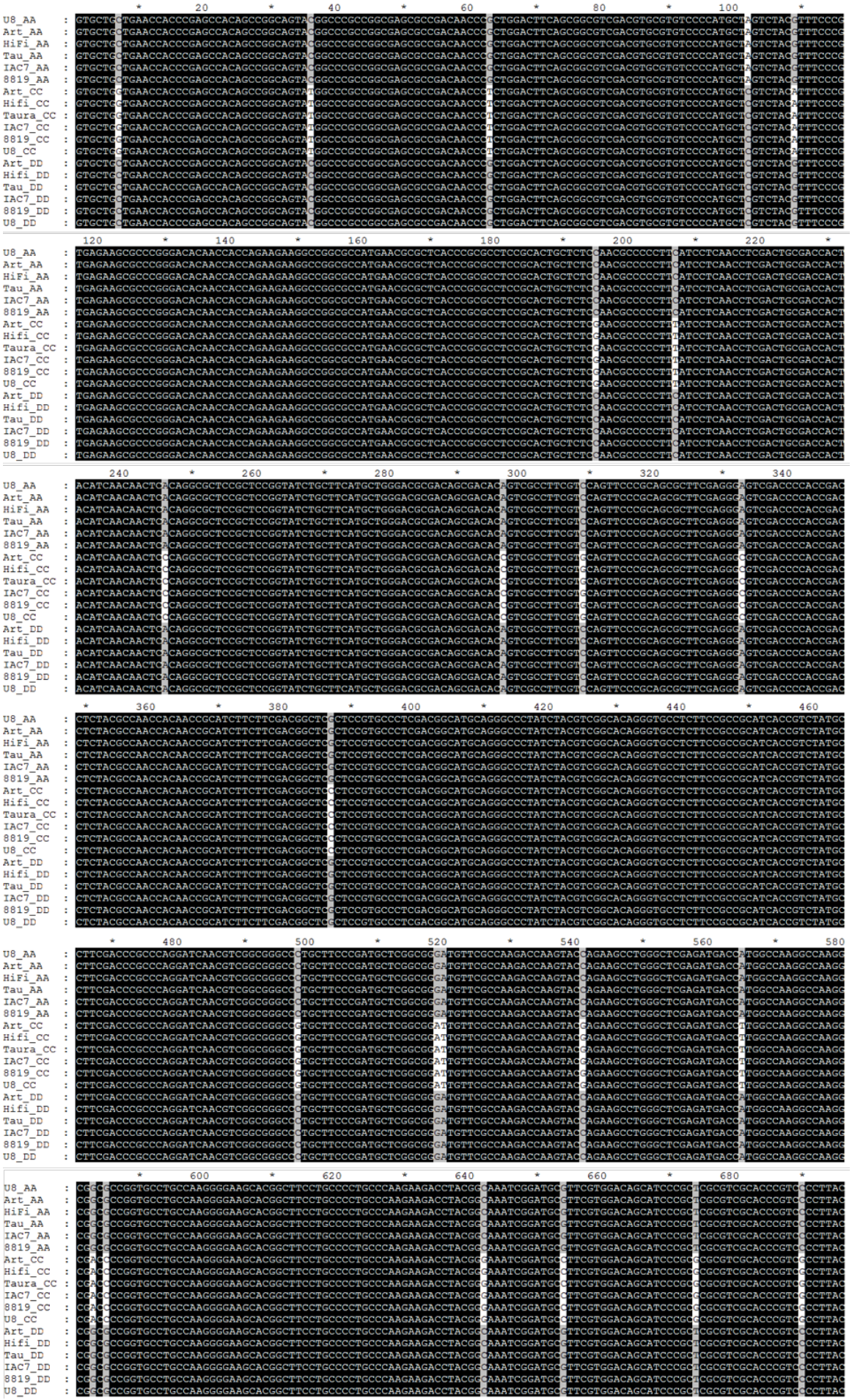



SUPPLEMENTARY FIGURE 1. Rainfall (grey), mean daily maximum (red), and minimum (blue line) temperatures recorded from planting to maturity in 2017 (A) and 2018 (B). Source: local weather station at the UFRGS Experimental Agronomic Station.

