UNIVERSITA' DEGLI STUDI DI BARI ALDO MORO

Dipartimento di Scienze del Suolo, della Pianta e degli Alimenti

DOTTORATO DI RICERCA IN BIODIVERSITÀ, AGRICOLTURA E AMBIENTE

Curriculum: Genetica e biotecnologie vegetali

CICLO XXXVII

SSD AGRI-06/A

Genetic studies of new grains improved with quality traits

Dottorando: Dott. Davide Caranfa

Coordinatore: Chiar.ma Prof.ssa Cinzia Montemurro

Supervisor: Chiar.ma Prof.ssa Agata Gadaleta

Co-supervisor: Chiar.ma Dott.ssa Ilaria Marcotuli

ESAME FINALE 2025

Borsa di Dottorato finanziata nell'ambito del Programma Operativo Nazionale Ricerca e Innovazione 2014-2020 (CCI 2014IT16M2OP005), risorse FSE REACT-EU, Azione IV.4 "Dottorati e contratti di ricerca su tematiche dell'innovazione" e Azione IV.5 "Dottorati su tematiche Green"

Contents

Summary

In a global context increasingly focused on sustainability and food quality, the search for sustainable and nutrient-rich raw materials is gaining growing interest. In cereal cultivation, worldwide researchers are shifting their focus to ancient cereals and wild species of domesticated crops. Among these, *Aegilops caudata*, a wild species, represents a promising genetic resource due to its nutritional characteristics that make it valuable for wheat improvement. Specifically, *Aegilops caudata* has been used to create amphiploid lines useful for studying qualitative traits such as β-glucan content, a key compounds for human health.

Chapter I of this thesis analyzes a key gene in β-glucan synthesis, the cellulose synthase *CslF6* gene, by determining the gene sequence and protein structure in *Ae. caudata*, highlighting evolutionary differences compared to other grass species and examining how variations in the amino acid sequence may influence the β-glucan content. Comparative analysis with other monocots revealed significant similarities with the wheat genome, with variations in amino acid motifs within the catalytic portions of the polypeptide chain, suggesting potential unique functional adaptations for *Ae. caudata*.

Chapter II explores the role of the *CslF6* gene in β-glucan synthesis, the total β-glucan content, the glycosidic linkages within the β-glucan structure, and the profile of major oligosaccharides in bread and durum wheat genotypes, wild parental species (*Aegilops caudata* and *Dasypyrum villosum*), and their hybrids. The analysis of *CslF6* gene expression revealed a correlation between gene expression and β-glucan accumulation. The results indicate that wild species, such as *Ae. caudata* and *Dasypyrum villosum*, have a higher β-glucan content compared to wheat, and hybrids show a significant increase in β-glucan levels. Additionally, through glycomics analysis, the study highlighted differences in the monosaccharide composition of the cell wall and the glycosidic linkages in the analyzed genotypes, providing new tools for the selection of wheat lines with high β-glucan content.

Chapter III shifts the focus on two important wheat qualitative traits: protein content (GPC) and yellow index (YI), which are critical for pasta production. The study, conducted on 144 tetraploid wheat accessions (*Triticum turgidum* L.), identified superior alleles associated with these qualitative traits through a Genome-Wide Association Study (GWAS). The analysis identified several association regions (QTLs) between SNP markers and the traits considered (GPC and YI). Bioinformatics analysis revealed candidate gene sequences co-localized within these QTL regions. Notably, a key gene involved in nitrogen metabolism, *Glutamine synthetase 2* (*GS2*), and a key gene for carotenoid biosynthesis, *Phytoene synthase 1* (*Psy-A1*), were identified. Furthermore, the analysis led to the identification of new haplotypes, contributing to a better understanding of the genetic mechanisms controlling protein content and yellow color in wheat, facilitating the development of superior wheat varieties.

These three studies, while addressing different aspects of wheat genetics and its wild relatives, are interconnected within a research framework aimed at improving the nutritional and qualitative traits of wheat. The exploration of the *CslF6* gene and its role in β-glucan biosynthesis, alongside the characterization of qualitative traits such as protein content and yellow index, contributes to a deeper understanding of the genetic potential of wheat species, particularly those from wild genotypes such as *Ae. caudata*. The information obtained from this research offers new opportunities for the selection and development of more nutritious, sustainable, and consumer-oriented wheat varieties worldwide.

Sintesi

In un contesto globale sempre più attento alla sostenibilità e alla qualità degli alimenti, la ricerca di materie prime sostenibili e nutrienti sta guadagnando sempre maggiore interesse. Nel mondo della cerealicoltura l'attenzione dei ricercatori di tutto il mondo si sta spostando sui cereali antichi e sulle specie selvatiche delle colture domesticate ricche in componenti bioattive. Tra questi, l'*Aegilops caudata*, una specie selvatica, rappresenta una risorsa genetica promettente, grazie alle sue caratteristiche nutrizionali che la rendono utile per il miglioramento del frumento. In particolare, l'*Aegilops caudata* è stata utilizzata per la costituzione di linee anfiploidi utili per lo studio di caratteri qualitativi quali il contenuto in β-glucani, composti chiave per la salute umana.

Il **capitolo I** della presente tesi analizza il gene chiave nella sintesi dei β-glucani andando a determinarne la sequenza genica e la struttura proteica in *Ae. Caudata* del gene *cellulosa sintasi CslF6*, mettendo in evidenza le differenze evolutive rispetto ad altre specie di graminacee e analizzando come le variazioni nella sequenza amminoacidica possano influire sul contenuto di β-glucani. L'analisi comparativa con altre monocotiledoni ha rivelato somiglianze significative con il genoma di frumento con variazioni di motivi amminoacidici nelle porzioni catalitiche della catena polipeptidica, suggerendo potenziali adattamenti funzionali unici per *Ae. Caudata*.

Il **capitolo II** analizza il ruolo che il gene *CslF6* ha nella sintesi dei β-glucani, il contenuto totale dei β-glucani, i legami glicosidici che intercorrono nella struttura dei β-glucani e il profilo dei principali oligosaccaridi in genotipi di frumento tenero e duro, parentali selvatici (*Aegilops caudata* e *Dasypyrum villosum*) e i loro ibridi.

L'analisi dell'espressione del gene *CslF6* ha rivelato una correlazione tra l'espressione genica e l'accumulo di β-glucani. I risultati evidenziano che le specie selvatiche, come *Ae. caudata* e *Dasypyrum villosum*, presentano un contenuto di β-glucani superiore rispetto ai frumenti, e che gli ibridi mostrano un aumento significativo del contenuto di β-glucani. Inoltre, mediante analisi glicomiche, lo

studio ha evidenziato differenze sulla composizione monosaccaridica della parete cellulare e dei legami glicosidici nei genotipi analizzati fornendo nuovi strumenti per la selezione di linee di frumento ad alto contenuto di β-glucani.

Nel **capitolo III**, l'attenzione si sposta su due importanti caratteristiche qualitative del frumento: il contenuto proteico (GPC) e l'indice giallo (YI), che sono determinanti per la produzione di pasta. Lo studio, condotto su 144 accessioni di frumenti tetraploidi (*Triticum turgidum L. var durum*), ha identificato alleli superiori associati a questi caratteri qualitativi attraverso una analisi di associazione mediante Genome-Wide Association Study (GWAS). L'analisi ha identificato diverse regioni di associazione (QTL) tra marcatori SNP e i caratteri presi in considerazione (GPC e YI). Mediante analisi bioinformatica sono state identificate le sequenze di geni candidati co-localizzati nelle regioni QTL. In particolare, sono stati identificati un gene chiave coinvolto nel metabolismo dell'azoto, la *glutammina sintetasi* 2 (*GS*2), e un gene chiave per la sintesi dei carotenoidi, la *fitoene sintasi* 1 (*Psy*-A1). Infine, l'analisi ha permesso l'identificazione di nuovi aplotipi che contribuiscono a comprendere meglio i meccanismi genetici che controllano il contenuto proteico e il colore giallo nei frumenti facilitando lo sviluppo di nuove varietà di frumento.

Questi tre studi, pur affrontando aspetti differenti della genetica del frumento e dei suoi parentali selvatici, si intrecciano in un quadro di ricerca volto a migliorare le caratteristiche nutrizionali e qualitative dei frumenti. L'esplorazione del gene *CslF*6 e del suo ruolo nella biosintesi dei β-glucani, insieme alla caratterizzazione dei caratteri qualitativi quali contenuto proteico e indice di giallo, contribuiscono ad una comprensione più approfondita delle potenzialità genetiche delle specie progenitrici del genere *Triticum*, in particolare quelle provenienti da genotipi selvatici come *Ae. caudata*. Le informazioni ottenute da queste ricerche offrono nuove opportunità per la selezione e lo sviluppo di varietà di frumento più nutrienti, più sostenibili e più adatte alle esigenze dei consumatori di tutto il mondo.

Introduction

Durum wheat genetics and breeding

The global durum wheat supply was tight at the beginning of the 2023/24 season due to drought conditions in the Middle East and Canada, as well as quality issues within the European Union (EU), which reduced the availability of durum wheat and exerted upward pressure on global prices. However, as the season progressed, non-traditional durum wheat exports from Russia and Turkey began flooding the markets with low-priced wheat, weighing on global durum wheat values.

According to data from the International Grains Council (IGC), global durum wheat production decreased by 10% to 31.4 million metric tons (MMT) in 2023/24, while consumption exceeded production by 2.6 MMT, leading to a tight durum wheat balance. In Canada, the world's largest exporter, durum wheat production dropped by 30% to 4.0 million metric tons due to drought in key growing areas, as reported in the Canadian Outlook for Principal Field Crops. In the EU, drought in Spain reduced production, resulting in a total EU production of 7.0 million metric tons, 7% lower than the previous year, according to the IGC. At the same time, rainfall during harvest in Italy compromised the quality of the durum wheat crop. As a result, durum wheat imports into the EU increased by 26% to 3.4 million metric tons.

The combined impact of reduced durum wheat availability and increased import requirements supported global durum wheat prices during the last quarter of 2023 and the early months of 2024 (U.S. Wheat Associates).

The domestication of durum wheat began in the Fertile Crescent, an area encompassing parts of modern-day southwest Asia, including regions of Iraq, Iran, Turkey, Syria, and Jordan (Peng 2011). This early cultivation (Shewry 2009), which dates to ancient Mesopotamia, involved selecting wild progenitors of durum wheat, such as *T. turgidum* subsp. *dicoccoides*, with an emphasis on yield and desirable traits such as non-brittle rachis and free-threshing kernels.

The modern history of durum wheat breeding in the Mediterranean area started in the early 20th century, evolving alongside advancements in agricultural science. Breeders have always aimed for high yield, robust end-use qualities, and resilience to both biotic and abiotic stresses. Initially, breeding efforts focused on utilizing local genetic resources, but the Green Revolution introduced new, high-yielding cultivars that were short-statured and adapted to intensive agriculture. These cultivars became widely adopted in breeding programs worldwide. The result was a notable increase in durum wheat yields through the second half of the 20th century (Isidro 2011).

In addition to yield, durum wheat breeding has played a major role in enhancing grain quality (De Vita 2007, Motzo 2004). Studies comparing cultivars from different breeding periods suggest that while genetic improvements have led to reduced protein content in the grain (as a result of higher yields), pasta quality has not been compromised (Waddington 1987). Modern cultivars, although lower in protein content, have larger kernels and higher carbohydrate content, which can dilute protein levels but improve pasta quality. Furthermore, these

modern cultivars show an increased gluten index, which further boosts pasta-making qualities (De Santis 2017).

The evolutionary history of wheat has been shaped by the allopolyploidization of species within Triticum and Aegilops, leading to the emergence of tetraploid and hexaploid species. The cultivation of the wild tetraploid emmer wheat, *Triticum dicoccoides*, gave rise to *Triticum dicoccon* and eventually to T. durum. Tetraploid species and the hexaploid bread wheat *T. aestivum*, despite lacking a directly identified hexaploid wild relative, expanded their cultivation through human migration and trade (Kilian 2010, Zohary 2012). However, migration and intensive agricultural improvement have diminished the genetic diversity of *T. aestivum* and *T. durum* (Tanksley & McCouch 1997, Ross-Ibarra 2007). Presently, climate change poses a significant threat to wheat productivity (Zhao 2017, Iizumi 2021). Both wild and domesticated relatives of wheat, well adapted to harsh habitats, are essential resources for enhancing resistance to abiotic stresses and improving tolerance to pests and diseases such as wheat midge and powdery mildew (Nsarellah 2003, Li 2020, Fatima 2020). The genetic diversity of these wild relatives is crucial for the future of wheat cultivation, with conservation efforts in genetic banks serving as a key resource (Dempewolf 2014, Kilian 2021). Recently, biofortification of wheat has become a focal point, particularly in addressing “hidden hunger,” referring to micronutrient deficiencies such as iron (Fe) and zinc (Zn), which are vital for human health (Graham 2001, 2007). The biofortification approach, which involves enriching crops with micronutrients through agronomic and selection methods, has

highlighted that wild relatives of wheat are highly efficient at accumulating micronutrients, making them valuable for seed enrichment (Arora 2019, Cakmak 2000, Chhuneja 2006). Wheat, consumed in various forms such as bread, pasta, and couscous, owes its quality to gluten, a protein complex composed of glutenins and gliadins. While glutenins determine dough elasticity, gliadins are responsible for its viscosity (Wieser 2007; Shewry, 2019). However, gliadins can trigger celiac disease (Biesiekierski 2017). A potential approach for improving wheat for celiac patients involves combining the properties of glutenins with less toxic gliadin variants, aiming to develop wheat that is safe for those affected by celiac disease. Finally, phenolic compounds in wheat, such as phenolic acids, are well-known for their health benefits (Laddomada 2015). These compounds are predominantly located in the aleurone and bran (Liyana-Pathirana & Shahidi 2006). Both domesticated einkorn and emmer have been identified as rich in phenolic acids (Serpen 2008, Barański 2020), suggesting that wild grains could be a valuable source for enhancing the phenolic content of the whole wheat kernel.

In summary, the evolution, selection, and introduction of beneficial traits from wheat's wild relatives are critical for addressing contemporary challenges in wheat cultivation, including climate change, nutritional deficiencies, and disease management.

Advances in wheat breeding have also made using molecular markers, which can help identifying genetic variation even among seemingly similar phenotypes. Molecular tools, such as Single Nucleotide

Polymorphisms (SNPs), have revolutionized the genetic approach to durum wheat (*Triticum turgidum* var. *durum*) (Maccaferri 2015). Due to their high density in the genome and relative, SNPs enable precise mapping of genetic loci associated with agronomically important traits, such as disease resistance, grain quality, and tolerance to environmental stresses. In particular, Marker-Assisted Selection (MAS) has made it possible to identify favourable genetic variants early in the breeding process, reducing selection time and increasing the efficiency of breeding programs.

Moreover, SNPs are ideal tools for studying genetic diversity among durum wheat varieties, facilitating the management of genetic resources and the improvement of agronomic traits through the introduction of new alleles. In this context, SNPs not only represent a key resource for improving durum wheat quality and productivity, but also for understanding the genetic mechanisms underlying plant adaptation to various environmental conditions.

The combination of molecular markers with genealogical data has enabled better exploitation of genetic diversity in both wild and cultivated wheat, enhancing the efficiency and precision of selection programs. Unlike traditional morphological markers, molecular markers can be applied at any developmental stage and are not influenced by environmental factors, offering significant advantages in selecting agronomically relevant traits.

New genetic variations identified through molecular technologies can also improve drought and heat tolerance in durum wheat traits that are

increasingly important in the face of climate change, particularly in Mediterranean climates.

Looking ahead, it is clear that agricultural production must adopt more sustainable practices, preserve crop biodiversity, and develop varieties suited to low-input agriculture, aligning breeding goals with the real needs of farmers and the market, particularly in Mediterranean regions. The use of genomic tools, such as Quantitative Trait Locus (QTL) mapping (Kissing 2015), has already provided valuable insights into the genetic basis of key traits like yield, grain quality, and disease resistance. These techniques are essential for understanding the genetic architecture of durum wheat and facilitating the identification of loci associated with traits of interest (Pozniak 2007).

However, there has been limited attention to improving certain health-related components of durum wheat, such as dietary fiber, arabinoxylans, and beta-glucan, which are important for human nutrition. Some studies suggest that intensive selection has had little impact on these nutritional aspects, but there is potential to further improve them through targeted selection efforts. The use of markers associated with these traits could help identify the best lines earlier in the selection process, allowing the development of wheat varieties with enhanced nutritional characteristics.

In summary, durum wheat breeding programs in the 20th century primarily focused on yield and pasta quality (Rossini 2018, Li 2018). However, with the intensification of selection, there is growing interest in improving the nutritional properties of wheat.

Wheat in nutrition and health

The growing global demand for wheat is largely driven by its ability to be transformed into a variety of unique food products, coupled with the rising consumption of these goods due to industrialization and westernization. A key factor is the distinctive characteristics of gluten, a protein found in wheat, which enables its use in the production of bread, baked goods, pasta, noodles, and various functional ingredients.

While wheat is often viewed primarily as a source of energy (carbohydrates), it also provides a substantial amount of other essential nutrients. These include proteins, fiber, and smaller amounts of fats, vitamins, minerals, and phytochemicals, all of which contribute to a balanced and healthy diet. As awareness of the health benefits of wheat-based products grows, there is increasing interest in tapping into the natural variations of these bioactive compounds. However, in some cases, the natural variability of certain traits may be limited or challenging to harness effectively, which calls for alternative strategies.

Presently, key focus areas for improving wheat's nutritional value include increasing mineral content, enhancing resistant starch, boosting antioxidant levels such as carotenoids, and improving grain protein content and dietary fiber. Despite its relatively modest fiber content (usually between 3-8%), wheat remains one of the most important sources of fiber in human and animal diets.

Wheat quality

The concept of "wheat quality" involves several factors that contribute to its overall assessment. Grain quality is determined by a variety of physical and compositional attributes, with specific threshold values set based on the intended end-use. For staple grains like wheat, physical properties such as grain size and shape play a significant role in determining milling yield and screening losses, which ultimately affect processing efficiency and the economic value of the grain. Additionally, the protein content and composition of the grain are critical indicators of quality, as they influence key aspects of dough properties like mixing behavior, rheological characteristics, and bread-making performance. This includes dough strength, development time, extensibility, breakdown, and loaf volume, all of which contribute to the quality of the final product and the efficiency of the bread-making process.

The breeding program begins with the evaluation of genotypes. The selection of parental plants is guided by an analysis of cultivar performance, which includes factors such as yield potential, end-use quality, and resistance to both abiotic and biotic stresses. These factors are closely linked to the genetic aspects of wheat quality. Additionally, the evaluation considers field performance. During this stage, proper grain conservation is essential. Kernel quality plays a key role in determining the quality of the final product from commercial wheat varieties, as it influences the types of products that can be produced. This distinction between commercial and technological value becomes

evident at the outset of wheat production and is closely tied to global market demands for various end-products.

Commercial value is primarily assessed by the milling rate, which refers to the amount of flour produced from 100 kg of wheat. This rate aims to maximize flour yield by evaluating the seed's size, thousand kernel weight, ash content (or mineral content), and the percentage of defective seeds (e.g., pre-germination, small or white seeds, or any pathological discoloration). Technological value, on the other hand, is assessed based on the strength of the flour protein when mixed with water to form dough, the amount of water needed to achieve a workable dough, and the color of the flour. The quantity and quality of the protein have a significant impact on the characteristics of the final food products. For example, higher protein content typically results in greater water absorption, which increases productivity and extends the shelf life of the final products (Jefremova 2015).

Dough water absorption is closely linked to starch content. A higher starch content affects the dough’s functional properties, influencing characteristics like gelatinization, pasting behavior, and baking performance (Seib 1994). This is particularly important in the production of baked goods such as cakes and certain types of bread, where it contributes to the product's tenderness.

Overall, whole-grain quality encompasses physical factors like milling yield, test weight, and screenings, which are influenced by both the genetic and environmental conditions.

Wheat as a source of carotenoids

Carotenoids are a group of organic compounds found in plants, algae, and certain bacteria, playing a crucial role in photosynthesis. These pigments are responsible for the yellow color of the endosperm in cereal grains. From a nutritional perspective, carotenoids are valued for their antioxidant properties and their role in supporting the immune system. High carotenoid intake has been linked to a lower risk of developing chronic conditions, including certain cancers, cardiovascular diseases, neurodegenerative disorders, and macular degeneration (Pérez-Gálvez 2005).

To date, more than 600 carotenoids have been discovered across animals, plants, and microorganisms. Carotenoids are commercially important not only for their contribution to the nutritional quality of food products but also for their use as natural colorants, making them highly research by the consumers.

One key area of research focuses on improving our understanding of the carotenoid biosynthesis pathway and the genetic factors regulating the production of yellow pigments. The antioxidant properties of carotenoids, along with their protein content, enhance the nutritional and technological properties of flour. For instance, whole wheat flour, which is rich in fiber, vitamins, and antioxidant compounds (carotenoids, tocopherols, and flavonoids) provides additional nutritional benefits when combined with protein-rich crops such as soybeans or oilseed residues.

Carotenoids are typically found in the chloroplasts and chromoplasts of plants and are synthesized from the isoprenoid pathway as C40-tetraterpenoid compounds. Their basic structure consists of a chain of conjugated double bonds, which may have cyclic end groups and can be modified by the addition of oxygen-containing functional groups (Stahl 2003). Carotenoids can be classified into two main categories based on their oxygen content: carotenes (which are hydrocarbons) and xanthophylls (which contain one or more oxygen atoms) (Van den Berg 2000) (Figure. 1.). This distinctive structure gives the carotenoids biological and chemical properties, including their ability to absorb light due to the conjugated double bond system. This light absorption also makes carotenoids chemically unstable, especially when exposed to light. Non-cyclic carotenoids (in the trans configuration) can convert into a mixture of stereoisomers, where some double bonds may switch to the cis form. Consequently, the trans configuration is the predominant form in nature. As lipophilic compounds, carotenoids are stored in cell membranes or lipoproteins, and their affinity for lipids affects their absorption, transport, and excretion in the body (Stahl 2003).

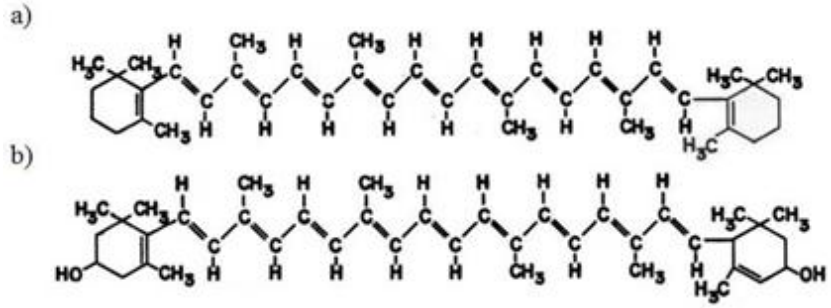


Figure. 1. Example of a molecular structure of some carotenoid molecules: in a) β-carotene from carotenes class and in b) lutein from xanthophylls group.

According to Konopka (Konopka 2004), lutein, the most prevalent carotenoid in cereal grains, has a concentration range of 1.48-1.80 ppm in the kernel and 1.47-2.11 ppm in wheat flour. This suggests that lutein is more concentrated in the endosperm. Durum wheat contains higher levels of lutein (1.69-2.75 ppm) compared to bread wheat (0.13-2.23 ppm). On the other hand, ß-carotene is more abundant in the grain (0.04-0.25 ppm) than in the flour, where it is often absent due to the milling process that removes the bran, where ß-carotene is primarily concentrated. In durum wheat, ß-carotene levels are typically around 5.44 ppm in whole flour and 5.16 ppm in semolina. Overall, the average carotenoid concentration in wheat kernels is 1.88 ppm, with a slightly lower concentration of 1.78 ppm in the flour (Konopka 2004).

Protein in wheat

In wheat grains, protein content may range between 10% - 18% of the total dry matter (Wieser 2020). Even though proteins represent less than 15% of the mature wheat seeds, they have been extensively studied, especially the ones involved in the gluten production, which is

responsible in determining dough viscoelasticity affecting the quality of the derived product as bread or pasta. For this reason, seed storage protein quantity and quality are not only important for their nutritional value, but also for their commercial significance. Wheats, with a protein content higher than 13%, produced satisfying quality products, while genotypes with a grain protein content less than 11% gave poor quality products (Melash & Ábrahám 2022). Protein content is an important parameter in grain quality evaluation. Wheat storage proteins can be classified based on their solubility in two groups: gluten and non-gluten proteins (Fig. 2).

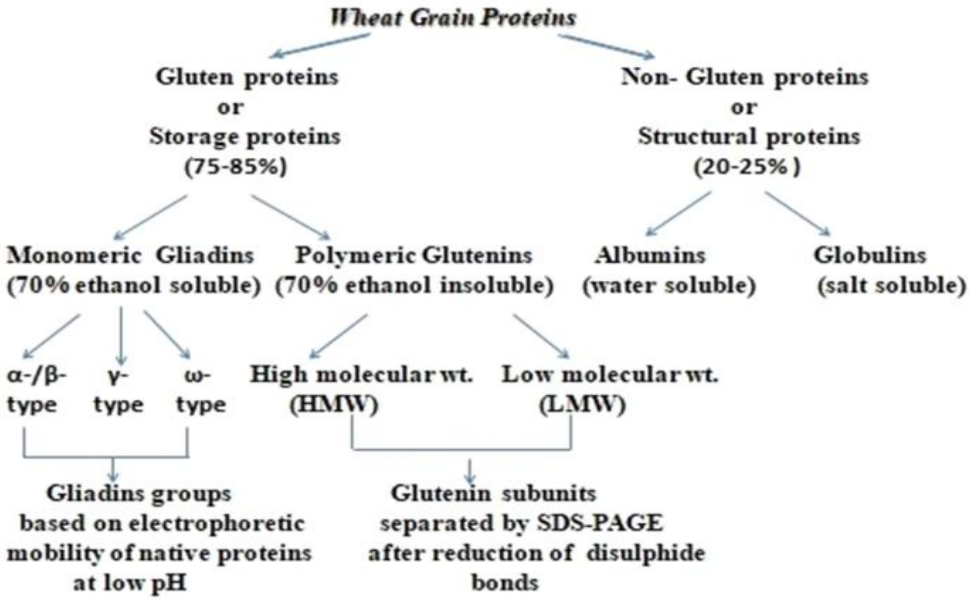


Figure 2 Classification of wheat proteins based on their solubility: gluten and non-gluten proteins.

Non-gluten proteins

Albumins and globulins, defined as the non-gluten proteins, represent 20–25% of the total grain protein in cereals. As revied by Sharma

(Sharma 2020), most of them are monomeric (Merlino 2009) and show high content of lysine, methionine and triptophane (Žilić 2011). Despite these proteins have molecular weights usually below 25 KDa, a significant fraction has higher molecular weights, between 60–70 KDa (Žilić 2011). It has also been shown that non-gluten proteins significantly affect the texture and crumb properties of the bread (Caballero 2007).

Gluten proteins

Gluten proteins are also known as prolamins as they have a proportion of proline and aspartic acid aminoacids as their structural components (Barak 2015). Gliadins and glutenins, the so-called gluten proteins, cross link originating the matrix of the wheat gluten. Basically, gluten, which is highly affected by the type and amount of gluten proteins, is responsible for determining the different rheological properties of dough and bread quality (Branković 2018, Kiszonas & Morris 2018, Fleitas 2018). Gluten quality has been considered to be the most important quality parameter for wheat, because gluten protein (water-insoluble protein complex) determines the elasticity of dough and the end-use properties of cereal based products (Barak 2013; Guo 2018).

A number of baking tests have been carried out to analyze the strong association between bread making and gluten content (Sharma 2020). The identification of all individual gluten components represented an important milestone for molecular breeding research to select genotypes useful for improving gluten quality (Kiszonas & Morris 2018). Proteins are polypeptides consisting of the combination of 20

amino acids, 10 of which are considered to be essential as cannot be synthesized by animals and therefore must be provided by the diet.

The most recent studies on the required amino acid amounts have set new values both for the two aromatic amino acids, tyrosine and phenylalanine, which are biosynthetically related, and for the two sulphur-containing amino acids: methionine, which is truly essential, and cysteine which can be synthesized from methionine (Sharma 2020).

Whole wheat grain and flour are actually deficient only in lysine, while other essential amino acids are present in amounts even higher than required. However, despite the lysin content in wheat shows very high variability, it has been observed that in high protein cereals a decrease in relative lysine content occurs, due to the proportional increase of lysine-poor gluten proteins when an excess of N is available (for example, when applying a fertilizer to increase cereal yield and protein content) and also explains the lower lysine content of white flour (gluten proteins are located in endospermatic starch tissue).

Mixed-Linkage Glucans

Durum wheat is one of the primary sources of fiber for human and animal feed, particularly with respect to arabinoxylan and, to a lesser extent, (1,3;1,4)-β-glucan. Over a decade of intensive research, the role of dietary fibers in reducing the risk of diet-related chronic diseases has been well-established (Fincher 2009 b, Fedacko 2019).

Dietary fiber refers to the edible plant components that resist enzymatic digestion in the small intestine, and includes cellulose, non-cellulosic

polysaccharides (hemicellulose, pectic substances, gums, mucilages), and non-carbohydrate components. It has been demonstrated that the structural properties of non-starch polysaccharides are strongly related to the digestibility, charge, and fermentability of foods (Bader 2018, Zhang D., 2015). The main non-starch polysaccharides in cereals are heteroxylans (predominantly arabinoxylans) and (1,3;1,4)-β-glucans, both of which are long and highly asymmetric molecules, thus forming high-viscosity solutions. The solubility of these polysaccharides is influenced by the degree of arabinosyl substitution in the xylan backbone for arabinoxylans, and by the distribution of the (1,3) and (1,4) linkages for (1,3;1,4)-β-glucans.

The (1,3;1,4)-β-D-glucans are prevalent in the cell walls of *Poaceae* and their fine structure is crucial for the functionality of plant cell walls (Burton 2010). These polysaccharides are unbranched and unsubstituted, consisting of β-D-glucopyranosil monomers linked via (1,3)- and (1,4)-linkages (see Figure 3).

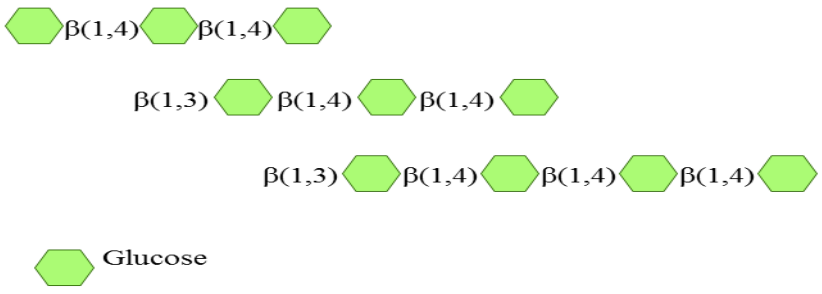


Figure 3 Mixed linkage glucans backbone schematic structure.

Among grasses, grains such as barley (*Hordeum vulgare*), oat (*Avena sativa*), and rye (*Secale cereale*) are particularly rich in (1,3;1,4)-β-D-glucans, while wheat, rice (*Oryza sativa*), and maize (*Zea mays*) contain lower levels of this polysaccharide (Fincher 2009 b).

The proportion of (1,4)- to (1,3)- linkages differs across species. For wheat, this ratio ranges from 3.0:1 to 4.5:1; in barley, it is between 2.9:1 and 3.4:1; rye averages about 2.7:1; and oats have a ratio from 1.8:1 to 2.3:1 (Fincher & Stone 2004). This specific arrangement of linkages leads to an irregular distribution of the polymer chains. Consequently, the molecules do not align over long distances and remain soluble when their degree of polymerization (DP) exceeds 1000 (Woodward 1983). This asymmetrical structure plays a role in the polysaccharide's function as a gel-like matrix component of the cell wall, allowing it to provide support while remaining flexible, pliable, and porous enough to facilitate the transfer of water, nutrients, and other small molecules during developmental stages (Fincher 2009a). Additionally, this asymmetry contributes to the high viscosity of (1,3;1,4)-β-D-glucan and its positive effects on human health and nutrition (Brennan & Cleary 2005). The significance of (1,3;1,4)-β-D-glucan structures lies in the molecular mechanisms proposed for their biosynthesis, which must take into account the chemical and physicochemical properties of their final form in the cell wall (Fincher 2009a).

Burton (Burton 2010) proposed a two-phase assembly system: the first phase entails the formation of glucose polymers (cellodextrins) associated with macromolecules linked to the Golgi membrane. This

phase involves the synthesis of (1,4)-β-oligoglucosidases by cellulose synthase-like *CslF* or *CslH* enzymes. In the second phase of the model, the membrane-bound (1,4)-β-oligoglucosidases are probably transferred to the plasma membrane, where they interact with other enzymes (such as callose synthase, xyloglucan endotransglycosylase, or *CslF* isoenzyme) and are connected via (1,3)-β-linkages (see Figure 4).

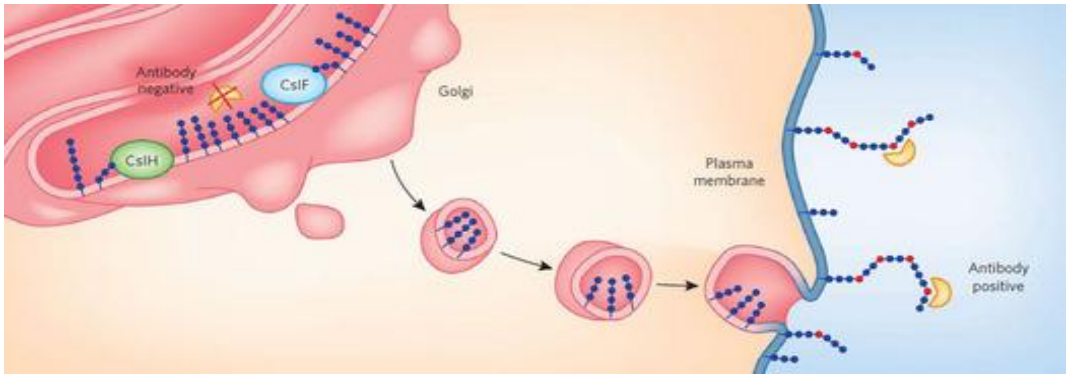


Figure.4 Two-phase mechanism for (1,3;1,4)-β-D-glucan assembly (Burton et al. 2010b)

Physical-chemical and functional property of (1,3;1,4) -β-D-glucan

The fine structure of (1,3;1,4)-β-D-glucans, as discussed earlier, plays a crucial role in the overall conformation and biological functionality of glucans. The molecular kinks introduced by the (1,3)-β-glucosyl residues become more evenly distributed within the polysaccharide, resulting in decreased solubility and reduced suitability for gel èformation in the wall's matrix phase. This irregular conformation of cell wall (1,3;1,4)-β-D-glucans seems to be a characteristic unique to the polysaccharides found in the *Poaceae* family. Within this family notable variations exist in the fine structures of these glucans. For instance, the glucans present in wheat starchy endosperm are

significantly less soluble compared to those found in barley and oat grains (Fincher 2004). The degree of conformational regularity or irregularity in (1,3;1,4)-β-D-glucans dictates their properties and physicochemical behavior in the cell wall matrix. As mentioned previously, the irregularly spaced kinks contribute to an asymmetrical shape that does not align or aggregate easily into fibrils. Instead, this structure remains soluble and can form limited intermolecular junction zones (Fincher 2009a), creating a gel-like substance that provides some structural support while also maintaining flexibility and porosity. These interactions within the cell wall can be affected by other factors, such as connections with various polysaccharides or proteins. In substituted wall polysaccharides, the enzymes responsible for synthesizing the backbone may interact with those that add monosaccharide or oligosaccharide substituents. Recent findings highlight that different isoforms of the same family can guide the synthesis of wall polysaccharides with varying structures and, consequently, distinct physicochemical properties (Burton 2010).

In the *Poaceae*, (1,3;1,4)-β-D-glucans are predominantly found in actively growing cells and are largely absent from mature tissues (Buckeridge 2004, Carpita 2001, Gibeaut 2005). They are particularly present in the expanding cells of organs like the coleoptile (Gibeaut & Carpita 1993, Gibeaut 2005) and also in some vascular and fiber cells of leaves, suggesting a structural function in secondary cell walls (Trethewey 2005).

In the starchy endosperm walls, where secondary thickening is not observed, (1,3;1,4)-β-D-glucans can be found in high concentrations, contributing up to 18% of the total glucose stored in the grain (Morrall & Briggs 1978). Therefore, they play a vital role as storage polysaccharides that are mobilized during grain germination.

Due to their solubility and viscosity in aqueous environments, (1,3;1,4)-β-D-glucans also enhance the functioning of the human digestive system. They are recognized not only as valuable dietary fiber components but also for their potential in preventing and reducing the risk of various diseases (Fox & Henry 1995, Topping 2007). Specifically, they contribute to decrease bowel transit time, alleviate constipation, lower blood cholesterol levels (Gallaher & Plate 2005), reduce the risk of colorectal cancer (Bingham 2000), and manage blood glucose levels in diabetes (Bornet 1987). Additionally, there is significant potential to utilize (1,3;1,4)-β-D-glucans as functional food ingredients to enhance the sensory attributes of low-fat dairy products (Brennan & Cleary 2005).

The importance of genotyping and phenotyping for association mapping

The plant phenotype encompasses all morphological, physiological, and biochemical characteristics that reflect a plant's structure, composition, and growth (Johannsen 2014, Fiorani &Schurr 2013). It includes agronomic traits such as size, structure, and color, as well as physiological aspects that evolve throughout development. In contrast,

genes are nucleotide sequences that encode polypeptides or functional RNA molecules and serve as the fundamental genetic units controlling trait expression. Alleles are alternative forms of a gene located at the same position on a pair of homologous chromosomes, each regulating different variants of a particular trait. Alleles can be classified as dominant or recessive, and they produce functional RNA or proteins that determine whether a trait will be dominant or recessive. The genotype, which is the complete set of genes inherited from both parents, represents a plant's genetic print. The phenotype of a plant is shaped by both its genotype and environmental factors (Rebolledo 2016).

Moreover, the expression of alleles is influenced by environmental factors such as temperature, light, and soil (Honsdorf 2013), meaning that dominant traits may only appear under specific environmental conditions. As a result, plant phenotypes are the product of a dynamic, three-dimensional (3D) spatiotemporal interaction between the genotype and the environment. This interaction leads to a variety of phenotypic outcomes due to the selective expression of genetic information under different environmental conditions.

Methods that link phenotypes to genotypes, such as quantitative trait locus (QTL) mapping, candidate-gene association studies, and genome-wide association studies (GWAS), are widely used to investigate various aspects of plant architecture, development, and environmental responses (Mir 2019). GWAS, for instance, provides high-resolution genetic data and is particularly effective at identifying small-effect

genes or QTL across the entire genome. The integration of GWAS with phenotyping has significantly advanced our understanding of plant growth and has enhanced crop breeding efforts.

Association mapping

One of the primary goals of molecular genetics is to identify and isolate genes that control important traits. Three main strategies are predominantly reported in the literature for cloning a gene of interest: the positional cloning method (Rommens 1989, Tanksley 1993), insertional mutagenesis (Bechtold 1993), and, lastly, the candidate gene approach (de Vienne 1999).

The first two methods have been used to identify major genes, i.e., genes that have the greatest influence on phenotypic variation, while the candidate gene approach is used for genes whose biological function is assumed to be known. The idea is to propose a gene with a known sequence or function in closely related species for the characterization or cloning of Mendelian traits or quantitative trait loci (QTL). Candidate genes may be structural genes or genes involved in the regulation of metabolic processes. The principle underlying this approach is that a molecular polymorphism within the candidate gene corresponds to phenotypic variation.

Several agronomically important traits exhibit quantitative inheritance, often resulting from the action of multiple genes combined with environmental influences. Due to the large number of genes defining complex traits, their partial effects on phenotypic variation, and their

imprecise localization on genetic maps, the candidate gene approach is better suited for QTL characterization compared to positional cloning or insertional mutagenesis.

In plant genetics, the most common strategy for identifying candidate genes is to look for co-segregation between the candidate gene and loci that influence the trait. Several studies conducted association analysis between molecular polymorphisms in the candidate gene and variation in the trait of interest. Final validation of candidate genes can be performed through physiological analyses, genetic transformations, and/or sexual complementation.

The relationship between SNPs (Single Nucleotide Polymorphisms) and GWAS (Genome-Wide Association Studies) is crucial in modern genetics, particularly for identifying genetic loci associated with complex traits such as β-glucan content. As previously mentioned, SNPs are genetic variations at the single nucleotide level, widely distributed across the genome and commonly present in all populations. These variants serve as molecular markers in GWAS to map traits of interest.

SNPs have high genomic density and relatively low cost compared to other markers, such as microsatellites (SSRs), and are particularly advantageous for conducting large-scale association studies with hundreds of thousands of markers, enabling comprehensive genome coverage (McCharty 2008). In a GWAS, SNP genotyping is correlated with phenotypic traits using statistical methods such as linear regression or mixed linear models (MLM), with the goal of identifying genetic loci

that are significantly associated with phenotypic variability, thereby enhancing our understanding of the genetic basis of diseases, agronomic traits, or other biological characteristics (Visscher 2017).

GWAS, through the use of SNPs, has facilitated significant advancements in the field of complex trait genetics. Several studies have reported loci associated with quantitative traits (QTLs) through GWAS analysis, such as protein content (Zhang N., 2015), yield (Marathi 2012), and fiber content (Shu & Rasmussen 2014).

Candidate genes for β-glucan

The biosynthesis of β-glucan in cereal crops has garnered significant research interest, particularly concerning the functions of associated genes and enzymes, and the effects of mutations on β-glucan content across various grains. β-glucan synthesis in grasses is catalyzed by members of glycosyltransferase family 2 that use *uridine diphosphate glucose* (UDP-Glc) for the synthesis process (Bain 2020). The genes responsible for β-glucan synthesis share sequence homology with cellulose synthases (*CesAs*) and belong to the *Cellulose synthase-like* (*Csl*) sub-families F, H, and J (Burton 2006; Doblin 2009; Farrokhi 2006; Little 2019). Notably, the *CslF* and *CslH* groups were found exclusive to monocotyledons and are believed to either directly or indirectly regulate the accumulation and fine structure of β-glucan in the grain and other parts of the plant (Burton 2011; Farrokhi 2006)

(Figure 5). Mutations of *CslF6* in barley, rice, and wheat were found

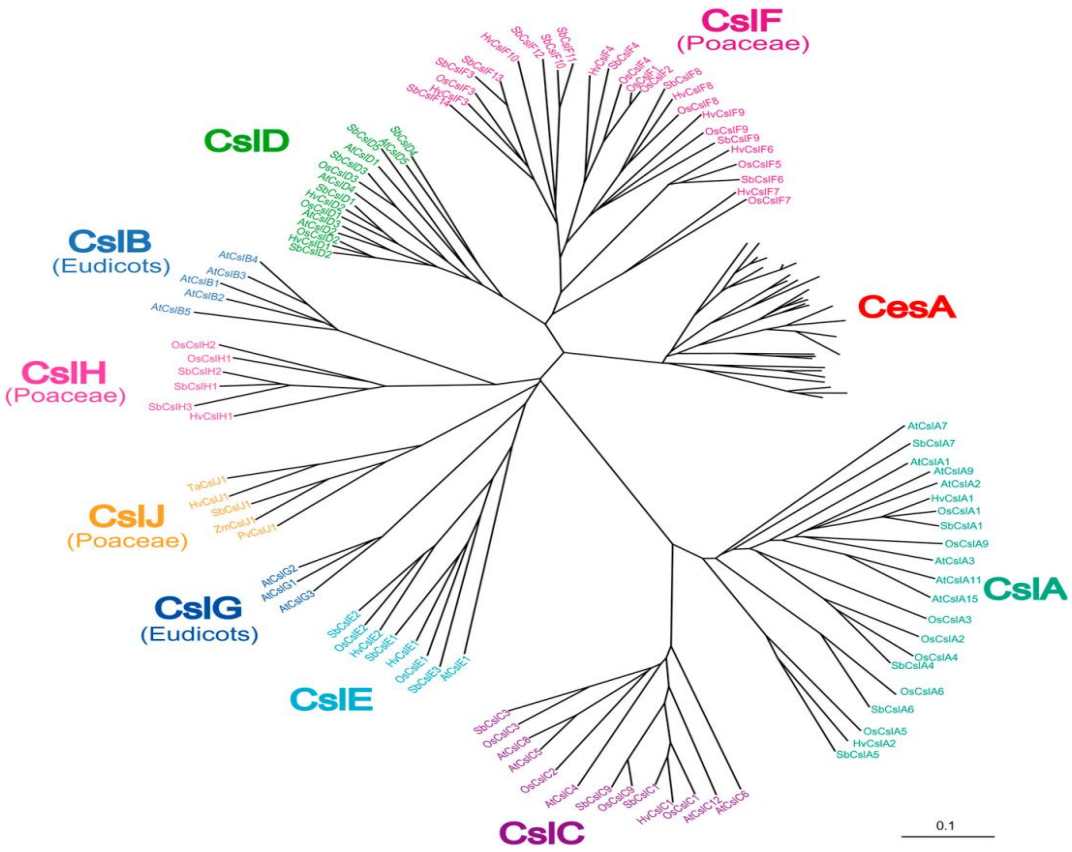


Figure 5 Phylogenetic tree of the CesA superfamily in higher plants (Fincher et al., 2009).

associated with reductions in β-glucan content, but these reductions generally had only moderate effects on vegetative growth and were well-tolerated (Cory 2012; Hu 2014; Nemeth 2010; Taketa 2012; Tonooka 2009; Wong 2015). In wheat, it was observed that targeted RNA knockdowns of *CslF6*, driven by an endosperm-specific promoter, resulted in an 30-53% decrease in β-glucan content in whole grain flour (Nemeth 2010). In another study, knockout mutants of *CslF6* in rice exhibited a 97% reduction in MLG levels in coleoptiles, with undetectable levels in other tissues (Vega-Sanchez 2012). Compared to other cereal species, limited data on biosynthetic pathway of β-glucan is available for durum wheat, its wild relatives, and their hybrid.