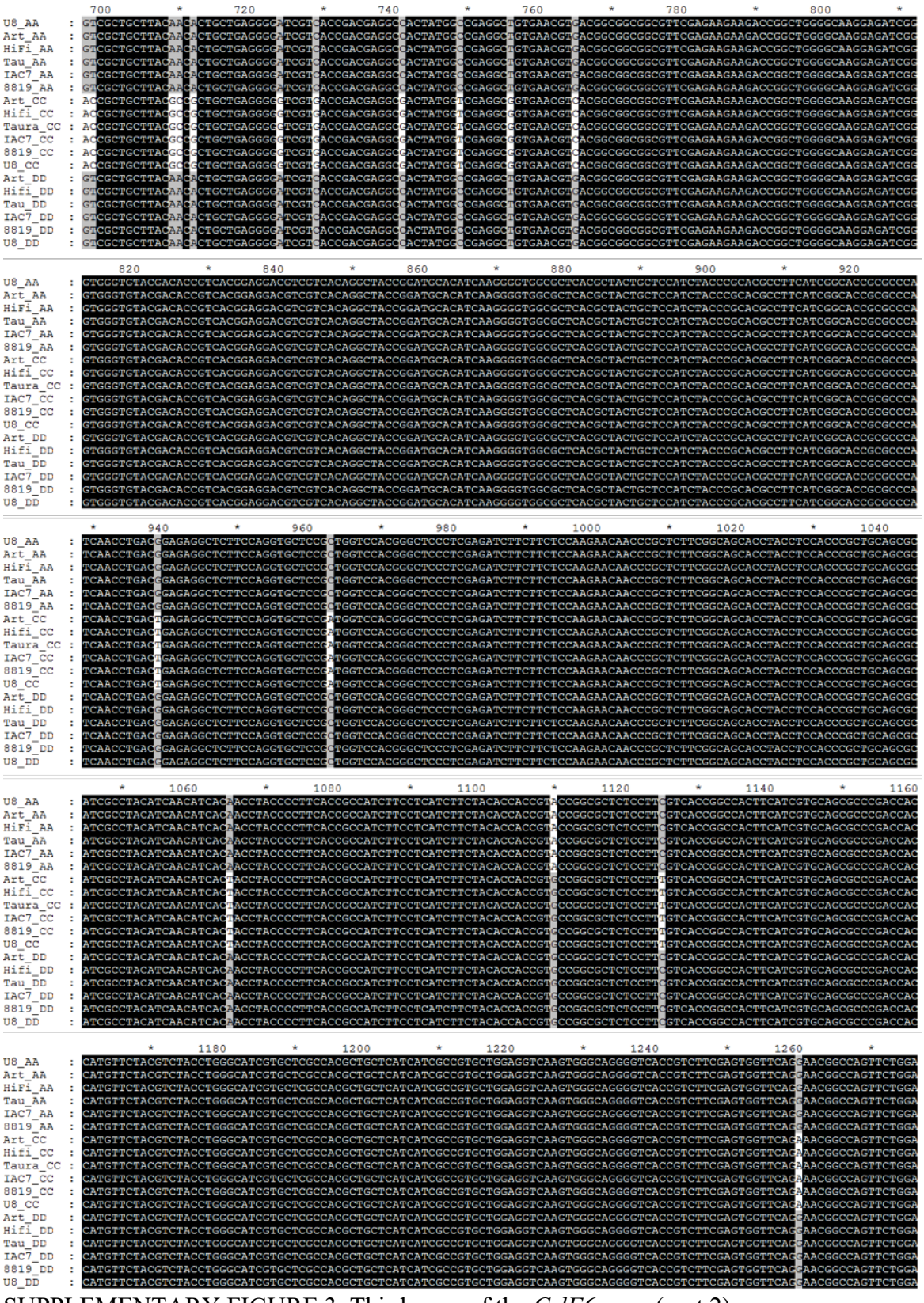
## 8.4 Chapter 4



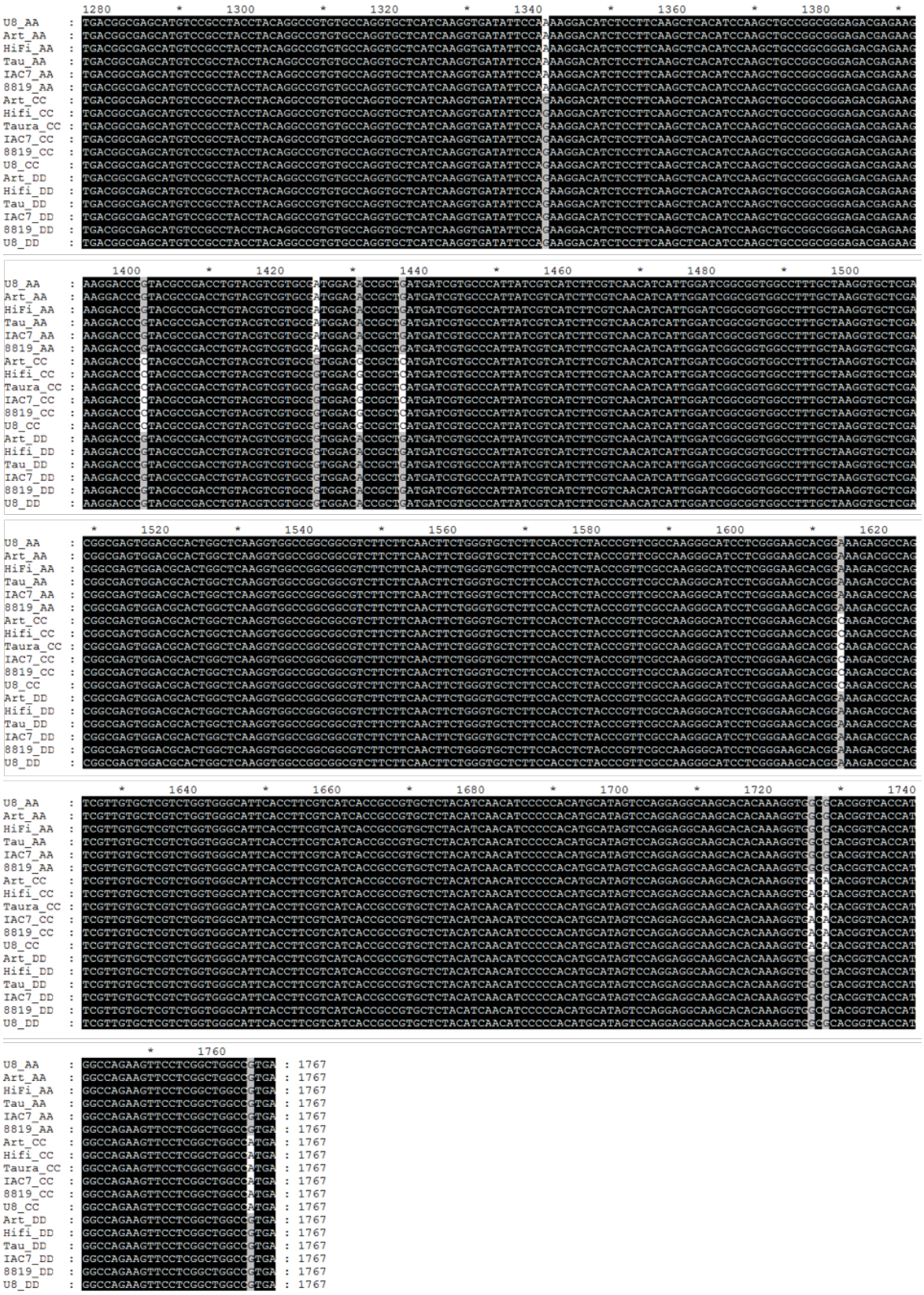

SUPPLEMENTARY FIGURE 1. Pluviometric precipitation, mean daily maximum, and minimum temperatures recorded from planting to maturity at two planting dates: 06/12/2017 (A); and 07/07/2017 (B) Horizontal lines represent the kernel collect moment in each genotype.



SUPPLEMENTARY FIGURE 2. Third exon of the *CslF6* gene (part 1).



SUPPLEMENTARY FIGURE 3. Third exon of the *CslF6* gene (part 2).



SUPPLEMENTARY FIGURE 4. Third exon of the *CslF6* gene (part 3).

# 9 VITA

Cristiano Mathias Zimmer is son of Ildemar Genesio Zimmer (*in memorian*) and Alvete Maria Zimmer. He was born on July 24th, 1991, in Salto do Lontra, Paraná state, Brazil. Studied most part of his high school at the Irmã Maria Margarida School, in his home town. Moved to Pelotas city, Rio Grande do Sul state, Brazil, on January 2008, where finished his high school at the Érico Veríssimo School in the same year. Started his Bachelor of Science degree in Agronomy at the College of Agronomy 'Eliseu Maciel', Federal University of Pelotas, graduating agronomist on March 14th, 2014. Cristiano has been working with plant breeding since the first semester of his bachelor degree and participated in the 'Ciências Sem Fronteiras' program spending two semesters at the Rovira I Virgili University, Spain, from September 2012 to May 2013. In 2014, entered the Master of Science graduate course in Plant Breeding at the Federal University of Rio Grande do Sul, in Porto Alegre city, Rio Grande do Sul state, Brazil, supported by the CAPES foundation. Earned a Master of Science degree on February 18th, 2016, and started the doctoral graduate course in Plant Breeding at the Federal University of Rio Grande do Sul, supported by the CNPq. Cristiano was enrolled as research scholar at the University of Minnesota, in the United States, from September 2018 to August 2019, supported by the CAPES foundation, where developed part of his dissertation.