Studies have highlighted the probability that several genomic regions associated with β-glucan accumulation in wheat seeds are predominantly located within chromosomes 2 and 7. Similarly, a significant QTL on chromosome 7H of barley, responsible for 39% of the variation in β-glucan content in barley seeds (Li 2008), further supports the influence of chromosome 7 on β-glucan content in wheat seeds. This was evident when chromosome 7H from barley was introduced into the wheat genome, resulting in a substantial increase in β-glucan content in wheat seeds (Colasuonno 2020). Furthermore, (Manickavelu 2011) identified QTLs for β-glucan content on chromosomes 3A, 1B, 5B, and 6D in wheat through RILs derived from a cross between Chinese Spring and spelt wheat. Similarly, (Marcotuli 2016) used a 230-genotype array of tetraploid wheat to predict QTL for β-glucan content. These results revealed QTL on chromosomes 1A, 2A, 2B, 5B, and 7A, also identifying a subgroup of putative genes that exert direct or indirect influences on β-glucan content in seeds. This focus on chromosomes 2 and 7 reinforces their crucial role in β-glucan accumulation, consistent with previous research.

Among the wide range of cereal crops, wheat (*Triticum aestivum* L.) plays a crucial role as a global staple, significantly contributing to daily caloric and protein intake (FAO 2020). However, the content of soluble β-glucans in wheat grains is relatively low, especially when compared to the higher concentrations found in oat and barley grains (Beresford & Stone 1983, Havrlentová & Kraic 2006, Collins 2010). Despite concerted efforts, the primary and secondary gene pools of bread wheat

have demonstrated limited genetic diversity that is conducive to the expression of high β-glucan content (Pritchard 2011, Marcotuli 2017).

Furthermore, neither bread wheat (T. aestivum L. 2n = 6x = 42) nor its closest relatives exhibit genetic diversity for high β-glucan levels in the grains (Friebe 1996). This lack of genetic diversity extends to both the primary and secondary gene pools, restricting the feasibility of establishing a breeding program aimed at increasing the β-glucan content in cultivated wheat varieties. This conclusion is based on comprehensive studies involving over 500 wheat germplasms of diverse origins, including primitive, elite, and synthetic lines of hexaploid, tetraploid, and diploid wheat, as well as related species (Pritchard 2011, Marcotuli 2019).

To improve the quality of wheat grains, a potential strategy involves incorporating genes from more distantly related relatives, which may belong to the tertiary wheat gene pool and include both wild and cultivated species. This can be achieved through chromosome engineering and interspecific hybridization (Cseh 2011, 2013, Rakszegi 2017, Danilova 2018, Türkösi 2018). Among the members of the tertiary gene pool, barley exhibits the highest β-glucan concentration (up to 11%), making it a promising candidate for interspecific hybridization. Notably, certain species of *Aegilops*, which are also part of the wheat tertiary gene pool, display significant β-glucan contents (up to 7.1%) in their grains. These species include diploid types such as *Ae. Umbellulata* (Zhuk.), as well as various tetraploid species containing U and C genomes (Rakszegi 2017, Marcotuli 2019). In

contrast, most bread wheat (*T. aestivum* L.) cultivars and durum wheat (*Triticum turgidum* L. ssp. *durum*) exhibit lower levels of β-glucan in their grains (Marcotuli 2016, Mohebbi 2018).

The practice of utilizing genetic traits from distantly related species to enhance the genetic robustness and resilience of wheat has been previously demonstrated, particularly in the context of conferring resistance to various biotic stresses (Friebe 1996, 2000, Marais 2005, Qi 2007, Kuraparthy 2007a, b, Schneider 2008).

References

1. Arora, S., Cheema, J., Poland, J., Uauy, C., & Chhuneja, P. (2019). Genome-wide association mapping of grain micronutrients concentration in Aegilops tauschii. *Frontiers in plant science*, *10*, 54. https://doi.org/10.3389/fpls.2019.00054
2. Bader Ul Ain, H., Saeed, F., Ahmad, N., Imran, A., Niaz, B., Afzaal, M., Imran, M., Tufail, T., & Javed, A. (2018). Functional and health-endorsing properties of wheat and barley cell wall's non-starch polysaccharides. *International Journal of Food Properties*, *21*(1), 1463-1480.https://doi.org/10.1080/10942912.2018.1489837
3. Bain, M., van de Meene, A., Costa, R., & Doblin, M. S. (2020). Characterisation of Cellulose Synthase Like F6 (CslF6) Mutants Shows Altered Carbon Metabolism in beta-D-(1,3;1,4)-Glucan Deficient Grain in Brachypodium distachyon. *Front Plant Sci, 11*, 602850. https://doi.org/10.3389/fpls.2020.602850.
4. Barak, S., Mudgil, D., & Khatkar, B. S. (2015). Biochemical and functional properties of wheat gliadins: a review. *Critical reviews in food science and nutrition*, *55*(3), 357-368. https://doi.org/10.1080/10408398.2012.654863
5. Barański, M., Lacko-Bartošová, M., Rembiałkowska, E., & Lacko-Bartošová, L. (2020). The effect of species and cultivation year on phenolic acids content in ancient wheat. *Agronomy*, *10*(5), 673. https://doi.org/10.3390/agronomy10050673
6. Bechtold, N. (1993). In planta Agrobacterium-mediated gene transfer by infiltration of adult Arabidopsis thaliana plants. CR Acad. *Sci. Paris, Life Sci.*, *316*, 1194-1199.Beresford, G., and Stone, B. A. (1983). (1→3), (1→4)-β-D-glucan content of Triticum grains. *J. Cereal Sci.* 1, 111–114. https://doi:10.1016/S0733-5210(83)80028-9

7. Biesiekierski, J. R. (2017). What is gluten?. *Journal of gastroenterology and hepatology*, *32*, 78-81. https://doi.org/10.1111/jgh.13703
8. Bingham, S. A. (2000). Diet and colorectal cancer prevention. *Biochemical Society Transactions*, *28*(2), 12-16. https://doi.org/10.1042/bst0280012
9. Bornet, F.R., Costagliola, D., Rizkalla, S.W., Blayo, A., Fontvieille, A.M., Haardt, M.J., Letanoux, M., Tchobroutsky, G. And Slama, G. (1987). Insulinemic and Glycemic Indexes of Six Starch-Rich Foods Taken Alone and in a Mixed Meal by Type 2 Diabetics. *The American Journal of Clinical Nutrition,* 45(3), 588-595. https://doi.org/10.1111/j.1467-7652.2010.00532.x.
10. Branković, G., Pajić, V., Zivanović, T., Dodig, D., Kandić, V., Knežević, D., & Đurić, N. (2018). Genetic parameters of Triticum aestivum and Triticum durum for technological quality properties in Serbia. *Zemdirbyste-agriculture*, *105*(1), 39-48. DOI 10.13080/z-a.2018.105.006
11. Brennan, C. S., & Cleary, L. J. (2005). The potential use of cereal (1→ 3, 1→ 4)-β-D-glucans as functional food ingredients. *Journal of cereal science*, *42*(1), 1-13. https://doi.org/10.1016/j.jcs.2005.01.002
12. Buckeridge, M. S., Rayon, C., Urbanowicz, B., Tiné, M. A. S., & Carpita, N. C. (2004). Mixed linkage (1→ 3),(1→ 4)-β-d-glucans of grasses. *Cereal Chemistry*, *81*(1), 115-127. https://doi.org/10.1094/CCHEM.2004.81.1.115
13. Burton, R. A., Collins, H. M., Kibble, N. A. J., Smith, J. A., Shirley, N. J., Jobling, S. A., Henderson M., Singh, R. R., Pettolino, F., Wilson S. M., Bird. A. R., Topping, D. L., Bacic, A,. Fincher, G. B. (2011). Over-expression of specific HvCslF cellulose synthase-like genes in

transgenic barley increases the levels of cell wall (1,3;1,4)-β-d-glucans and alters their fine structure. *Plant Biotechnology Journal, 9*(2), 117-135. https://doi.org/10.1111/j.1467-7652.2010.00532.x.

14. Burton, R. A., Gidley, M. J., & Fincher, G. B. (2010). Heterogeneity in the chemistry, structure and function of plant cell walls. *Nature chemical biology*, *6*(10), 724-732.
15. Burton, R. A., Wilson, S. M., Hrmova, M., Harvey, A. J., Shirley, N. J., Medhurst, A., Stone, B. A., Newbigin, E. J., Bacic, A., Fincher, G. B. (2006). Cellulose synthase-like *CslF* genes mediate the synthesis of cell wall (1,3;1,4)-β-D-glucans. *Science, 311*(5769), 1940-1942. https://doi.org/doi:10.1126/science.1122975.
16. Caballero, P. A., Gómez, M., & Rosell, C. M. (2007). Bread quality and dough rheology of enzyme-supplemented wheat flour. *European Food Research and Technology*, *224*, 525-534.
17. Cakmak, I., Ozkan, H., Braun, H. J., Welch, R. M., & Romheld, V. (2000). Zinc and iron concentrations in seeds of wild, primitive, and modern wheats. *Food and nutrition bulletin*, *21*(4), 401-403. https://doi.org/10.1177/156482650002100411
18. Carpita, N. C., Defernez, M., Findlay, K., Wells, B., Shoue, D. A., Catchpole, G., ... & McCann, M. C. (2001). Cell wall architecture of the elongating maize coleoptile. *Plant physiology*, *127*(2), 551-565. https://doi.org/10.1104/pp.010146
19. Chhuneja, P., Dhaliwal, H. S., Bains, N. S., & Singh, K. (2006). Aegilops kotschyi and Aegilops tauschii as sources for higher levels of grain iron and zinc. *Plant breeding*, *125*(5), 529-531. https://doi.org/10.1111/j.1439-0523.2006.01223.x
20. Colasuonno, P., Marcotuli, I., Cutillo, S., Simeone, R., Blanco, A., & Gadaleta, A. (2020). Effect of barley chromosomes on the β-glucan

content of wheat. *Genetic Resources and Crop Evolution*, *67*, 561-567. https://doi:10.1007/s10722-019-00829-y

21. Collins, H. M., Burton, R. A., Topping, D. L., Liao, M. L., Bacic, A., & Fincher, G. B. (2010). Variability in fine structures of noncellulosic cell wall polysaccharides from cereal grains: potential importance in human health and nutrition. *Cereal Chemistry*, *87*(4), 272-282. https://doi:10.1094/CCHEM-87-4-0272
22. Cory, A. T., Båga, M., Anyia, A., Rossnagel, B. G., & Chibbar, R. N. (2012). Genetic markers for *CslF6* gene associated with (1,3;1,4)-β-glucan concentration in barley grain. *Journal of Cereal Science, 56*(2), 332-339. https://doi.org/10.1016/j.jcs.2012.02.003
23. Cseh, A., Soós, V., Rakszegi, M., Türkösi, E., Balázs, E., & Molnár-Láng, M. (2013). Expression of HvCslF9 and HvCslF6 barley genes in the genetic background of wheat and their influence on the wheat β-glucan content. *Annals of Applied Biology*, *163*(1), 142-150. https://doi:10.1111/AAB.12043
24. Cseh, A., Kruppa, K., Molnár, I., Rakszegi, M., Doležel, J., & Molnár-Láng, M. (2011). Characterization of a new 4BS. 7HL wheat–barley translocation line using GISH, FISH, and SSR markers and its effect on the β-glucan content of wheat. *Genome*, *54*(10), 795-804. https://doi:10.1139/g11-044
25. Danilova, T. V., Friebe, B., Gill, B. S., Poland, J., & Jackson, E. (2018). Development of a complete set of wheat–barley group-7 Robertsonian translocation chromosomes conferring an increased content of β-glucan. Theoretical and Applied Genetics, 131, 377-388. https://doi:10.1007/s00122-017-3008-z
26. De Santis, M. A., Giuliani, M. M., Giuzio, L., De Vita, P., Lovegrove, A., Shewry, P. R., & Flagella, Z. (2017). Differences in gluten protein

composition between old and modern durum wheat genotypes in relation to 20th century breeding in Italy. *European Journal of Agronomy*, *87*, 19-29. https://doi.org/10.1016/j.eja.2017.04.003

27. De Vienne, D., Leonardi, A., Damerval, C., & Zivy, M. (1999). Genetics of proteome variation for QTL characterization: application to drought-stress responses in maize. *Journal of Experimental Botany*, *50*(332), 303-309. https://doi.org/10.1093/jxb/50.332.303
28. De Vita, P., Nicosia, O. L. D., Nigro, F., Platani, C., Riefolo, C., Di Fonzo, N., & Cattivelli, L. (2007). Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. *European Journal of Agronomy*, *26*(1), 39-53. https://doi.org/10.1016/j.eja.2006.08.009
29. Dempewolf, H., Eastwood, R. J., Guarino, L., Khoury, C. K., Müller, J. V., & Toll, J. (2014). Adapting agriculture to climate change: a global initiative to collect, conserve, and use crop wild relatives. *Agroecology and Sustainable Food Systems*, *38*(4), 369-377. https://doi.org/10.1080/21683565.2013.870629
30. Doblin, M. S., Pettolino, F. A., Wilson, S. M., Campbell, R., Burton, R. A., Fincher, G. B., Newbigin, E., Bacic, A. (2009). A barley cellulose *synthase-like CSLH* gene mediates (1,3;1,4)-β-D-glucan synthesis in transgenic *Arabidopsis*. *Proceedings of the National Academy of Sciences, 106*(14), 5996-6001. https://doi.org/doi:10.1073/pnas.0902019106.
31. FAO, (2020). Food and Agriculture Organisation. Available at: https://www.fao.org/home/en)
32. Farrokhi, N., Burton, R. A., Brownfield, L., Hrmova, M., Wilson, S. M., Bacic, A., & Fincher, G. B. (2006). Plant cell wall biosynthesis: Genetic, biochemical and functional genomics approaches to the

identification of key genes. *Plant Biotechnology Journal, 4*(2), 145-167. https://doi.org/10.1111/j.1467-7652.2005.00169.x

33. Fatima, F., McCallum, B. D., Pozniak, C. J., Hiebert, C. W., McCartney, C. A., Fedak, G., You, F. M., & Cloutier, S. (2020). Identification of new leaf rust resistance loci in wheat and wild relatives by array-based SNP genotyping and association genetics. *Frontiers in Plant Science*, *11*, 583738. https://doi.org/10.3389/fpls.2020.583738
34. Fedacko, J., Takahashi, T., Singh, R. B., Pella, D., Chibisov, S., Hristova, K., Pella, D., Elkilany G. N., Tomar, R. S., & Juneja, L. R. (2019). Globalization of diets and risk of noncommunicable diseases. In *The role of functional food security in global health* (pp. 87-107). Academic Press.Fincher G.B. 2009a. Revolutionary Times in Our Understanding of Cell Wall Biosynthesis and Remodeling In the Grasses. Plant Physiol 149: 27-37. https://doi.org/10.1016/B978-0-12-813148-0.00006-2
35. Fincher, G. B. (2009). Exploring the evolution of (1, 3; 1, 4)-β-d-glucans in plant cell walls: comparative genomics can help!. *Current opinion in plant biology*, *12*(2), 140-147. https://doi.org/10.1016/j.pbi.2009.01.002
36. Fincher, G., & Stone, B. (2004). Chemistry of nonstarch polysaccharides. Elsevier Academic Press.
37. Fiorani, F., & Schurr, U. (2013). Future scenarios for plant phenotyping. *Annual review of plant biology*, *64*(1), 267-291. https://doi.org/10.1146/annurev-arplant-050312-120137
38. Fleitas, M. C., Schierenbeck, M., Gerard, G. S., Dietz, J. I., Golik, S. I., Campos, P. E., & Simón, M. R. (2018). How leaf rust disease and its control with fungicides affect dough properties, gluten quality and

loaf volume under different N rates in wheat. *Journal of cereal science*, *80*, 119-127. https://doi.org/10.1016/j.jcs.2018.02.003

39. Fox, G. P., & Henry, R. J. (1995). ANALYSIS OF WATER-SOLUBLE PROTEINS FROM BARLEY BY ION-EXCHANGE HIGH PERFORMANCE LIQUID CHROMATOGRAPHY. *Journal of the Institute of Brewing*, *101*(3), 181-185. https://doi.org/10.1002/j.2050-0416.1995.tb00859.x
40. Friebe, B., Jiang, J., Raupp, W. J., McIntosh, R. A., & Gill, B. S. (1996). Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica*, *91*, 59-87. https://doi:10.1007/bf00035277
41. Friebe, B., Qi, L. L., Nasuda, S., Zhang, P., Tuleen, N. A., & Gill, B. S. (2000). Development of a complete set of Triticum aestivum-Aegilops speltoides chromosome addition lines. *Theoretical and applied genetics*, *101*, 51-58. https://doi:10.1007/s001220051448
42. Gallaher, D. D., & Plate, A. Y. A. (2005, March). Reduction in liver cholesterol by a beta-glucan concentrate in rats: correlation with intestinal contents viscosity. In *FASEB JOURNAL* (Vol. 19, No. 4, pp. A90-A90). 9650 ROCKVILLE PIKE, BETHESDA, MD 20814-3998 USA: FEDERATION AMER SOC EXP BIOL.
43. Gibeaut, D. M., & Carpita, N. C. (1993). Synthesis of (1--> 3),(1--> 4)-beta-D-glucan in the Golgi apparatus of maize coleoptiles. *Proceedings of the National Academy of Sciences*, *90*(9), 3850-3854. https://doi.org/10.1073/pnas.90.9.3850
44. Gibeaut, D. M., Pauly, M., Bacic, A., & Fincher, G. B. (2005). Changes in cell wall polysaccharides in developing barley (Hordeum vulgare) coleoptiles. *Planta*, *221*, 729-738.

45. Graham, R. D., Welch, R. M., & Bouis, H. E. (2001). Addressing micronutrient malnutrition through enhancing the nutritional quality of staple foods: principles, perspectives and knowledge gaps. https://doi.org/10.1016/S0065-2113(01)70004-1
46. Graham, R. D., Welch, R. M., Saunders, D. A., Ortiz-Monasterio, I., Bouis, H. E., Bonierbale, M., De Haan, S., Burgos, G., Thiele, G., Liria, R., Meisner, C. A., Beebe, S. E., Potts, M. J., Kadian, M., Hobbs, P. R., Gupta, R., & Twomlow, S. (2007). Nutritious subsistence food systems. *Advances in agronomy*, *92*, 1-74. https://doi.org/10.1016/S0065-2113(04)92001-9
47. Guo, X., Sun, X., Zhang, Y., Wang, R., & Yan, X. (2018). Interactions between soy protein hydrolyzates and wheat proteins in noodle making dough. *Food Chemistry*, *245*, 500-507.
48. Havrlentová, M., & Kraic, J. Á. N. (2006). Content of β-D-glucan in cereal grains.
49. Honsdorf, N., March, T. J., Berger, B., Tester, M., & Pillen, K. (2014). High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. *PloS one*, *9*(5), e97047. https://doi.org/10.1371/journal.pone.0097047
50. Hu, G., Burton, C., Hong, Z., & Jackson, E. (2014). A mutation of the cellulose-synthase-like (CslF6) gene in barley (Hordeum vulgare L.) partially affects the β-glucan content in grains. *Journal of Cereal Science*, *59*(2), 189-195. https://doi.org/10.1016/j.jcs.2013.12.009
51. Iizumi, T., Ali-Babiker, I. E. A., Tsubo, M., Tahir, I. S., Kurosaki, Y., Kim, W., Gorafi, Y. S. A., Idris, A. A. M., & Tsujimoto, H. (2021). Rising temperatures and increasing demand challenge wheat supply in Sudan. *Nature Food*, *2*(1), 19-27. https://doi.org/10.1038/s43016-020-00214-4

52. Isidro, J., Alvaro, F., Royo, C., Villegas, D., Miralles, D. J., & García del Moral, L. F. (2011). Changes in duration of developmental phases of durum wheat caused by breeding in Spain and Italy during the 20th century and its impact on yield. *Annals of Botany*, *107*(8), 1355-1366. https://doi.org/10.1093/aob/mcr063
53. Jefremova, O., Radenkovs, V., Kunkulberga, D., & Klava, D. (2015). Technological properties of dough from wheat flour and fermented bran. *Chemical Technology*, *66*(1), 13-18. https://doi.org/10.5755/j01.ct.66.1.12360
54. Johannsen, W. (2014). The genotype conception of heredity. Int J Epidemiol, 43 (4):989-1000.
55. Kilian, B., Dempewolf, H., Guarino, L., Werner, P., Coyne, C., & Warburton, M. L. (2021). Crop Science special issue: Adapting agriculture to climate change: A walk on the wild side. *Crop Science*, *61*(1), 32-36. https://doi.org/10.1002/csc2.20418
56. Kilian, B., Martin, W., & Salamini, F. (2010). Genetic diversity, evolution and domestication of wheat and barley in the Fertile Crescent. *Evolution in action: case studies in adaptive radiation, speciation and the origin of biodiversity*, 137-166.
57. Kiszonas, A. M., & Morris, C. F. (2018). Wheat breeding for quality: A historical review. *Cereal Chemistry*, *95*(1), 17-34. https://doi.org/10.1094/CCHEM-05-17-0103-FI
58. Konopka, I., Kozirok, W., & Rotkiewicz, D. (2004). Lipids and carotenoids of wheat grain and flour and attempt of correlating them with digital image analysis of kernel surface and cross-sections. *Food Research International*, *37*(5), 429-438. https://doi.org/10.1016/j.foodres.2003.12.009

59. Kucek, L. K., Veenstra, L. D., Amnuaycheewa, P., & Sorrells, M. E. (2015). A grounded guide to gluten: how modern genotypes and processing impact wheat sensitivity. *Comprehensive Reviews in Food Science and Food Safety*, *14*(3), 285-302. https://doi.org/10.1111/1541-4337.12129

60. Kuraparthy, V., Sood, S., Chhuneja, P., Dhaliwal, H. S., Kaur, S., Bowden, R. L., & Gill, B. S. (2007). A cryptic wheat–Aegilops triuncialis translocation with leaf rust resistance gene Lr58. *Crop Science*, *47*(5), 1995-2003. https://doi:10.2135/CROPSCI2007.01.0038

61. Kuraparthy, V., Chhuneja, P., Dhaliwal, H. S., Kaur, S., Bowden, R. L., & Gill, B. S. (2007). Characterization and mapping of cryptic alien introgression from Aegilops geniculata with new leaf rust and stripe rust resistance genes Lr57 and Yr40 in wheat. *Theoretical and Applied Genetics*, *114*, 1379-1389. https://doi:10.1007/s00122-007-0524-2

62. Laddomada, B., Caretto, S., & Mita, G. (2015). Wheat bran phenolic acids: Bioavailability and stability in whole wheat-based foods. *Molecules*, *20*(9), 15666-15685. https://doi.org/10.3390/molecules200915666

63. Li, J., Båga, M., Rossnagel, B. G., Legge, W. G., & Chibbar, R. N. (2008). Identification of quantitative trait loci for β-glucan concentration in barley grain. *Journal of Cereal Science*, *48*(3), 647-655. https://doi:10.1016/j.jcs.2008.02.004

64. Li, L., Niu, Y., Ruan, Y., DePauw, R. M., Singh, A. K., & Gan, Y. (2018). Agronomic advancement in tillage, crop rotation, soil health, and genetic gain in durum wheat cultivation: A 17-year Canadian story. *Agronomy*, *8*(9), 193. https://doi.org/10.3390/agronomy8090193

65. Li, M., Dong, L., Li, B., Wang, Z., Xie, J., Qiu, D., Li, Y., Shi, W., Yang, L., Wu, Q., Chen, Y., Lu, P., Guo, G., Zhang, H., Zhang, P., Zhu, K., Li, Y., Zhang, Y., Wang, R., Yuan, C., Liu, W., Yu, D., Luo, M., Fahima, T., Nevo, E., Li, H., & Liu, Z. (2020). A CNL protein in wild emmer wheat confers powdery mildew resistance. *New Phytologist*, *228*(3), 1027-1037. https://doi.org/10.1111/nph.16761
66. Little, A., Lahnstein, J., Jeffery, D. W., Khor, S. F., Schwerdt, J. G., Shirley, N. J., Hooi, M., Xing, X., Burton, R. A., & Bulone, V. (2019). A novel (1, 4)-β-linked glucoxylan is synthesized by members of the cellulose synthase-like F gene family in land plants. *ACS Central Science*, *5*(1), 73-84. https://doi.org/10.1021/acscentsci.8b00568
67. Liyana-Pathirana, C. M., & Shahidi, F. (2006). Importance of insoluble-bound phenolics to antioxidant properties of wheat. *Journal of agricultural and food chemistry*, *54*(4), 1256-1264. https://doi.org/10.1021/jf052556h
68. Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., Casadio, R., Akhunov, E., Scalabrin, S., Vendramin, V., Ammar, K., Blanco, A., Desiderio, F., Distelfeld, A., Dubcovsky, J., Fahima, T., Faris, J., Korol, A., Massi, A., Mastrangelo, A. M., Morgante, M., Pozniak, C., N'Diaye, A., Xu, S., & Tuberosa, R. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. Plant biotechnology journal, 13(5), 648-663. https://doi.org/10.1111/pbi.12288
69. Manickavelu, A., Kawaura, K., Imamura, H., Mori, M., & Ogihara, Y. (2011). Molecular mapping of quantitative trait loci for domestication traits and β-glucan content in a wheat recombinant inbred line

population. *Euphytica*, *177*, 179-190. https://doi:10.1007/s10681-010-0217-9

70. Marais, G. F., McCallum, B., Snyman, J. E., Pretorius, Z. A., & Marais, A. S. (2005). Leaf rust and stripe rust resistance genes Lr54 and Yr37 transferred to wheat from Aegilops kotschyi. *Plant Breeding*, *124*(6), 538-541. https://doi:10.1111/J.1439-0523.2005.01116.X

71. Marathi, B., Guleria, S., Mohapatra, T., Parsad, R., Mariappan, N., Kurungara, V. K., Atwal, S. S., Prabhu, K. V., Singh, N. K., & Singh, A. K. (2012). QTL analysis of novel genomic regions associated with yield and yield related traits in new plant type based recombinant inbred lines of rice (Oryza sativa L.). *BMC plant biology*, *12*, 1-19.

72. Marcotuli, I., Colasuonno, P., Cutillo, S., Simeone, R., Blanco, A., & Gadaleta, A. (2019). β-glucan content in a panel of Triticum and Aegilops genotypes. *Genetic Resources and Crop Evolution*, *66*, 897-907. https://doi:10.1007/s10722-019-00753-1

73. Marcotuli, I., Gadaleta, A., Mangini, G., Signorile, A. M., Zacheo, S. A., Blanco, A., Simeone, R., & Colasuonno, P. (2017). Development of a high-density SNP-based linkage map and detection of QTL for β-glucans, protein content, grain yield per spike and heading time in durum wheat. *International journal of molecular sciences*, *18*(6), 1329. https://doi.org/10.3390/ijms18061329

74. Marcotuli, I., Houston, K., Schwerdt, J. G., Waugh, R., Fincher, G. B., Burton, R. A., Blanco, A, & Gadaleta, A. (2016). Genetic diversity and genome wide association study of β-glucan content in tetraploid wheat grains. *PLoS One*, *11*(4), e0152590. https://doi.org/10.1371/journal.pone.0152590

75. McCarthy, M. I., & Hirschhorn, J. N. (2008). Genome-wide association studies: potential next steps on a genetic journey. *Human molecular genetics*, *17*(R2), R156-R165.
76. Melash, A. A., & Ábrahám, É. B. (2022). Barriers and levers to enhance end-use functional properties of durum wheat (Triticum turgidum L.) grain: An agronomic implication. *Heliyon*, *8*(5).
77. Merlino, M., Leroy, P., Chambon, C., & Branlard, G. (2009). Mapping and proteomic analysis of albumin and globulin proteins in hexaploid wheat kernels (Triticum aestivum L.). *Theoretical and Applied Genetics*, *118*(7), 1321-1337.
78. Mir, R. R., Reynolds, M., Pinto, F., Khan, M. A., & Bhat, M. A. (2019). High-throughput phenotyping for crop improvement in the genomics era. *Plant Science*, *282*, 60-72. https://doi.org/10.1016/j.plantsci.2019.01.007
79. Mohebbi, Z., Homayouni, A., Azizi, M. H., & Hosseini, S. J. (2018). Effects of beta-glucan and resistant starch on wheat dough and prebiotic bread properties. *Journal of food science and technology*, *55*, 101-110. https://doi:10.1007/s13197-017-2836-9
80. Morrall, P., & Briggs, D. E. (1978). Changes in cell wall polysaccharides of germinating barley grains. *Phytochemistry*, *17*(9), 1495-1502. https://doi.org/10.1016/S0031-9422(00)94628-4
81. Motzo, R., Fois, S., & Giunta, F. (2004). Relationship between grain yield and quality of durum wheats from different eras of breeding. *Euphytica*, *140*, 147-154.
82. Nemeth, C., Freeman, J., Jones, H. D., Sparks, C., Pellny, T. K., Wilkinson, M. D., Dunwell J., Andersson, A. A. M., Aman, P., Guillon, F., Saulnier L., Mitchell R. A. C., & Shewry, P. R. (2010). Down-regulation of the CSLF6 gene results in decreased (1, 3; 1, 4)-β-D-

glucan in endosperm of wheat. *Plant physiology*, *152*(3), 1209-1218. https://doi.org/10.1104/pp.109.151712

83. Nsarellah, N., Amri, A., Nachit, M. M., Bouhssini, M. E., Lhaloui, S., & Lorenzoni, C. (2003). New durum wheat with Hessian fly resistance from Triticum araraticum and T. carthlicum in Morocco. *Plant Breeding*, *122*(5), 435-437. https://doi.org/10.1046/j.1439-0523.2003.00871.x

84. Peng, J. H., Sun, D., & Nevo, E. (2011). Domestication evolution, genetics and genomics in wheat. *Molecular Breeding*, *28*, 281-301.

85. Pérez-Gálvez, A., & Mínguez-Mosquera, M. I. (2005). Esterification of xanthophylls and its effect on chemical behavior and bioavailability of carotenoids in the human. *Nutrition Research*, *25*(7), 631-640. https://doi.org/10.1016/j.nutres.2005.07.002

86. Pozniak, C. J., Knox, R. E., Clarke, F. R., & Clarke, J. (2007). Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. *Theoretical and Applied Genetics*, *114*, 525-537.

87. Pritchard, J. R., Lawrence, G. J., Larroque, O., Li, Z., Laidlaw, H. K., Morell, M. K., & Rahman, S. (2011). A survey of β-glucan and arabinoxylan content in wheat. *Journal of the Science of Food and Agriculture*, *91*(7), 1298-1303. https://doi:10.1002/JSFA.4316

88. Qi, L., Friebe, B., Zhang, P., & Gill, B. S. (2007). Homoeologous recombination, chromosome engineering and crop improvement. Chromosome Research, 15, 3-19. https://doi:10.1007/s10577-006-1108-8

89. Rakszegi, M., Molnár, I., Lovegrove, A., Darkó, É., Farkas, A., Láng, L., Bedő Z., Dolžel, J., Shewry, P. (2017). Addition of Aegilops U and M chromosomes affects protein and dietary fiber content of

wholemeal wheat flour. *Frontiers in Plant Science*, *8*, 1529. https://doi:10.3389/fpls.2017.01529

90. Rebolledo, M. C., Peña, A. L., Duitama, J., Cruz, D. F., Dingkuhn, M., Grenier, C., & Tohme, J. (2016). Combining image analysis, genome wide association studies and different field trials to reveal stable genetic regions related to panicle architecture and the number of spikelets per panicle in rice. *Frontiers in plant science*, *7*, 1384. https://doi.org/10.3389/fpls.2016.01384
91. Rommens, J. M., Iannuzzi, M. C., Kerem, B. S., Drumm, M. L., Melmer, G., Dean, M., Rozmahel, R., Cole, J. L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Tsui, L., Riordan, J. R., & Collins, F. S. (1989). Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science*, *245*(4922), 1059-1065. https://DOI: 10.1126/science.2772657
92. Rossini, F., Provenzano, M. E., Sestili, F., & Ruggeri, R. (2018). Synergistic effect of sulfur and nitrogen in the organic and mineral fertilization of durum wheat: Grain yield and quality traits in the Mediterranean environment. *Agronomy*, *8*(9), 189. https://doi.org/10.3390/agronomy8090189
93. Ross-Ibarra, J., Morrell, P. L., & Gaut, B. S. (2007). Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proceedings of the National Academy of Sciences*, *104*(suppl_1), 8641-8648. https://doi.org/10.1073/pnas.0700643104
94. Schneider, A., Molnár, I., & Molnár-Láng, M. (2008). Utilisation of Aegilops (goatgrass) species to widen the genetic diversity of cultivated wheat. *Euphytica*, *163*, 1-19. https://doi:10.1007/s10681-007-9624-y

95. Seib, P. A. (1994). Wheat starch: Isolation, structure and properties. *Journal of Applied Glycoscience*, *41*(1), 49-69. https://doi.org/10.11541/jag1994.41.49

96. Serpen, A., Gökmen, V., Karagoz, A., & Köksel, H. (2008). Phytochemical quantification and total antioxidant capacities of emmer (Triticum dicoccon Schrank) and einkorn (Triticum monococcum L.) wheat landraces. *Journal of agricultural and food chemistry*, *56*(16), 7285-7292. https://doi.org/10.1021/jf8010855

97. Sharma, A., Garg, S., Sheikh, I., Vyas, P., & Dhaliwal, H. S. (2020). Effect of wheat grain protein composition on end-use quality. *Journal of food science and technology*, *57*, 2771-2785.

98. Shewry, P. (2019). What is gluten—why is it special?. *Frontiers in nutrition*, *6*, 101. https://doi.org/10.3389/fnut.2019.00101

99. Shewry, P. R. (2009). Wheat. *Journal of experimental botany*, *60*(6), 1537-1553. https://doi.org/10.1093/jxb/erp058

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

101. Stahl, W., & Sies, H. (2003). Antioxidant activity of carotenoids. molecular aspects of medicine, 24: 345-351. *DOI*, *10*, S0098-2997.

102. Taketa, S., Yuo, T., Tonooka, T., Tsumuraya, Y., Inagaki, Y., Haruyama, N., Larroque, O., & Jobling, S. A. (2012). 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, *63*(1), 381-392. https://doi.org/10.1093/jxb/err285

103. Tanksley, S. D. (1993). Mapping polygenes. *Annual review of genetics*, *27*(1), 205-233.

104. Tanksley, S. D., & McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *science*, *277*(5329), 1063-1066. https://doi.org/10.1126/science.277.5329.1063

105. Tonooka, T., Aoki, E., Yoshioka, T., & Taketa, S. (2009). A novel mutant gene for (1-3, 1-4)-β-D-glucanless grain on barley (Hordeum vulgare L.) chromosome 7H. *Breeding Science*, *59*(1), 47-54. https://doi.org/10.1270/jsbbs.59.47

106. Topping, D. (2007). Cereal complex carbohydrates and their contribution to human health. *Journal of Cereal Science*, *46*(3), 220-229. https://doi.org/10.1016/j.jcs.2007.06.004

107. Trethewey, J. A., Campbell, L. M., & Harris, P. J. (2005). (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*, *92*(10), 1660-1674. https://doi.org/10.3732/ajb.92.10.1660

108. Türkösi, E., Darko, E., Rakszegi, M., Molnár, I., Molnár-Láng, M., & Cseh, A. (2018). Development of a new 7BS. 7HL winter wheat-winter barley Robertsonian translocation line conferring increased salt tolerance and (1, 3; 1, 4)-β-D-glucan content. *PLoS One*, *13*(11), e0206248. https://doi:10.1371/JOURNAL.PONE.0206248

109. U.S. Wheat Associates. Available online: https://www.uswheat.org/wheatletter/shifted-trade-flows-black-sea-durum-enters-the-global-market/ (accessed on 20 January 2025).

110. Van den Berg, H., Faulks, R., Granado, H. F., Hirschberg, J., Olmedilla, B., Sandmann, G., Southon, S., & Stahl, W. (2000). The potential for the improvement of carotenoid levels in foods and the likely systemic effects. *Journal of the Science of Food and Agriculture*, *80*(7), 880-912. https://doi.org/10.1002/(SICI)1097-0010(20000515)80:7<880::AID-JSFA646>3.0.CO;2-1

111. Vega-Sánchez, M. E., Verhertbruggen, Y., Christensen, U., Chen, X., Sharma, V., Varanasi, P., Jobling, S. A., Talbot, M., White, R. G., Joo, M., Singh, S., Auer, M., Scheller, H. V., & Ronald, P. C. (2012). 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*, *159*(1), 56-69. https://doi.org/10.1104/pp.112.195495

112. Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 years of GWAS discovery: biology, function, and translation. *The American Journal of Human Genetics*, *101*(1), 5-22.

113. Waddington, S. R., Osmanzai, M., Yoshida, M., & Ransom, J. K. (1987). The yield of durum wheats released in Mexico between 1960 and 1984. *The Journal of Agricultural Science*, *108*(2), 469-477.

114. Wieser, H., Koehler, P., & Scherf, K. A. (2020). *Wheat-An exceptional crop: Botanical features, chemistry, utilization, nutritional and health aspects*. Elsevier.

115. Wieser, H. (2007). Chemistry of gluten proteins. *Food microbiology*, *24*(2), 115-119. https://doi.org/10.1016/j.fm.2006.07.004

116. Wong, S. C., Shirley, N. J., Little, A., Khoo, K. H., Schwerdt, J., Fincher, G. B., Burton, R. A., & Mather, D. E. (2015). Differential

expression of the HvCslF6 gene late in grain development may explain quantitative differences in (1, 3; 1, 4)-β-glucan concentration in barley. *Molecular Breeding*, *35*, 1-12.

117. Woodward, J. R., Phillips, D. R., & Fincher, G. B. (1983). Water-soluble (1→3),(1→4)-β-D-glucans from barley (Hordeum vulgare) endosperm. I. Physicochemical properties. *Carbohydrate Polymers*, *3*(2), 143-156. https://doi.org/10.1016/0144-8617(83)90004-8

118. Zhang, D., Williams, B. A., Mikkelsen, D., Li, X., Keates, H. L., Lisle, A. T., Collins, H. M., Fincher, G. B., Bird, A. R., Topping, D. L., Gidley, M. J., & Bryden, W. L. (2015). Soluble arabinoxylan alters digesta flow and protein digestion of red meat-containing diets in pigs. *Nutrition*, *31*(9), 1141-1147. https://doi.org/10.1016/j.nut.2015.03.006

119. Zhang, N., Gibon, Y., Wallace, J. G., Lepak, N., Li, P., Dedow, L., Chen, C., So, Y., Kremling, K., Bradbury, P. J., Brutnell, T., Stitt, M., & Buckler, E. S. (2015). Genome-wide association of carbon and nitrogen metabolism in the maize nested association mapping population. *Plant Physiology*, *168*(2), 575-583. https://doi.org/10.1104/pp.15.00025

120. Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., Huang, M., Yao, Y., Bassu, S., Ciais, P., Durand, J., Elliott, J., Ewert, F., Janssens, I. A., Li, T., Lin, E., Liu, Q., Martre, P., Muller, C., Peng, S., Penuelas, J., Ruane, A. C., Wallach, D., Wang, T., Wu, D., Liu, Z., Zhu, Y., Zhu, Z., & Asseng, S. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of sciences*, *114*(35), 9326-9331. https://doi.org/10.1073/pnas.1701762114

121. Žilić, S., Barać, M., Pešić, M., Hadži-Tašković-Šukalović, V., Dodig, D., Mladenović-Drinić, S., & Janković, M. (2011). Genetic variability of albumin-globulin content, and lipoxygenase, peroxidase activities among bread and durum wheat genotypes. *Genetika*, *43*(3), 503-516. **10.2298/GENSR1103503Z**

122. Zohary, D., Hopf, M., & Weiss, E. (2012). *Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199549061.001.0001

CHAPTER I

Exploring *Aegilops caudata*: A Comprehensive Study of the *CslF6* Gene and β-Glucan

BACKGROUND: The search for sustainable, nutrient food has led to increased interest in ancient grains and wild relatives of cultivated cereals. *Aegilops caudata*, a wild wheatgrass species, holds significant genetic potential for crop improvement and offers promising nutritional benefits.

AIM: This study focused on analyzing the *CslF6* gene sequence and its protein structure in *Ae. caudata* using comparative genomic methods with other grass species, to identify genetic variations that could influence β-glucan biosynthesis. The *CslF6* gene in *Ae. caudata* was the isolate and characterize, defining thegenomic sequence, predicting the protein structure, and carrying out the comparative genomics to identify evolutionary relationships and pinpointed amino acid variations in the *CslF6* enzyme.

RESULTS: Comparative genomic analysis indicates that the *CslF6* gene sequence of *Ae. caudata* shared high sequence identity with wheat genomes, confirming a close evolutionary relationship with other monocots. Key amino acid motifs within the *CslF6* enzyme, especially near catalytic regions, exhibited variability across monocots, which likely impacts β-glucan biosynthesis and the DP3:DP4 ratio. These findings suggest genetic divergence or unique functional adaptations in *Ae. caudata*. The study highlighted the importance of specific amino acid residues in modulating enzymatic activity and influencing polysaccharide structure, offering deeper insights into β-glucan biosynthesis.

genes

MDPI

Article

Exploring *Aegilops caudata*: A Comprehensive Study of the *CslF6* Gene and β-Glucan

Ilaria Marcotuli *, Davide Caranfa, Pasqualina Colasuonno, Stefania Lucia Giove and Agata Gadaleta

Department of Soil, Plant and Food Sciences, University of Bari Aldo Moro, Via G. Amendola 165/A, 70126 Bari, Italy; davide.caranfa@uniba.it (D.C.); pattybiotec@yahoo.it (P.C.); stefanialucia.giove@uniba.it (S.L.G.); agata.gadaleta@uniba.it (A.G.)

* Correspondence: ilaria.marcotuli@uniba.it; Tel.: +39-080-544-2996

Abstract: In the quest for sustainable and nutritious food sources, exploration of ancient grains and wild relatives of cultivated cereals has gained attention. *Aegilops caudata*, a wild wheatgrass species, stands out as a promising genetic resource due to its potential for crop enhancement and intriguing nutritional properties. This manuscript investigates the *CslF6* gene sequence and protein structure of *Aegilops caudata*, employing comparative analysis with other grass species to identify potential differences impacting β-glucan content. The study involves comprehensive isolation and characterization of the *CslF6* gene in *Ae. caudata*, utilizing genomic sequence analysis, protein structure prediction, and comparative genomics. Comparisons with sequences from diverse monocots reveal evolutionary relationships, highlighting high identities with wheat genomes. Specific amino acid motifs in the CslF6 enzyme sequence, particularly those proximal to key catalytic motifs, exhibit variations among monocot species. These differences likely contribute to alterations in β-glucan composition, notably impacting the DP3:DP4 ratio, which is crucial for understanding and modulating the final β-glucan content. The study positions *Ae. caudata* uniquely within the evolutionary landscape of *CslF6* among monocots, suggesting potential genetic divergence or unique functional adaptations within this species. Overall, this investigation enriches our understanding of β-glucan biosynthesis, shedding light on the role of specific amino acid residues in modulating enzymatic activity and polysaccharide composition.

Keywords: *Aegilops caudata*; *CslF6* gene; comparative genomics; β-glucan biosynthesis

Citation: Marcotuli, I.; Caranfa, D.; Colasuonno, P.; Giove, S.L.; Gadaleta, A. Exploring *Aegilops caudata*: A Comprehensive Study of the *CslF6* Gene and β-Glucan. *Genes* **2024**, *15*, 168. https://doi.org/10.3390/genes15020168

Academic Editors: Patrizia Galeffi and Giuseppina Del Signore

Received: 19 December 2023
Revised: 24 January 2024
Accepted: 25 January 2024
Published: 27 January 2024

1. Introduction

In the pursuit of sustainable and nutritious food sources, researchers have turned their attention to ancient grains and wild relatives of cultivated cereals [1]. Among these, *Ae. caudata*, a wild wheatgrass species, has gained prominence due to its potential as a valuable genetic resource for crop improvement and its intriguing nutritional properties. In particular, the seeds of *Ae. caudata* have been a subject of increasing interest for their unique nutrient content and their association with genes responsible for β-glucan biosynthesis [2].

Ae. caudata belongs to the genus *Aegilops* within the grass family *Poaceae*. It is native to regions of the Middle East and is closely related to cultivated wheat species (*Triticum* spp.). Unlike cultivated wheat, *Ae. caudata* has not undergone extensive domestication, making it a valuable genetic resource for breeding programs aimed at improving the resilience, disease resistance, and nutritional content of wheat and related cereals. Its genetic diversity and adaptation to various environmental conditions make it an intriguing subject for research in the context of sustainable agriculture [3,4].

Ae. caudata seeds have garnered attention for their unique nutrient composition. These seeds are a rich source of essential nutrients, including proteins, vitamins, minerals, and dietary fiber. The protein content of *Ae. caudata* seeds is of particular interest, as it contains a balanced amino acid profile, including essential amino acids. Furthermore, these seeds

are known to be rich in micronutrients such as iron, zinc, and folate, which are critical for human health. The dietary fiber content, including β-glucans, in *Ae. caudata* seeds further adds to their nutritional appeal [5,6].

B-glucans are a group of polysaccharides found in the cell walls of various cereal grains, including wheat, oats, and barley [7–9]. They are known for their potential health benefits, particularly lowering cholesterol levels and supporting digestive health [10–12]. The molecular structure of β-glucan dictates its function. According to current understanding, β-1,3/1,4 glycosidic bonds link cereal β-glucans. In the structure of cereal β-glucan, the β-1,4-linked glucose chain is interspersed with β-1,3 linkages. While trimers and tetramers are the most common consecutive β-1,4 segments, longer cellulose-like segments are also present in β-glucan molecules [13]. Significantly, two primary oligomer units, DP3 and DP4, account for over 90% of cereal β-glucan structures, varying among species and between genotypes [14].

Understanding the genetic basis of β-glucan biosynthesis is crucial for crop improvement and the development of cereal varieties with enhanced β-glucan content [7].

In recent years, significant progress has been made in identifying and characterizing the genes involved in β-glucan biosynthesis in various cereal species [7,10,15,16]. In the case of *Ae. caudata*, research efforts have focused on elucidating the genetic mechanisms responsible for the synthesis of β-glucans in its seeds. These insights can potentially lead to the development of wheat varieties with improved β-glucan content through breeding or genetic modification techniques [6].

B-glucans exhibit diverse physiological and nutritional functions in plants. They are vital components of cell walls, where they contribute to structural integrity and resistance against environmental stresses [17,18]. Moreover, β-glucans function as storage polysaccharides in various plant organs, such as seeds, grains, and tubers, serving as an energy reserve during germination and growth [19,20].

Our knowledge regarding how and where β-glucans are produced remains largely incomplete, as well as which genes are involved, their functions, and interactions, and the specific activities of the enzymes. A group of genes, forming a superfamily, plays a significant role in the synthesis of these polysaccharides. This superfamily comprises the *cellulose synthase* (*Ces*) [21–23] and *cellulose-synthase-like* (*Csl*) [24] families.

The *Csl* superfamily is responsible for synthesizing several plant cell wall polysaccharides, organized into subfamilies labeled A to H, each of which consists of multiple genes [25]. For instance, in rice (*Oryza sativa* L.), there are a total of 37 *Csl* genes [26], whereas *Arabidopsis* has 30 [27]. Notably, not all *Csl* subfamilies are represented in all higher plant groups. The *CslB* and *CslG* subfamilies are exclusive to dicotyledons and gymnosperms, while the *CslF* and *CslH* groups are only found in monocotyledons [28]. These subfamilies directly or indirectly regulate the abundance and fine structure of β-glucans in both grain and other parts of the plant [29–31].

Research conducted by Burton et al. [32] revealed that over-expressing a *CslF* gene, under the control of an endosperm-specific promoter, led to an increase in β-glucan content and a significant reduction in starch in transgenic grains. Given that the β-glucan and starch pathways compete for the initial substrate—glucose—used in their synthesis, *Brachypodium distachyon*, with over 40% of its grain weight as β-glucan and only about 6% starch, further supports a regulatory connection between starch and β-glucan synthesis [29]. Additionally, when *Arabidopsis* was transformed with the *OsCslF6* gene, it produced mixed-linkage glucan in the cell wall, indicating the capability of *CslF6* to synthesize β-glucan [33]. In barley, four corresponding *CslF* genes were mapped to chromosome 2H (*HvCslF3*, *HvCslF4*, *HvCslF8*, *HvCslF10*), with two other genes on chromosomes 1H (*HvCslF9*) and 7H (*HvCslF6*), corresponding to quantitative trait loci (QTL) for grain β-glucan content [34,35].

In rice, knockout mutants of *OsCslF6* synthesize minimal β-glucan content. Nemeth et al. [36] identified the *CslF6* gene in wheat and demonstrated that transgenic manipulation through iRNA could modify the amounts and properties of β-glucan in wheat. Other studies showed that the addition of barley chromosome 7H (where *HvCslF6* is located) to

the wheat genome increases β-glucan production [37]. The wheat–barley addition lines, obtained through hybridization, contain a genetic background of common wheat, allowing the genetic analysis of a single barley chromosome affecting the final phenotype. They have been used for several trait studies as the production of bioactive compounds [38–40].

In this scenario, *Ae. caudata* stands as a promising genetic resource with unique nutrient content and the potential to contribute to the enhancement of cereal crops. The investigation of genes involved in β-glucan biosynthesis in this species not only offers insights into its nutritional properties but also flags the way for improving the health-promoting aspects of cultivated wheat. This paper investigates the cellulose synthase F6 gene sequence and protein structure of *Ae. caudata* and carries out a comparative analysis with other grasses to identify possible differences among species and correlate them with the final β-glucan content.

2. Materials and Methods

2.1. Plant Material and DNA Extraction

The *Aegilops caudata* genotype from the GenBank of the Department of Soil, Plant and Food Sciences (University of Bari, Italy) was used to characterize the *CslF6* gene. The genotype was grown in Valenzano (Bari, Department of Soil, Plant and Food Science, University of Bari Aldo Moro), and leaves were harvested at tillering time. Genomic DNA was extracted from fresh leaves using the technique outlined in Sharp et al. [41] and subsequently underwent purification through phenol-chloroform extraction. The quality and concentration of the DNA were assessed via spectrophotometric analysis using the NanoDrop2000 (Thermo Scientific™, Thermo Fisher Scientific, Waltham, MA, USA 02451) at 260 and 280 nm, with an A260/A280 ratio falling within the range of 1.6 to 1.8, and checked by agarose gel-electrophoresis.

For fragment sequencing, DNA amplifications were conducted in 25 mL reaction mixtures, with each mixture containing 25 ng of template DNA, 2 mM of each primer, 200 mM of each dNTP, 2.5 mM of MgCl2, 1X PCR buffer (10 mM TRIS-HCl, pH 8.3, 10 mM KCl), and 0.5 units of Taq DNA polymerase. The PCR protocol used in a Perkin Elmer DNA Thermal Cycler (Norwalk, CT, USA) was as follows: initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 95 °C for 1 min, annealing at 55 °C/65 °C for 2 min, and extension at 72 °C for 1 min, with a final extension step at 72 °C for 15 min. The PCR products were visualized through 1.5% agarose gel electrophoresis.

2.2. Cellulose Synthase Gene (CslF6) Isolation and Characterization

To isolate the complete sequences of the *Ae. caudata CslF6* gene, we used the sequences of durum wheat, previously isolated by our lab [42] as an initial query probe to blast the Persephone® multi-genome browser (https://web.persephonesoft.com/?data=genomes/TA1851, accessed on 8 November 2023), which contain the *Ae. umbellulata* genome assemblies. From the Persephone® browser the sequence with 99% of similarity located on chromosome 7U (*chr7U_TA1851*) was used as a probe for the primer design.

In order to obtain the entire gene sequences, a complete set of genome specific primer pairs were designed by using 'Primer3 Input' (version 0.4.0)' software. Single PCR fragments were directly purified with an EuroGold Cycle Pure Kit and sequenced in both frame direction (5′->3′ and 3′->5′) using BigDye chemistry (Applied Biosystems) in a 96 capillary automatic sequencer ABI PRISM 3500. Gaps and uncertain sequence were resolved by primer walking. Regions of less coverage or ambiguous reads were rechecked with additional primers. Sequence assembly was obtained with 'Codone Code Aligner' (version 11.0.2) and 'Geneious' (version 2023.2.1) assembly programs.

Gene prediction was conducted with the FGENESH program (http://linux1.softberry.com/berry.phtml?topic=fgeneshandgroup=programsandsubgroup=gfind, accessed on 27 September 2023). Consensus exon/introns boundaries were confirmed using grass expressed sequence tag sequences aligned to the genomic sequence.

2.3. Aegilops caudata CslF6 Protein Sequence and Structure

To predict the protein sequence and structure based on the genomic sequence, Geneious software (version 2023.2.1) was employed for the sequence translation and prediction of transmembrane regions, coiled coil regions, conserved regions between wheat and barley proteins. Additionally, the *Ae. caudata* CslF6 protein newly obtained was aligned with the corresponding sequences from the alignment of the CslF6 from *Oryza sativa*, *Setaria italica*, *Sorghum bicolor*, *Zea mays*, *Brachypodium*, *Avena sativa*, *Hordeum vulgare*, wheat (A, B and D genomes), *Aegilops strangulata*, *Trititucm dicoccoides* and *Trititcum urartu*. Protein sequences were additionally employed in the phylogenetic examination using the neighbor joining method (NJ), and its topology was evaluated using 1000 bootstrap replicates implemented in 'Geneious'. Differences in the residues of the principal motif of the proteins were investigated to elucidate the potential impact of particular amino acid variations and their spatial arrangements near the active site on the intricate structure of the (1,3;1,4)-β-glucan synthesized. Homology models with cellulose synthases of bacteria (BCSA) were also implemented using the sequences from rice, maize, sorgo and *Setaria*, *Triticum* subspecies, *Brachypodium*, oat, and barley.

2.4. Promoter Cis-Acting Element Distribution Analysis

The 2000 -bp sequences upstream of the start codon of *Ae. caudata CslF6* gene was extracted as the promoter region and submitted to the PlantCARE database (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on 11 January 2024) for prediction of the cis-acting elements.

3. Results

3.1. Isolation and Characterization of Cellulose Synthase-like F6 Gene (CslF6) in Ae. caudata

The sequence corresponding to the *CslF6* gene from the durum wheat genomic sequences, previously isolated by our group [42], was used as a query blast to the *Ae. umbellulata* assembly deposit in the Persephone® multi-genome browser to design specific primer pairs for the isolation of the *CslF6* full gene sequence in *Ae. caudata*. All the amplification analyses were carried out on DNA extracted from leaves of the *Ae. caudata* genotype grown in Valenzano (Bari, Italy). Through the genotype used for the analysis, the C genome sequence of *CslF6* was isolated with the corresponding cDNA. The genomic sequence was 5357 bp, including an mRNA of 2829 bp and a protein of 942 aa (Figure 1). Fgenesh++ (version 2.1) software was used for gene prediction to define the intron/exon structure, predicting a gene structure composed of three exons and two introns (Figure 1). Using Phytozome (version 13) software, a comparison between wheat and barley sequences was carried out through blast analysis. Considering the *Ae. caudata CslF6* with the wheat genome sequences, the identities detected were 96.8% with the A genome, 97.1% with the B genome, and 98.9% with the D genome, while considering the cDNA, the similarities were 97.4% with the A genome, 97.6% with the B genome, and 99.2% with the D genome.

Using the newly obtained sequence, we were able to localize the *CslF6* gene on the *Ae. umbellulata* genome through the Persephone® multi-genome browser page (https://web.persephonesoft.com/?data=genomes/TA1851, *accessed on 08/11/2023*). The gene was localized on chromosome chr7U_TA1851 at the physical position from 204.549.972 to 204.555.410 bp. Comparing the position of the gene with other species, the CslF6 was located on: chromosome 7H for barley (*HORVU7Hr1G070010*); chromosome group 7 for wheat (*TraesCS7A02G298600*, *TraesCS7B02G188400*, and *TraesCS7D02G294300* for bread wheat; *TRITD7Av1G149750* and *TRITD7Bv1G108090* for durum wheat) and *T. dicoccoides* (*TRIDC7AG041550* and *TRIDC7BG030910*); chromosome 7 for *Sorghum* (*SORBI_3007G050600*); chromosome group 7 for oat (*AVESA.00001b.r3.7Ag0002427*, *AVESA.00001b.r3.7Cg0002511*, *AVESA.00001b.r3.7Dg0001419*); chromosome 7 for *urartu* (*LOC125522276*); chromosome 7D for *Ae. strangulata* (*LOC109773098*); chromosome 6 for *S. italica* (*XM_004972717*); chromosome 10 for maize (*GRMZM2G110145*); chromosome Bd3 for *Brachypodium* (*BRADI_3g16307v3*); chromosome 8 for rice (*LOC_Os08g06380*) (Table 1).

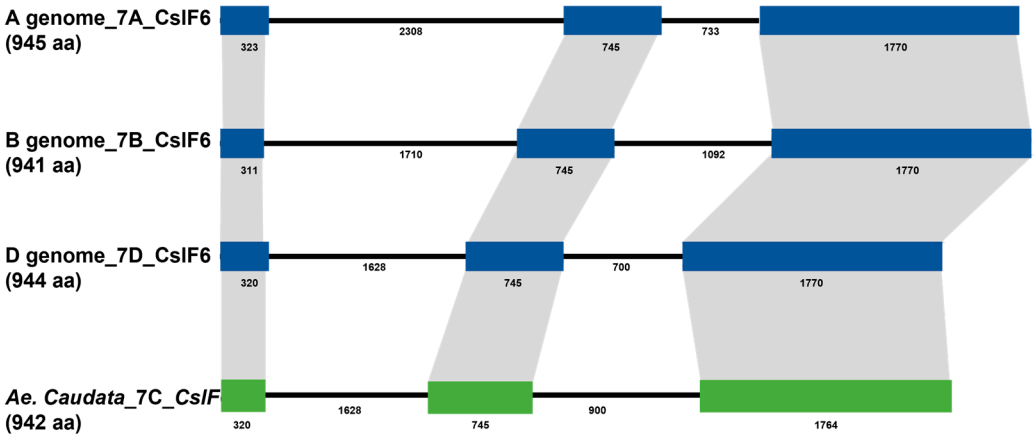


Figure 1. Comparison of gene structures between *Ae. caudata* and *Triticum* genome sequences (A, B and D) based on colored boxes, highlighting conserved exons. Intron and exon sizes are shown, as well as the whole gene (in brackets). The *CslF6* genes, in the genomes reported, are composed of three exons of conserved sizes and two introns.

Table 1. Chromosome location of CslF6 gene in a set of monocotyledons used for the analysis. For each species, the accession number relative to the sequence was reported.

Species	Chromosome Location	Accession Number
H. vulgare	7H	HORVU7Hr1G070010
T. turgidum ssp. *Durum*	7A	TRITD7Av1G149750
	7B	TRITD7Bv1G108090
T. aestivum	7A	TraesCS7A02G298600
	7B	TraesCS7B02G188400
	7D	TraesCS7D02G294300
T. dicoccoides	7A	TRIDC7AG041550
	7B	TRIDC7BG030910
S. bicolor	7	SORBI_3007G050600
A. sativa	7A	AVESA.00001b.r3.7Ag0002427
	7C	AVESA.00001b.r3.7Cg0002511
	7D	AVESA.00001b.r3.7Dg0001419
T. urartu	7A	LOC125522276
A. strangulata	7D	LOC109773098
S. italica	6	XM_004972717
Z. mays	10	GRMZM2G110145
Brachypodium	Bd3	BRADI_3g16307v3
O. sativa	8	LOC_Os08g06380

The newly obtained sequence was used for the determination of the amino acid sequence through Geneious (version 2023.2.1) software (Figure 2).

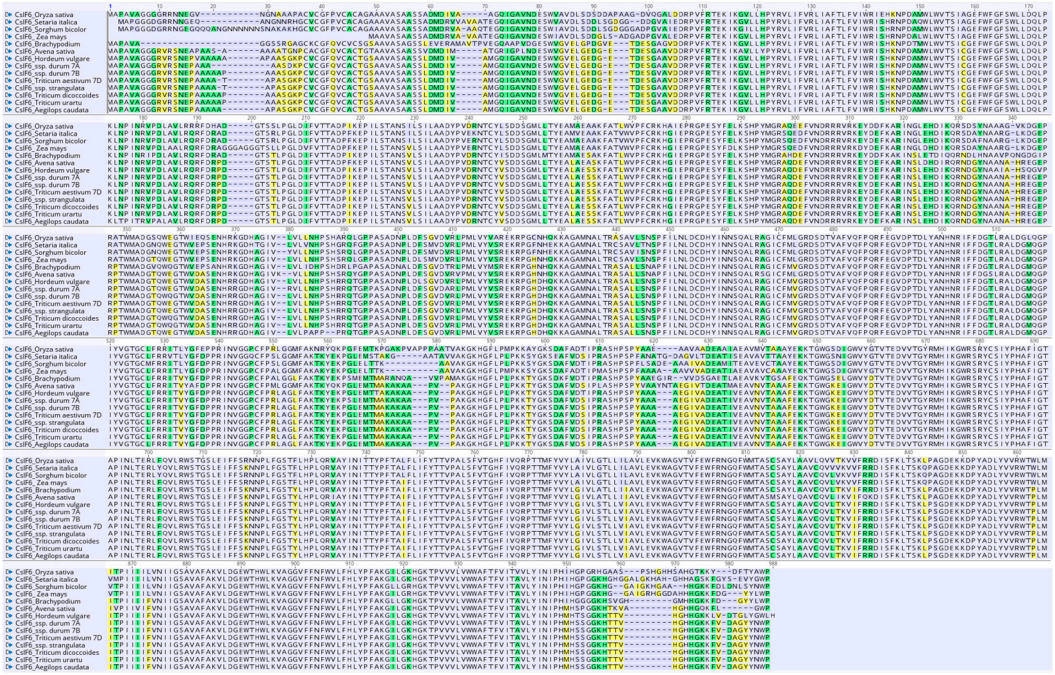


Figure 2. Plant CslF6 amino acid sequences for the mature protein are aligned with Clustal V: *Ae. caudata* (present report), wheat (A, B and D genomes), *T. urartu* (*XP_048543309.1*), *T. dicoccoides* (*XP_037457806.1*), *Brachypodium* (*XP_003573454.1*), rice (*AKJ66179.1*), maize (*AKJ66177.1*), *Setaria* (*XP_004972774.1*), *Sorghum* (*XP_002445102.1*), *Avena* (*AKJ66176.1*), barley (*ABZ01578.1*), *Ae. Stranguata* (*XP_020187382.1*).

3.2. Comparison of Amino Acid Sequences from Other Species

Once the *Ae. caudata CslF6* gene sequence was isolated, a protein structure prediction was performed to define the differences among a selection of monocot sequences available in public databases and relate them to the final β-glucan content and DP3/DP4 ratio in the different species. The DP3 and DP4 represent two major oligomer units obtained from the enzymatic digestion of the β-glucan, which explain more than 90% of cereal β-glucan structures and are strongly correlated to the degree of solubility of the polysaccharide [43].

The closest match was found with wheat genome D and *Aegilops taushii* ssp. *strangulata* (98%), followed by the other *Triticum* subspecies (~98%) and barley (97%). The more divergent sequences were the CslF6 from *Sorghum* and *Setaria*, at 81.4% and 81.6%, respectively. The amino acid length for each species differed, with some of them being of identical length. In order to investigate the evolutionary distances, the UPGMA tree was implemented in Geneious (version 2023.2.1) software. The tree showed two main clusters, one including the CslF6 protein sequences from rice, maize, sorghum, and *Setaria*, and a second one with all the other monocot species considered. Among the second group, the Triticum subspecies clustered all together, while *Brachypodium*, oat, and barley grouped independently. In the evolutionary tree, the *Ae. caudata* gene represents the outline of the *Triticum* cluster (Figure 3).

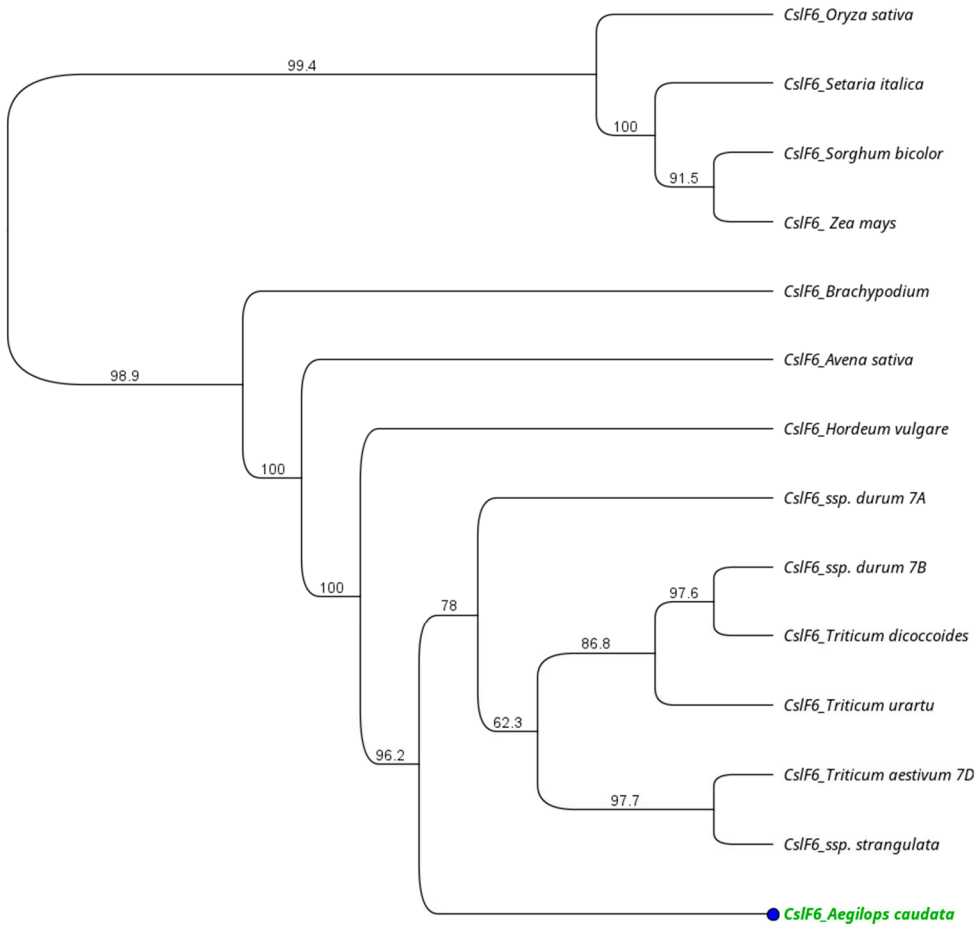


Figure 3. Phylogenetic relationships among the CslF6 polypeptide from *Ae. caudata* (green text), wheat (A, B and D genomes), *T. urartu*, *T. dicoccoides*, *Brachypodium*, rice, maize, *Setaria*, *Sorghum*, avena, barley, *Ae. Stranguata*.

3.3. Amino Acid Sequence, (1,3;1,4)-β-Glucan Amounts and DP3:DP4 Ratios

The amino acid sequences of the monocot sequences considered are highly conserved, and their 3D structures are likely to exhibit similarities (Figure 4). Considering that the amino acid sequence is of fundamental importance for the protein structure and β-glucan biosynthesis, a comparative analysis was carried out on the principal motifs using the homology model based on the *Rhodobacter sphaeroides* BCSA cellulose synthase subunit crystal structure [44].

As shown in Figure 4, the *Ae. caudata* protein sequence, the similarity in amino acid sequences between CslF6 enzymes from *Brachypodium*, barley, wheat, *T. dicoccoides*, *T. urartu* and *Ae. strangulata* suggests resemblances in their 3D structures.

Figure 4 showed that the G/D residue sits just before the anticipated "finger helix," adjacent to the TED motif believed to interact with the nascent polysaccharide's acceptor end.

Additionally, two other residues are reported on Figure 4, the W residue in the conserved QxxRW motif highlighting the catalytic pocket at the membrane distal side, which coordinates the translocation and elongation of the glucosidic units in the β-glucan biosynthesis, and the Y/F residue near the core QxxRW catalytic motif that interacts with CslF6's putative gating loop, housing an FxLTxK motif (Figure 4).

Both the differences highlighted in the residues close to the TED and QxxRW motif have an effect on contributing to the difference in the DP3:DP4 ratio [31].

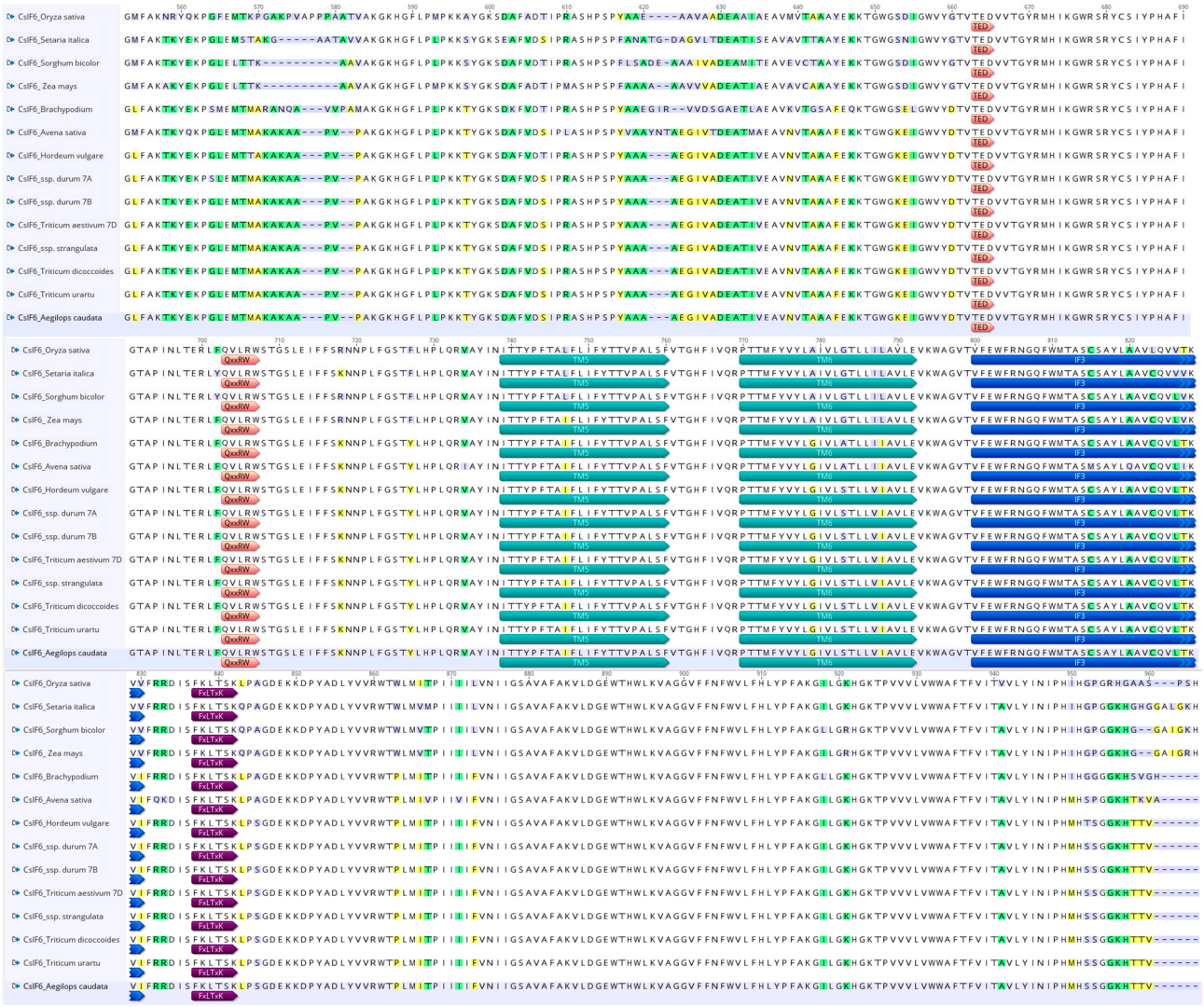


Figure 4. Comparison of CslF6 protein sequences from a selection of plants with colored residues highlighting sequence discrepancies. White boxes indicate a similarity of 100%, green boxes a similarity from 80% to 100%, yellow boxes a similarity from 60 to 80%, and purple boxes a similarity less than 60%. Shown are important predicted trans-membrane helices (TM5 and TM6), the FxLTxK motif, and the TED and QxxRW motifs. The cytoplasmic amphipathic helix (IF3) is also shown.

3.4. Analysis of Cis-Acting Elements in CslF6 Promoter Region of Ae. caudata Gene

The gene expression pattern primarily relies on cis-acting elements found within the regulatory regions of a promoter. Prediction of cis-acting elements of the *CslF6* gene promoter of *Ae. caudata* revealed a total of 18 variations (depicted in Figure 5). These elements encompassed a diverse range, including light-responsive (G-box, TCCC-motif and ATC-motif), defense and stress-related (ARE-element), phytohormone-responsive (ABRE), and growth and development-associated elements (CAT-box, GCN4-motif, MSA-like, O2-site and RY-element). Notably, G-box, ARE-element, TGACG-motif, and CGTCA-motif were prevalent, suggesting their pivotal role in stress resilience, growth, and development. Additionally, the gene appeared to be significantly regulated by cis-regulatory elements found to be involved in plant hormone responsiveness, including gibberellins, salicylic acid responsiveness and jasmonic acid.

Figure 5. Cis-acting regulatory element detected in the promoter sequence of the *CslF6* gene in *Ae. caudata*.

4. Discussion

The isolation and characterization of the *cellulose synthase-like F6* gene (*CslF6*) in *Ae. caudata* have unveiled significant insights into the molecular architecture and functional implications of this gene within the species. Our study employed a multifaceted approach,

integrating genomic sequence analysis, protein structure prediction, and comparative genomics to elucidate the distinct attributes of *CslF6* in *Ae. caudata*.

One key aspect, unveiled through this study, is the structural and functional conservation observed across related species. The high sequence identities between the CslF6 gene of *A. caudata* and wheat genomes (A, B, and D) underscore evolutionary links among these species. The conservation of intron/exon structures across monocots like durum wheat, maize, and oat, echoing that of *Aegilops caudata*, signifies an underlying genetic coherence in β-glucan biosynthesis among these cereals, useful for new breeding programs

We obtained the complete sequence of the *CslF6* gene in *Ae. caudata*, and we compared our data with *CslF6* sequences from other species. The gene had the same intron/exon structure as durum wheat, maize, and oats, including three exons and two introns [42,45,46].

Comparative analyses, particularly in the context of genomic and cDNA similarities with wheat genomes, underscored high identities between the *Ae. caudata CslF6* and the A, B, and D genomes of wheat, signifying evolutionary relationships and conservation across these species. Furthermore, protein structure predictions and comparative assessments with a spectrum of monocot sequences elucidated pivotal similarities and divergences. Notably, the closest matches were identified with the D wheat genome and *Ae. tauschii* ssp. *strangulata*, implying a close evolutionary association among these genomes. The results are strongly correlated to the genome's evolution; in fact, the ancestor of the bread wheat D genome is *Ae. tauschii* (Coss.) [47]. Additionally, we located the gene on chromosome 7C of the *Ae. caudata* genome, and we made a comparison with the location in other monocots. The synteny observed in the map locations suggested that these genes were highly conserved, as already reported by previous authors [48].

The investigation into amino acid sequences from various monocot species revealed significant conservation, suggesting analogous 3D structures among these CslF6 enzymes. Crucially, specific motifs within the amino acid sequences were pinpointed, notably the monocots used for the comparative analysis, which showed differences in the residue near the 'finger helix' (G/D) and the residue (Y/F) adjacent to the core QxxRW catalytic motif. No differences were underlined in the W residue within the conserved QxxRW motif. These residues are implicated in critical interactions in β-glucan biosynthesis, particularly influencing the DP3:DP4 ratio, thereby potentially affecting the final β-glucan content [31,44].

The observed variations in these residues, specifically those proximal to the TED and QxxRW motifs, are likely contributors to differences in the DP3:DP4 ratio, aligning with existing literature [31,44] and demonstrating their impact on β-glucan composition. This highlights the functional relevance of specific amino acid residues in shaping the enzymatic activity and structural features of CslF6, ultimately influencing the β-glucan synthesis pathway [33,44].

Moreover, the placement of *Ae. caudata* within an evolutionary tree, particularly as an outlier within the *Triticum* cluster, underscores its distinctiveness within the context of *CslF6* evolution among monocots, signifying potential unique functional adaptations or genetic divergence within this species as previously reported in literature [49].

The analysis of the cis-acting elements in the promotor region of the *CslF6* gene in the *Ae. caudata* genome highlighted the presence of many motifs associated with plant growth and development, stress response, and hormone regulation. These results suggested that the *CslF6* gene family may participate in plant growth and development as well as stress tolerance, and the family members are regulated by plant hormones as already reported in other species [50–52].

The potential applications of these findings extend to breeding strategies aimed at enhancing cereal varieties' nutritional content and resilience. Harnessing the genetic resources within *Ae. caudata* for targeted breeding or employing genetic modification techniques holds promise for developing cereals with improved β-glucan content, balanced amino acid profiles, and enhanced micronutrient richness. Moreover, this knowledge aids

in creating crops resilient to environmental stresses, thereby contributing to sustainable agricultural practices.

Furthermore, the detailed characterization of the *CslF6* gene and its protein structure in *Ae. caudata* provides a foundational framework for future studies. Exploring the regulatory mechanisms underlying β-glucan biosynthesis, investigating additional genetic factors influencing cereal nutritional content, and conducting functional validations of specific amino acid residues could offer deeper insights into refining cereal traits for human consumption.

5. Conclusions

The present work focuses on the cellulose synthase-like F6 gene (CslF6) in the *Ae. caudata* genome, providing valuable insights into the molecular processes of β-glucan biosynthesis and their potential impact on cereal nutrition. Utilizing genomic analysis, protein structure prediction, and comparative genomics, the research uncovers unique characteristics of CslF6 in *Ae. caudata*.

Comparative analysis of CslF6 amino acid sequences across monocot species reveals significant conservation, suggesting similar 3D structures. However, crucial motifs near the 'finger helix' and the QxxRW catalytic motif exhibit variations among species, influencing the DP3:DP4 ratio and, consequently, β-glucan content in cereals. These variations emphasize the role of specific residues in shaping CslF6 enzymatic activity and structural features.

The placement of *Ae. caudata* as an outlier in the evolutionary tree within the Triticum cluster signifies its distinctiveness in CslF6 evolution among monocots, hinting at potential unique functional adaptations or genetic divergence.

Exploring the promoter region of the CslF6 gene in the Ae. caudata genome reveals cis-acting elements associated with essential processes like plant growth, stress response, and hormone regulation. This implies multifaceted functions for the CslF6 gene family. The observed regulation by plant hormones aligns with known mechanisms in other species, highlighting evolutionary conservation.

In conclusion, this research deepens our understanding of β-glucan biosynthesis and suggests leveraging *Ae. caudata*'s genetic resources for improving cereal nutrition through breeding or genetic modification. The study's insights pave the way for further exploration and exploitation of these genetic resources in crop improvement programs focused on sustainable and nutritious food production.

Author Contributions: Conceptualization, I.M. and A.G.; Methodology, I.M. and A.G.; Software, I.M.; Validation, I.M.; Formal Analysis D.C., P.C. and S.L.G.; Data Curation, I.M.; Writing—Original Draft Preparation, I.M.; Writing—Review and Editing, I.M., P.C. and A.G.; Funding Acquisition, A.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Prima MED 2022-2025, section 2, AgroEcoMed.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

Ces, cellulose synthase; Csl, cellulose-synthase-like; DP3, 3-O-β-cellobiosyl-D-glucose; DP4, 3-O-β-cellotriosyl-D-glucose; BCSA, bacterial cellulose synthase A; TED motif, threonine-glutamic acid-aspartic acid motif; QxxRW motif, binding site for the terminal disaccharide of the glucan; TM5, trans-membrane helix 5; TM6, trans-membrane helix 6; IF3, cytoplasmic amphipathic helix; G-box, CACGTG box; TCCC-motif, element involved in defense and stress responsiveness; ATC-motif, conserved DNA module involved in light responsiveness; ARE-element, adenylate-uridylate-rich elements; ABRE, ABA response elements; CAT-box, element related to meristem

expression; GCN4-motif, 'leucine-zipper' recognition; MSA-like, mitosis-specific activator; O2-site, cis-acting regulatory element involved in zein metabolism regulation; RY-element, element involved in seed-specific regulation.

References

1. Dinelli, G.; Segura-Carretero, A.; Di Silvestro, R.; Marotti, I.; Arraez-Roman, D.; Benedettelli, S.; Ghiselli, L.; Fernadez-Gutierrez, A. Profiles of phenolic compounds in modern and old common wheat varieties determined by liquid chromatography coupled with time-of-flight mass spectrometry. *J. Chromatogr. A* **2011**, *1218*, 7670–7681. [CrossRef] [PubMed]
2. Marcotuli, I.; Colasuonno, P.; Cutillo, S.; Simeone, R.; Blanco, A.; Gadaleta, A. β-glucan content in a panel of Triticum and Aegilops genotypes. *Genet. Resour. Crop Evol.* **2019**, *66*, 897–907. [CrossRef]
3. King, J.; Grewal, S.; Yang, C.-y.; Hubbart Edwards, S.; Scholefield, D.; Ashling, S.; Harper, J.A.; Allen, A.M.; Edwards, K.J.; Burridge, A.J.; et al. Introgression of Aegilops speltoides segments in Triticum aestivum and the effect of the gametocidal genes. *Ann. Bot.* **2017**, *121*, 229–240. [CrossRef] [PubMed]
4. King, J.; Grewal, S.; Yang, C.-y.; Hubbart, S.; Scholefield, D.; Ashling, S.; Edwards, K.J.; Allen, A.M.; Burridge, A.; Bloor, C.; et al. A step change in the transfer of interspecific variation into wheat from Amblyopyrum muticum. *Plant Biotechnol. J.* **2017**, *15*, 217–226. [CrossRef]
5. Schneider, A.; Molnár, I.; Molnár-Láng, M. Utilisation of *Aegilops* (goatgrass) species to widen the genetic diversity of cultivated wheat. *Euphytica* **2008**, *163*, 1–19. [CrossRef]
6. Kishii, M. An Update of Recent Use of Aegilops Species in Wheat Breeding. *Front. Plant Sci.* **2019**, *10*, 585. [CrossRef]
7. Marcotuli, I.; Colasuonno, P.; Hsieh, Y.S.Y.; Fincher, G.B.; Gadaleta, A. Non-Starch Polysaccharides in Durum Wheat: A Review. *Int. J. Mol. Sci.* **2020**, *21*, 2933. [CrossRef] [PubMed]
8. Colasuonno, P.; Marcotuli, I.; Gadaleta, A.; Soriano, J.M. From Genetic Maps to QTL Cloning: An Overview for Durum Wheat. *Plants* **2021**, *10*, 315. [CrossRef]
9. Berski, W.; Krystyjan, M.; Buksa, K.; Zięć, G.; Gambuś, H. Chemical, physical and rheological properties of oat flour affected by the isolation of beta-glucan preparation. *J. Cereal Sci.* **2014**, *60*, 533–539. [CrossRef]
10. Dhingra, D.; Michael, M.; Rajput, H.; Patil, R.T. Dietary fibre in foods: A review. *J. Food Sci. Technol.* **2012**, *49*, 255–266. [CrossRef]
11. Hase, W.-J.L.K. Gut microbiota–generated metabolites in animal health and disease. *Nat. Chem. Biol.* **2014**, *10*, 416–424.
12. Lafiandra, D.; Riccardi, G.; Shewry, P.R. Improving cereal grain carbohydrates for diet and health. *J. Cereal Sci.* **2014**, *59*, 312–326. [CrossRef] [PubMed]
13. Hu, X.; Zhao, J.; Zhao, Q.; Zheng, J. Structure and Characteristic of β-Glucan in Cereal: A Review. *J. Food Process. Preserv.* **2014**. [CrossRef]
14. De Arcangelis, E.; Djurle, S.; Andersson, A.A.; Marconi, E.; Messia, M.C.; Andersson, R. Structure analysis of β-glucan in barley and effects of wheat β-glucanase. *J. Cereal Sci.* **2019**, *85*, 175–181. [CrossRef]
15. Taketa, S.; Yuo, T.; Tonooka, T.; Tsumuraya, Y.; Inagaki, Y.; Haruyama, N.; Larroque, O.; Jobling, S.A. Functional characterization of barley betaglucanless mutants demonstrates a unique role for CslF6 in (1,3;1,4)-β-D-glucan biosynthesis. *J. Exp. Bot.* **2011**, *63*, 381–392. [CrossRef]
16. Hu, G.; Burton, C.; Hong, Z.; Jackson, E. A mutation of the cellulose-synthase-like (CslF6) gene in barley (*Hordeum vulgare* L.) partially affects the β-glucan content in grains. *J. Cereal Sci.* **2014**, *59*, 189–195. [CrossRef]
17. Rayapuram, C.; Jensen, M.K.; Maiser, F.; Shanir, J.V.; Hornshoj, H.; Rung, J.H.; Gregersen, P.L.; Schweizer, P.; Collinge, D.B.; Lyngkjaer, M.F. Regulation of basal resistance by a powdery mildew-induced cysteine-rich receptor-like protein kinase in barley. *Mol. Plant Pathol.* **2012**, *13*, 135–147. [CrossRef]
18. Ha, Y.-s.; Covert, S.F.; Momany, M. FsFKS1, the 1,3-b-glucan synthase from the caspofungin-resistant fungus *Fusarium solani*. *Eukaryot. Cell* **2006**, *5*, 1036–1042. [CrossRef]
19. Fincher, G.B.; Burton, R.A. Evolution and Development of Cell Walls in Cereal Grains. *Front. Plant Sci.* **2014**, *5*, 456. [CrossRef]
20. Bernal, A.J.; Yoo, C.M.; Mutwil, M.; Jensen, J.K.; Hou, G.; Blaukopf, C.; Sorensen, I.; Blancaflor, E.B.; Scheller, H.V.; Willats, W.G. Functional analysis of the cellulose synthase-like genes CSLD1, CSLD2, and CSLD4 in tip-growing *Arabidopsis* cells. *Plant Physiol.* **2008**, *148*, 1238–1253. [CrossRef]
21. Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; Hofte, H.; Plazinski, J.; Birch, R.; et al. Molecular analysis of cellulose biosynthesis in *Arabidopsis*. *Science* **1998**, *279*, 717–720. [CrossRef]
22. Burton, R.; Gidley, M.; Fincher, G. Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat. Chem. Biol.* **2010**, *6*, 724–732. [CrossRef]
23. Scheible, W.R.; Pauly, M. Glycosyltransferases and cell wall biosynthesis: Novel players and insights. *Curr. Opin. Plant Biol.* **2004**, *7*, 285–295. [CrossRef] [PubMed]
24. Burton, R.; Wilson, S.; Hrmova, M.; Harvey, A.; Shirley, N.; Medhurst, A.; Stone, B.; Newbigin, E.; Bacic, A.; Fincher, G. Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)-D-glucans. *Am. Assoc. Adv. Sci.* **2006**, *311*, 1940–1942. [CrossRef] [PubMed]

25. Little, A.; Schwerdt, J.G.; Shirley, N.J.; Khor, S.F.; Neumann, K.; O'donovan, L.A.; Lahnstein, J.; Collins, H.M.; Henderson, M.; Fincher, G.B.; et al. Revised phylogeny of the cellulose synthase gene superfamily: Insights into cell wall evolution. *Plant Physiol.* **2018**, *177*, 1124–1141. [CrossRef] [PubMed]
26. Hazen, S.P.; Scott-Craig, J.S.; Walton, J.D. Cellulose synthase-like genes of rice. *Plant Physiol.* **2002**, *128*, 336–340. [CrossRef]
27. Somerville, C.; Bauer, S.; Brininstool, G.; Facette, M.; Hamann, T.; Milne, J.; Osborne, E.; Paredez, A.; Persson, S.; Raab, T.; et al. Toward a systems approach to understanding plant cell walls. *Science* **2004**, *306*, 2206–2211. [CrossRef] [PubMed]
28. Farrokhi, N.; Burton, R.; Brownfield, L.; Hrmova, M.; Wilson, S.; Bacic, A.; Fincher, G. Plant cell wall biosynthesis: Genetic, biochemical and functional genomics approaches to the identification of key genes. *Plant Biotechnol. J.* **2006**, *4*, 145–167. [CrossRef]
29. Bain, M.; van de Meene, A.; Costa, R.; Doblin, M.S. Characterisation of Cellulose Synthase Like F6 (CslF6) Mutants Shows Altered Carbon Metabolism in beta-D-(1,3;1,4)-Glucan Deficient Grain in Brachypodium distachyon. *Front. Plant Sci.* **2020**, *11*, 602850. [CrossRef]
30. Burton, R.; Jobling, S.; Harvey, A.; Shirley, N.; Mather, D.; Bacic, A.; Fincher, G. The Genetics and Transcriptional Profiles of the Cellulose Synthase-Like HvCslF Gene Family in Barley. *Plant Physiol.* **2008**, *146*, 1821–1833. [CrossRef]
31. Dimitroff, G.; Little, A.; Lahnstein, J.; Schwerdt, J.G.; Srivastava, V.; Bulone, V.; Burton, R.A.; Fincher, G.B. (1,3;1,4)-beta-Glucan Biosynthesis by the CSLF6 Enzyme: Position and Flexibility of Catalytic Residues Influence Product Fine Structure. *Biochemistry* **2016**, *55*, 2054–2061. [CrossRef] [PubMed]
32. Burton, R.; Collins, H.; Kibble, N.; Smith, J.; Shirley, N.; Jobling, S.; Henderson, M.; Singh, R.; Pettolino, F.; Wilson, S.; et al. Over-expression of specific HvCslF cellulose synthase-like genes in transgenic barley increases the levels of cell wall (13;14)-beta-D-glucans and alters their fine structure. *Plant Biotechnol. J.* **2011**, *9*, 117–135. [CrossRef] [PubMed]
33. Doblin, M.S.; Pettolino, F.A.; Wilson, S.M.; Campbell, R.; Burton, R.A.; Fincher, G.B.; Newbigin, E.; Bacic, A. A barley *cellulose synthase-like CSLH* gene mediates (1,3;1,4)-beta-D-glucan synthesis in transgenic *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5996–6001. [CrossRef] [PubMed]
34. Molina-Cano, J.L.; Moralejo, M.; Elia, M.; Munoz, P.; Russell, J.R.; Perez-Vendrell, A.M.; Ciudad, F.; Swanston, J.S. QTL analysis of a cross between European and North American malting barleys reveals a putative candidate gene for beta-glucan content on chromosome 1H. *Mol. Breed.* **2007**, *19*, 275–284. [CrossRef]
35. Tonooka, T.; Aoki, E.; Yoshioka, T.; Taketa, S. A novel mutant gene for (1-3,1-4)-β-D-glucanless grain on barley (*Hordeum vulgare* L.) chromosome 7H. *Breed. Sci.* **2009**, *59*, 47–54. [CrossRef]
36. Nemeth, C.; Freeman, J.; Jones, H.D.; Sparks, C.; Pellny, T.K.; Wilkinson, M.D.; Dunwell, J.; Andersson, A.A.; Aman, P.; Guillon, F.; et al. Down-regulation of the CSLF6 gene results in decreased (1,3;1,4)-beta-D-glucan in endosperm of wheat. *Plant Physiol.* **2010**, *152*, 1209–1218. [CrossRef] [PubMed]
37. Colasuonno, P.; Marcotuli, I.; Cutillo, S.; Simeone, R.; Blanco, A.; Gadaleta, A. Effect of barley chromosomes on the β-glucan content of wheat. *Genet. Resour. Crop Evol.* **2020**, *67*, 561–567. [CrossRef]
38. Islam, A.K.M.R.; Sheperd, K.W. Incorporation of barley chromosomes into wheat. In *Biotechnology in Agriculture and Forestry*; Bajaj, Y.P.S., Ed.; Springer: Berlin, Germany, 1990; Volume 13, pp. 128–151.
39. Islam, A.; Shepherd, K.W.; Sparrow, D.H.B. Isolation and characterization of euplasmic wheat-barley chromosome addition lines. *Heredity* **1981**, *46*, 161–174. [CrossRef]
40. Tang, J.; Ohyama, K.; Kawaura, K.; Hashinokuchi, H.; Kamiya, Y.; Suzuki, M.; Muranaka, T.; Ogihara, Y. A new insight into application for barley chromosome addition lines of common wheat: Achievement of stigmasterol accumulation. *Plant Physiol.* **2011**, *157*, 1555–1567. [CrossRef]
41. Sharp, P.J.; Kreis, M.; Shewry, P.R.; Gale, M.D. Location of β-amylase sequences in wheat and its relatives. *Theor. Appl. Genet.* **1988**, *75*, 286–290. [CrossRef]
42. Marcotuli, I.; Colasuonno, P.; Blanco, A.; Gadaleta, A. Expression analysis of cellulose synthase-like genes in durum wheat. *Sci. Rep.* **2018**, *8*, 15675. [CrossRef] [PubMed]
43. Collins, H.M.; Burton, R.A.; Topping, D.L.; Liao, M.-L.; Bacic, A.; Fincher, G.B. Variability in Fine Structures of Noncellulosic Cell Wall Polysaccharides from Cereal Grains: Potential Importance in Human Health and Nutrition. *Cereal Chem.* **2010**, *87*, 272–282. [CrossRef]
44. Morgan, J.L.; Strumillo, J.; Zimmer, J. Crystallographic snapshot of cellulose synthesis and membrane translocation. *Nature* **2013**, *493*, 181–186. [CrossRef] [PubMed]
45. Zimmer, C.M.; Oliveira, G.; Pacheco, M.T.; Federizzi, L.C. Characterization and absolute quantification of the Cellulose synthase-like F6 homoeologs in oats. *Euphytica* **2020**, *216*, 173. [CrossRef]
46. Li, Y.; Cheng, X.; Fu, Y.; Wu, Q.; Guo, Y.; Peng, J.; Zhang, W.; He, B. A genome-wide analysis of the cellulose synthase-like (Csl) gene family in maize. *Biol. Plant.* **2019**, *63*, 721–732. [CrossRef]
47. McFadden, E.R.; Sears, E.R. The origin of Triticum spelta and its free-threshing hexaploid relatives. *J. Hered.* **1946**, *37*, 81–89. [CrossRef]
48. Schwerdt, J.G.; MacKenzie, K.; Wright, F.; Oehme, D.; Wagner, J.M.; Harvey, A.J.; Shirley, N.J.; Burton, R.A.; Schreiber, M.; Halpin, C.; et al. Evolutionary Dynamics of the Cellulose Synthase Gene Superfamily in Grasses. *Plant Phy. 2015 168*, 968–983. [CrossRef]
49. Schreiber, M.; Wright, F.; MacKenzie, K.; Hedley, P.; Schwerdt, J.; Little, A.; Burton, R.; Fincher, G.; Marshall, D.; Waugh, R.; et al. The barley genome sequence assembly reveals three additional members of the CslF (1,3;1,4)-beta-glucan synthase gene family. *PLoS ONE* **2014**, *9*, e90888. [CrossRef]

50. Zhang, S.; Hu, H.; Cui, S.; Yan, L.; Wu, B.; Wei, S. Genome-wide identification and functional analysis of the cellulose synthase-like gene superfamily in common oat (*Avena sativa* L.). *Phytochemistry* **2024**, *218*, 113940. [CrossRef]
51. Huang, H.; Zhao, S.; Chen, J.; Li, T.; Guo, G.; Xu, M.; Liao, S.; Wang, R.; Lan, J.; Su, Y.; et al. Genome-wide identification and functional analysis of Cellulose synthase gene superfamily in Fragaria vesca. *Front. Plant Sci.* **2022**, *13*, 1044029. [CrossRef]
52. Xi, H.; Liu, J.; Li, Q.; Chen, X.; Liu, C.; Zhao, Y.; Yao, J.; Chen, D.; Si, J.; Liu, C.; et al. Genome-wide identification of Cellulose-like synthase D gene family in Dendrobium catenatum. *Biotechnol. Biotechnol. Equip.* **2021**, *35*, 1163–1176. [CrossRef]

CHAPTER II

Is the *CslF6* gene involved in the accumulation of (1,3;1,4)-β-D-glucan in wheats, their wild relatives and their hybrids?

Background: Mixed linkage (1,3;1,4)-β-D-glucan (MLG) is a well-recognized bioactive carbohydrate and dietary fibre with expanding applications in the food industry. MLG are small components of the cell wall of vegetative tissues of cereals, synthesized by members of the *Cellulose Synthase-Like* genes (*Csl*). Within this gene family, *CslF6* has been the major contributor in wheat. Enhancing (1,3;1,4)-β-D-glucan content in wheat, a staple grain with naturally low (1,3;1,4)-β-D-glucan levels, has significant health and economic benefits.

AIM: This study investigated the role of the *CslF6* gene in (1,3;1,4)-β-D-glucan synthesis and analyzed total (1,3;1,4)-β-D-glucan contents, cell wall monosaccharide composition, glycosidic linkage profiles, and the major oligosaccharides comprising (1,3;1,4)-β-D-glucan in various wheat genotypes, their wild relatives (*Aegilops caudata* and *Dasypyrum villosum*), and hybrids between them. The aim was to assess the relationship between *CslF6* gene expression and (1,3;1,4)-β-D-glucan accumulation across different wheat lines.

Results: The study observed a clear relationship between *CslF6* gene expression and (1,3;1,4)-β-D-glucan accumulation in the different wheat lines. While *Aegilops caudata* and *Dasypyrum villosum* exhibited higher (1,3;1,4)-β-D-glucan content compared to other genotypes, hybrid breeding resulted in a significant increase in (1,3;1,4)-β-D-glucan content: 24.4% in durum wheat and 43.3% in *T. aestivum*. Additionally, variations were found in the ratios of major oligosaccharides released from (1,3;1,4)-β-D-glucan by lichenase treatment, as well as in the composition of cell wall monosaccharides and glycosidic linkages. This study demonstrates that HPAEC-PAD and GC-MS-based glycomics were invaluable tools for breeders in selecting high (1,3;1,4)-β-D-glucan lines.

Food Chemistry: Molecular Sciences 9 (2024) 100212

Contents lists available at ScienceDirect

Food Chemistry: Molecular Sciences

journal homepage: www.sciencedirect.com/journal/food-chemistry-molecular-sciences

Is the *CslF6* gene involved in the accumulation of (1,3;1,4)-β-D-glucan in wheats, their wild relatives and their hybrids?

Ilaria Marcotuli [a,*], Xiaohui Xing [b], Davide Caranfa [a], Stefania L. Giove [a], Yves S.Y. Hsieh [c,d], Shu-Chieh Chang [c], D. Wade Abbott [b], Agata Gadaleta [a]

[a] Department of Soil, Plant and Food Science, University of Bari Aldo Moro, Via G. Amendola 165/A, 70126, Bari, Italy
[b] Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, Alberta T1J 4B1, Canada
[c] Division of Glycoscience, Department of Chemistry, School of Engineering Sciences in Chemistry, Biotechnology and Health, Royal Institute of Technology (KTH), Stockholm SE-10691, Sweden
[d] School of Pharmacy, College of Pharmacy, Taipei Medical University, Taipei 11031, Taiwan

ARTICLE INFO

Keywords:
Wheat
Aegilops caudata
Dasypyrum villosum
Mixed-linkage glucan
CslF6 gene
Glycomics

ABSTRACT

Mixed linkage (1,3;1,4)-β-D-glucan (MLG) is a well-recognized bioactive carbohydrate and dietary fibre with expanding applications in food industry. The MLG are small components of the cell wall of vegetative tissues of cereals synthetized by members of the *Cellulose Synthase-Like* genes (*Csl*). Within the family, the *CslF6* has been the major contributor in wheat. It is of significant health and economic benefits to enhance MLG content in wheat, a staple grain with naturally low MLG levels. This study investigated the role of *CslF6* gene in MLG synthesis and analysed total MLG contents, cell wall monosaccharide, glycosidic linkage composition, and profile of major comprising oligosaccharides of MLG in various wheat genotypes, their wild relatives (*Aegilops caudata* and *Dasypyrum villosum*), and hybrids between them. We observed a relationship between *CslF6* gene expression and MLG accumulation across the different wheat lines. While *Aegilops caudata* and *Dasypyrum villosum* exhibited higher MLG content than other genotypes, hybrid breeding led to an increase in MLG content by 24.4% in durum wheat and 43.3% in *T. aestivum*. Variations in the ratios of major oligosaccharides released from MLG by lichenase treatment and in the compositions of cell wall monosaccharides and glycosidic linkages were also found. This study demonstrates that HPAEC-PAD and GC–MS-based glycomics are invaluable tools to assist breeders in selecting high MLG lines.

1. Introduction

Durum wheat (*Triticum turgidum* L. subsp. *durum*) is a widely cultivated cereal crop that plays a pivotal role in meeting daily caloric requirements in many countries and regions (Beres et al., 2020). Ranking as the 10th most important cereal globally, it accounts for 5 % of total wheat production and spans a planting area of 16 million hectares worldwide. Durum wheat is a rich source of proteins, starch and non-starch polysaccharides, vitamins, minerals, and other phytochemicals in the human diet (Grant, Cubadda, Carcea, Pogna, & Gazza, 2012; Marcotuli, Colasuonno, Hsieh, Fincher, & Gadaleta, 2020; Marcotuli, Soriano, & Gadaleta, 2022). Durum semolina, the product resulting from the milling of the hard-textured durum wheat kernel, is predominantly used in pasta production and also serves as a key ingredient in dishes like couscous and bulgur (Beres et al., 2020). Notably, in countries like Italy, regulatory standards mandate that pasta be made of 100 % durum semolina (Sopiwnyk, 2018).

Mixed linkage (1,3;1,4)-β-D-glucan (MLG), a non-starch polysaccharide characterized by a linear chain of 1,3- and 1,4-linked β-D-glucopyranosyl residues, is predominantly found in the kernels of cereal grains. MLG has demonstrated efficacy in reducing the risks of colorectal cancer, cardiovascular disease, and diabetes, the leading causes of morbidity and mortality in industrialized nations, offering both societal and individual benefits (Bhoite, Satyavrat, & Premasudha Sadananda). Globally recognized regulatory bodies, such as US Food and Drug Administration (FDA), Health Canada, and the European Commission have approved health claims highlighting its role in cholesterol management and associated cardiovascular complications (Mathews, Kamil, & Chu, 2020). These endorsements have enhanced the reputation of MLG, prompting its widespread incorporation in various functional food

* Corresponding author.
E-mail address: ilaria.marcotuli@uniba.it (I. Marcotuli).

https://doi.org/10.1016/j.fochms.2024.100212
Received 7 March 2024; Received in revised form 31 May 2024; Accepted 2 July 2024
Available online 5 July 2024

and nutraceutical formulas. Besides, MLG has been increasingly used as functional additives (*e.g.*, emulsion stabilizer, thickening agent, fat substitute) in food industry due to its water solubility, viscosity-enhancing properties, and gel-forming abilities (Maheshwari, Sowrirajan, & Joseph, 2017; Sun et al., 2021; Sushytskyi et al., 2023).

The biosynthesis of MLG in cereal crops has garnered significant research interest, particularly concerning the functions of associated genes and enzymes, and the effects of mutations on MLG content across various grains. MLG synthesis in grasses is catalyzed by members of Glycosyltransferase Family 2 that use uridine diphosphate glucose (UDP-Glc) for the synthesis process (Bain, van de Meene, Costa, & Doblin). The genes responsible for MLG synthesis share sequence homology with *cellulose synthases* (*CesAs*) and belong to the *Cellulose synthase-like* (*Csl*) sub-families F, H, and J (Burton et al., 2006; Doblin et al., 2009; Farrokhi et al., 2006; Little et al., 2019). Notably, the *CslF* and *CslH* groups were found exclusive to monocotyledons and are believed to either directly or indirectly regulate the accumulation and fine structure of MLG in the grain and other parts of the plant (Burton et al., 2011; Farrokhi et al., 2006). Additionally, it has been demonstrated that the barley *CslF6* is responsible for synthesis of β-1,4-linkage, but also as a conserved "switch motif" at the entrance of the enzyme's transmembrane channel, which is critical to generate (1,3)-linkages (Purushotham et al. 2022).

In *Brachypodium distachyon*, *CslF6* mutants showed altered carbon metabolism in MLG deficient grain (Bain, van de Meene, Costa, & Doblin, 2020). Mutations of *CslF6* in barley, rice, and wheat were found associated with reductions in MLG content, but these reductions generally had only moderate effects on vegetative growth and were well-tolerated (Cory, Båga, Anyia, Rossnagel, & Chibbar, 2012; Hu, Burton, Hong, & Jackson, 2014; Nemeth et al., 2010; Taketa et al., 2012; Tonooka, Aoki, Yoshioka, & Taketa, 2009; Wong et al., 2015). In wheat, it was observed that targeted RNA knockdowns of *CslF6*, driven by an endosperm-specific promoter, resulted in a 30–53 % decrease in MLG content in whole grain flour (Nemeth et al., 2010). In another study, knockout mutants of *CslF6* in rice exhibited a 97 % reduction in MLG levels in coleoptiles, with undetectable levels in other tissues (Vega-Sanchez et al., 2012). Compared to other cereal species, limited data on biosynthetic pathway of MLG is available for durum wheat, its wild relatives, and their hybrids.

Wheat naturally provides a low amount of MLG. For instance, barley and oat bran have MLG contents of 7 % and 5 %, respectively, whereas the MLG content of wheat bran is less than 1 % (Tiefenbacher, 2017). To introduce wheat-based products with naturally high MLG content to the market, innovative breeding strategies are needed to develop new wheat lines. Due to the limited genetic diversity within the wheat gene pool, incorporating wheat wild relatives into breeding programs is a promising strategy, achievable by introgressing small segments of their genome that carry desirable traits into wheat (Feuillet, Langridge, & Waugh, 2008). *Aegilops caudata* L. [syn. *Ae. markgrafii* (Greuter) Hammer], a diploid wild relative (2n = 2x = 14, CC) of hexaploid wheat, has been shown to harbor numerous resistance genes for both biotic and abiotic stresses and established as a valuable source of genetic variation for improving quality traits of wheat (Grewal et al., 2020; Ivanizs et al., 2022). There have been reports of the development of several wheats-*Aegilops* interspecific hybrids, addition lines, and translocation lines with many agronomically beneficial traits successfully integrated into the wheat gene pool from *Aegilops* (Schneider, Molnár, & Molnár-Láng, 2008). MLG was found to be an important cell wall component of *Aegilops* genus (Kishii, 2019). By incorporating chromosome 5U or group 7 chromosomes from the U and M genomes of *Aegilops geniculata* and *Aegilops biuncialis* species, the MLG content of *T. aestivum* significantly increased under various growth conditions (Rakszegi et al., 2017). For *Dasypyrum villosum*, another wild relative of wheat, there was only one MLG-related report that the incorporation of genes from the V chromosomes into durum wheat genomes improved both grain yield and grain quality; however, the durum wheat with the gene introgression exhibited MLG content similar to that of the wheat parental control (De Pace et al., 2014).

The current study was aimed at investigating and comparing the transcription levels of *CslF6* in the endosperm at various developmental stages across a range of genotypes, including *Aegilops caudata*, *Dasypyrum villosum*, durum wheat, *T. aestivum*, and amphyploid lines. We conducted enzymatic-colorimetric MLG assay, GC–MS-based monosaccharide and glycosidic linkage analyses, and HPAEC-PAD-based oligosaccharide profiling to examine the variations in content and structural features of MLG across samples. This study sought to establish a workflow for MLG-specific glycomics, assisting breeders, growers, and processors in selecting MLG-rich cereal products.

2. Materials and methods

2.1. Wheat materials

A set of six genotypes, described in Table 1, were grown in an experimental field in Valenzano, Bari, Italy. The seeds included the durum variety Creso, the *D. villosum* genotype, the durum-*villosum* amphyploid genotype (from now named as Amphyploid 1), the *T. aestivum* variety Alcedo, the *Ae. Caudata* genotype and the *T. aestivum-caudata* Amphyploid (from now named as Amphyploid 2). Seeds from these genotypes were germinated and, during the growing season, 10 g of nitrogen per m^2 was applied at the beginning of planting following standard cultivation practices. Single plants were hand-harvested at maturity, and the grain was stored at 4 °C. The grain was ground using a 1093 Cyclotec Sample Mill (Tecator Foss, Hillerød, Denmark) and passed through a 1-μm sieve. Endosperm samples from each genotype were collected at five developmental stages (6, 12, 18, 24, and 32 days post anthesis) and stored at −80 °C for subsequent RNA extraction and analysis.

2.2. Determination of MLG content using enzymatic colorimetric assay

The total content of MLG was determined in wholegrain flour obtained from mature kernels and in immature endosperm sampled at each developmental stage previously described using the commercially available Mixed-Linkage β-D-glucan Assay Kit (Megazyme International Ireland Ltd., Wicklow, Ireland), following the methodology described by the literature (McCleary & Codd, 1991). Barley material with known MLG content of 4.1 % (w/w) was included in the kit as a reference for the analysis. Three biological replicates, each with two technical replicates, were performed for each genotype.

2.3. Oligosaccharide profile of lichenase digest by high performance anion exchange chromatography coupled to amperometric detection (HPAEC-PAD)

HPAEC-PAD analysis was used to profile the oligosaccharides enzymatically released from MLG from kernels with different genotypes, as described in our previous report (Chang et al., 2023). This involved subjecting the oligosaccharides released by lichenase digestion in the Megazyme assay to solid phase extraction using Varian Bond Elut

Table 1
Genome, whole grain MLG content, and MLG major oligosaccharide ratios (DP3/DP4) for wheats, wild relatives, and their hybrid lines.

Line	Genome	MLG (w %)	St. dev.	DP3/DP4	St. dev.
T. durum	AABB	0.45	0.02	2.5	0.09
D. villosum	VV	4.36	0.04	1.8	0.09
Amphyploid 1	AABBVV	0.56	0.05	1.8	0.02
T. aestivum	AABBDD	0.90	0.01	1.8	0.04
Ae. caudata	CC	4.15	0.19	1.9	0.09
Amphyploid 2	AABBDDCC	1.29	0.03	1.8	0.03

Carbon columns (Agilent Technologies Inc., USA) with a concentration of 50 mg/mL. Elution was performed with 55 % (v/v) acetonitrile on a Dionex ICS-5000 HPAEC-PAD system (Thermo Fisher Scientific Inc., USA). The ratio of the trisaccharide G4G3G to the tetrasaccharide G4G4G3G was calculated and denoted as DP3/DP4.

2.4. Cell wall preparation

The whole cell walls of the durum wheats were prepared as de-starched alcohol insoluble residues (AIRs) following the published procedure (Wood et al., 2018), except that the hexane and ethyl acetate soaking steps were omitted, only thermostable amylase digestion was performed (replacing the combined use of amylase and amyloglucosidase), and the de-starched samples were dialyzed with a molecular weight cut off of 6,000–8,000 Da instead of 3,500 Da (Klassen et al., 2023). The dry pellets of AIRs were finely powdered by ball-milling in preparation for subsequent monosaccharide and linkage analyses.

2.5. Monosaccharide analysis

Monosaccharides were released from around 10 mg of dry AIR powder by heating in 2 mL of 1.25 M anhydrous methanolic hydrochloric acid at 80 °C for 16 h with gentle magnetic stirring, followed by evaporation under nitrogen until dryness and additional two rounds of evaporation to dryness in 2 mL of anhydrous methanol (Casillo, D'Angelo, Parrilli, Tutino, & Corsaro, 2022; Gottstein et al., 2021; Smith, O'Neill, Backe, York, Peña, & Urbanowicz, 2020). The samples were then magnetically stirred in 2 mL of deionized water containing 10 mg/mL of sodium borodeuteride at 4 °C for 24 h to carboxyl reduce uronic acid methyl esters to their C-6 dideuterated neutral sugar counterparts (Bacic, Moody, & Clarke, 1986; Badhan et al., 2022; Low et al., 2020). Excess reductant was quenched by glacial acetic acid, the reaction mixture was evaporated to dryness under nitrogen, and then the product was acetylated by heating in a mixture of 0.5 mL of trifluoroacetic acid and 2.5 mL of acetic anhydride at 60 °C for 1 h min (Chang et al., 2023; Jones et al., 2020; Robb et al., 2022; Voiges, Adden, Rinken, & Mischnick, 2012). After evaporation to dryness under nitrogen, the product were redissolved in 3 mL of dichloromethane followed by partition with 3 mL of saturated sodium bicarbonate deionized water one time then 3 mL of deionized water three times, passing the final lower phase through a glass wool plugged Pasteur pipette loaded with anhydrous sodium sulfate, and evaporation to dryness by nitrogen (Robb et al., 2022; Voiges et al., 2012; Yu et al., 2017). The resulting acetylated methyl glycosides were converted to alditol acetates by 2 M trifluoroacetic acid hydrolysis (100 °C, 2 h), reduction with sodium borodeuteride, and acetylation in heated mixture of trifluoroacetic acid and acetic anhydride (1:5, v/v) (Chang et al., 2023; Jones et al., 2020; Robb et al., 2022; Voiges et al., 2012). The derivatives were then cleaned up using partition and sodium sulfate column as described above, redissolved in ethyl acetate, and tested on a 6890 N GC-FID system (Agilent, United States) installed with an Supelco SP-2380 column (30 m × 0.25 mm × 0.20 μm, Sigma-Aldrich, United States) with oven temperature to start at 180 °C (hold 1 min) followed by increasing at 3 °C/min to 250 °C (hold 20 min). The same sample was also tested on a 7890B-5977B GC–MS system (Agilent, United States) installed with an Supelco SP-2380 column (100 m × 0.25 mm × 0.20 μm, Sigma-Aldrich, United States) with oven temperature to start at 220 °C followed by increasing at 1 °C/min to 250 °C (hold 30 min). Both systems had the same settings of inlet temperature (250 °C) and constant column helium flow of 0.8 mL/min. Splitless injection and split injection (10:1 ratio) were used for the GC-FID and GC–MS, respectively. The relative composition of alditol acetates was quantified using FID response factors obtained from standards, and the C-6 deuterated alditol acetates of uronic acids and alditol acetates generated from corresponding neutral monosaccharides were distinguished and determined based on GC–MS data, as described in the literature (Pettolino, Walsh, Fincher, & Bacic, 2012). Two separate experiments were conducted for each sample.

2.6. Glycosidic linkage analysis

Uronic acids in around 10 mg of dry AIR powder were carboxyl reduced to 6,6-dideuterated neutral sugars by sodium borodeuteride reduction of their methyl esters generated by weak methanolysis (0.5 M methanolic HCl, 80 °C, 20 min) (Chong, Cleary, Dokoozlian, Ford, & Fincher, 2019; Hosain, Ghosh, Bryant, Arivett, Farone, & Kline, 2019; Muhidinov et al., 2020). After quenching excess reductant with acetic acid and removal of borate by repeated evaporation to dryness in mixture of acetic acid and methanol (1:10, v/v) then in anhydrous methanol, the sample was methylated in a mixture of 1.2 mL of methyl iodide and 2 mL of DMSO with 100 mg of dry sodium hydroxide powder (Badhan et al., 2022; Jones et al., 2020; Low et al., 2020). The methylation product was cleaned up by partition between 3 mL of dichloromethane and 6 mL of 10 % acetic acids over ice two times then with 6 mL of deionized water two times. The lower phase was evaporated to dryness, and the methylation process was repeated one more time. For the first round of methylation, the sample was magnetically stirred in DMSO at 60 °C overnight, with the tube headspace filled with N_2, prior to the methylation reaction. Before the second round of methylation, the same process of stirring in DMSO was conducted overnight but at room temperature, instead of 60 °C. The methylated samples were then converted to partially methylated alditol acetates (PMAAs) by first underwent 2 M trifluoracetic acid hydrolysis (120 °C, 2 h). The hydrolysis product was then reduced by sodium borodeuteride before all free hydroxyl groups were acetylated under a mixture of trifluoroacetic acid and acetic anhydride (1:5, v/v) (Chang et al., 2023; Robb et al., 2022; Voiges et al., 2012). The PMAAs were redissolved in ethyl acetate and tested on the Agilent 7890B-5977B GC–MS system with the same 100 m Supelco SP-2380 column as described above for monosaccharide analysis, except that the oven temperature was programmed to start at 100 °C followed by increasing at 1.5 °C/min to 220 °C then at 1.25 °C/min to 250 °C (hold 20 min). The EI-MS spectra of the PMAAs were interpreted by comparing them with those of reference derivatives and by referring to the literature (Carpita & Shea, 1989). The glycosidic linkage composition (mol%) was calculated following the published protocol (Pettolino et al., 2012).

2.7. RNA extraction and cDNA synthesis

The expression analysis of the gene in the endosperm of the two wheat varieties, *Ae. caudata*, *D. villosum,* and amphyploid genotypes was conducted using the primer pairs described in our previous report (I. Marcotuli, P. Colasuonno, A. Blanco, & A. Gadaleta, 2018). To analyze the expression level of *CslF6*, total RNA was extracted from the endosperm of each genotype using the RNeasy Plant Mini Kit (Qiagen, Valencia, USA) and assessed on a 1.5 % denaturing agarose gel. All RNA samples were adjusted to the same concentration (1 μg/mL) and reverse-transcribed into double-stranded cDNA with the QuantiTect Reverse Transcriptase Kit (Qiagen, Valencia, USA). Data were normalized using three reference genes: *Cell Division Control AAA-Superfamily of ATPases* (*CDC*), *ADP-Ribosylation Factor* (*ADP-RF*), and *RNase L Inhibitor-like protein* (*RLI*). These genes had a stability value of approximately 0.035 when evaluated with NormFinder software (Andersen, Jensen, & Ørntoft, 2004).

2.8. qPCR for cellulose synthase gene

Quantitative real-time PCR (qRT-PCR) was performed using Cyber® GREEN on a CFX96TM Real-Time PCR Systems (Bio-Rad Laboratories, Hercules, USA), adhering to the amplification procedure described in our previous report (Marcotuli et al., 2018). The specificity of the amplicons was validated through several means: by observing a single band of the expected size for each primer pair on a 2 % (w/v) agarose

gel, by noting a singular peak in the melting curves of the PCR products, and by sequencing the amplified fragments using a 3500 Genetic Analyzer (Applied Biosystems, Waltham, USA). qRT-PCR data were derived from the mean values of three independent amplification reactions conducted on five distinct plants harvested at the same phenotypic stage (biological replicates). All calculations and analyses were conducted using the ΔCt method with the CFX Manager 2.1 software (Bio-Rad Laboratories, Hercules, USA). Standard deviations were utilized to normalize values for the highest or lowest individual expression levels as per the CFX Manager 2.1 software user manual. ANOVA and the LSD test were employed to evaluate significant differences in the expression of the *CslF6* gene between genotypes.

3. Results and discussion

The health and economic benefits associated with MLG have led to a search for wild alleles that can increase the level of MLG in wheat, primarily focusing on species within the genus *Triticum* (Marcotuli et al., 2015; Marcotuli et al., 2016). Wild relatives of wheat from the genus *Aegilops* and *Dasypyrum*, which showed MLG content around 4 % and are critical sources of new genes and alleles for wheat breeding, have been largely overlooked in breeding efforts aimed at developing high MLG wheat. The current study investigated the role of the *CslF6* gene in MLG accumulation in wheat, its wild relatives *Aegilops caudata* and *Dasypyrum villosum*, and their hybrids, and also aimed to enhance the efficiency and accuracy of identifying high MLG lines using MLG-specific glycomics.

3.1. *CslF6* expression in mature kernels and endosperm at different developmental stages

The transcript levels of *CslF6* mRNA were determined in the endosperm of bread and durum wheats, *Ae. caudata*, *D. villosum*, and the two corresponding amphyploid during various developmental stages, spanning from 6 to 32 days post anthesis (DPA) (Fig. 1). For *T. durum*, the normalized expression of *CslF6* mRNA remained relatively low during the initial days, peaking significantly at 18 DPA with a value of 9.0, then declining in subsequent measurements. *D. villosum* displayed a drastic surge at 24 DPA, reaching the highest expression level of 18.5 before descending to 10.7 at 32 DPA. The hybrid of *T. durum* and *D. villosum* (Amphyploid 1) exhibited a steady increase in expression up to 32 DPA, reaching a peak value of 16.7. *T. aestivum* showed a notable spike at 24 DPA with a value of 11.0. It seems that *Ae. Caudata* reached peak expressions similar to that of *D. villosum* at 24 DPA.. The hybrid of *T. aestivum* and *Ae. caudata* demonstrated high expression levels at 18 DPA (7.5), followed by a dramatic drop to 0.2 by 32 DPA. In comparison, both *T. durum* and *Ae. caudata* reached peak expressions around the middle of the developmental stages, while *D. villosum* and *T. aestivum* presented higher expressions slightly later, around 24 DPA. Amphyploid 1 showed a sharp increase at 18DPA rather than consistent increase. Amphyploid, while the Amphyploid 2 demonstrated a severe decline post its peak. Such variances in expression patterns could be indicative of the genetic and regulatory influences that each species or hybrid inherits and how these influences affect the expression of the *CslF6* gene during endosperm development. These results were in good agreement with previous studies on the expression patterns of *CslF6* in developing endosperm of durum wheat. For instance, similar trends of low *CslF6* expression were previously reported in the early stages of endosperm development, followed by a significant increase in expression during the mid to late stages (Ilaria Marcotuli, Pasqualina Colasuonno, Antonio Blanco, & Agata Gadaleta, 2018). These findings suggested a conserved regulatory mechanism governing *CslF6* expression during endosperm development across different genotypes. The observed increase in *CslF6* expression at 18 DPA was in good agreement with the previous report that a set of genes in *Brachypodium*, including *CslF6*, exhibited peak expression levels during this stage of endosperm development (Francin-Allami et al., 2023). The upregulation of *CslF6* at 18 DPA might indicate its involvement in specific developmental processes, such as cell division or cell wall synthesis, crucial for endosperm development during this period. A distinct expression pattern of *CslF6* in *Triticum aestivum*, showing a gradual increase in *CslF6* expression throughout development, was previously reported (Nemeth et al., 2010). It is worth noting

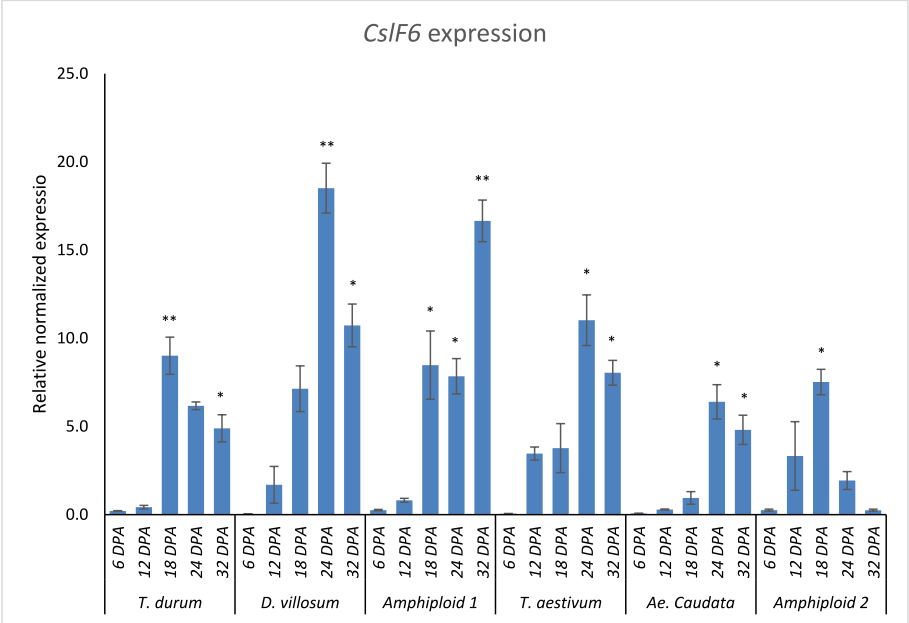


Fig. 1. Normalized expression levels for *CslF6* gene in the developing endosperm at various times (days) following the initiation of maturation at 6, 12, 18, 24, and 32 days post anthesis (DPA) analysed in durum wheat (*T. durum*), *T. aestivum*, *Ae. Caudata*, *D. villosum*, and amphyploid lines. Amphyploid 1 refers to the hybrid between *T. durum* and *D. villosum*. Amphyploid 2 denotes the hybrid between *T. aestivum* and *Ae. caudata*. The height of each bar represents the mean value, with the standard deviation denoted by error bars labelled at the top. Asterisks indicate genotypes significantly different within the developmental stages (** $P \leq 0.01$; * $P \leq 0.05$).

that while our study focused on the expression levels of *CslF6* mRNA during endosperm development, the functional implications of *CslF6* in this context should not be overlooked. A previous study demonstrated that down-regulation of *CslF6* expression resulted in altered cell wall composition and reduced grain filling in transgenic lines (Nemeth et al., 2010). This suggested that *CslF6* plays a critical role in endosperm development, particularly in the biosynthesis and organization of cell wall components. The current study sheds light on the role of *CslF6* in endosperm development, but further research is needed to pinpoint its precise functions, its contribution in various genotypes, and its interplay with genes involved in cell wall biosynthesis and grain filling.

3.2. Total MLG content determined by enzymatic colorimetric assay

The total MLG amount in mature grains was tested using a well-established enzymatic and colorimetric method with commercial kit from Megazyme (McCleary et al., 1991). The amount of MLG was expressed as weight percentage (w%) based on the dry weight of the kernel. Results showed that the whole grain MLG values ranged from 0.45 % (durum wheat) to 4.15 % (*Ae. caudata*), with significant differences ($p \leq 0.01$) between the genotypes (Table 1). The *T. aestivum-caudata* amphyploidamphyploid genotype displayed a much lower whole grain MLG content (1.29 %) than the *caudata* line (4.15 %). Similarly, the durum wheat-*villosum* amphyploidamphyploid genotype had a low whole grain MLG content of 0.56 %. Highest value was shown from *D. villosum* (4.36 %). Both amphyploidamphyploid lines exhibited low MLG amounts similar to those of the parental lines. The results were in good agreement with previous reports that bread and durum wheats were not significant sources of MLG because of their low content in the grain, typically less than 1 % on a dry weight basis (Collins, Burton, Topping, Liao, Bacic, & Fincher, 2010; Marcotuli et al., 2016). However, it is worth noting that higher MLG values, up to 2.3 %, have been reported in *T. aestivum* (Fincher & Stone, 2004). The relatively high concentration of MLG in wheat grain was mainly found in the sub-aleurone layer, with lower amounts present in the rest of the endosperm (Beresford & Stone, 1983).

In the developing endosperm (6, 12, 18, 24, and 32 DPA), the MLG contents across the different species showed distinct patterns (Fig. 2). *T. durum* started with an MLG content of 0.50 % at 6 DPA, saw a slight decline by 18 DPA to 0.27 %, and then gradually rebounded to 0.35 % by 32 DPA, suggesting relative stability in its low MLG levels over time. *D. villosum*, with an initial value of 0.98 % at 6 DPA, consistently increased, reaching its highest value of 3.54 % by 32 DPA. The hybrid of *T. durum* and *D. villosum* maintained a relatively stable trend, initiating at 0.25 % and increasing slowly to 0.54 % over the 32 days. *T. aestivum* exhibited a narrow range of fluctuation, starting at 0.60 % and showing a minor peak of 0.77 % at 32 DPA. *Ae. caudata*, with its onset at 0.74 %, saw a significant increase in its MLG content, reaching 4.06 % by 24 DPA, and preserved this level until 32 DPA. The hybrid lineage of *T. aestivum* and *Ae. caudata* had a relatively high value of 1.27 % at 6 DPA, but then stabilized around 1.0 %, concluding at a slightly reduced value of 0.82 % by 32 DPA. These results demonstrated the dynamic changes in MLG content during kernel development across different varieties. The variations observed in MLG content among the developmental stages and varieties suggested the potential influence of genetic factors and developmental processes on MLG accumulation in kernels. A noticeable rise in the content of MLG from 4 to 28 DPA during *T. aestivum* grain development was observed in previous study (Palmer, Cornuault, Marcus, Knox, Shewry, & Tosi, 2015). That study also revealed that the most pronounced increase occurred during the early stages of development (4–12 DPA), after which the growth rate decelerated as it approached maturity.

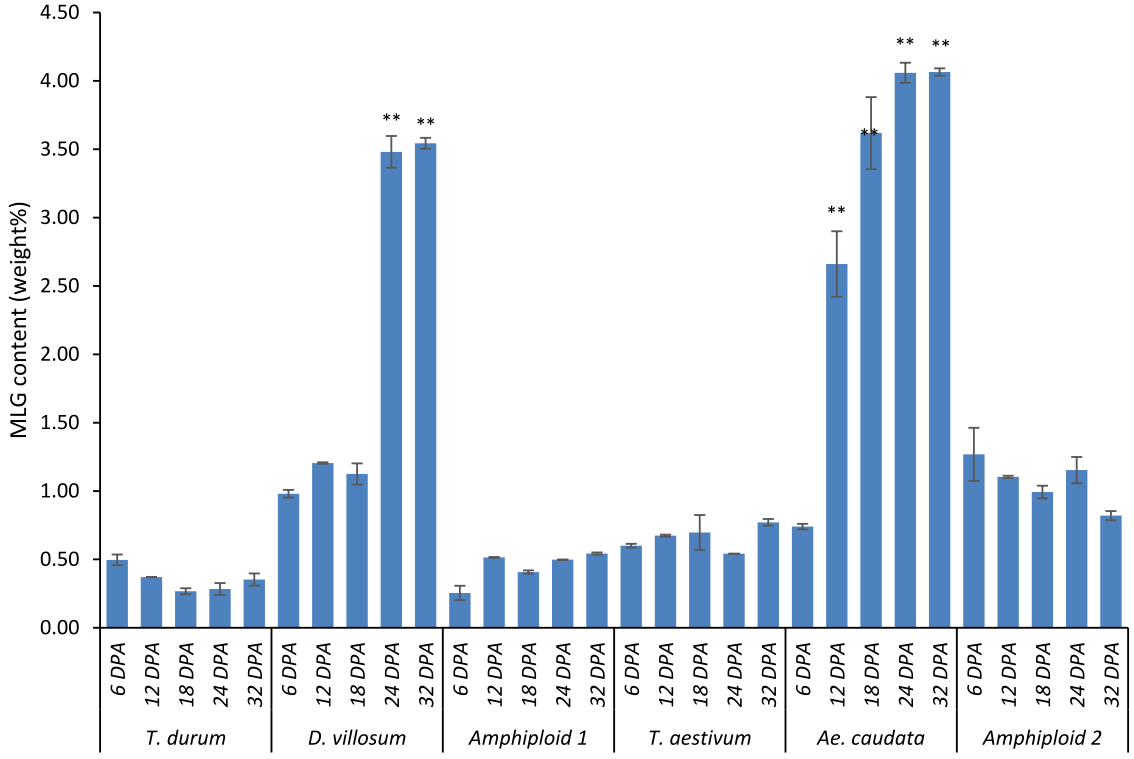


Fig. 2. MLG content detected in the developing endosperm at various times (days) following the initiation of maturation at 6, 12, 18, 24, and 32 days post anthesis (DPA) analysed in durum wheat (*T. durum*), *T. aestivum*, *Ae. Caudata*, *D. villosum*, and amphyploid lines. Amphyploid 1 refers to the hybrid between *T. durum* and *D. villosum*. Amphyploid 2 denotes the hybrid between *T. aestivum* and *Ae. caudata*. The height of each bar represents the mean value, with the standard deviation denoted by error bars labelled at the top. Asterisks indicate genotypes significantly different within the developmental stages (** $P \leq 0.01$; * $P \leq 0.05$).

3.3. HPAEC-PAD analysis of oligosaccharides released from kernels by lichanase treatment

The molecular traits of MLG are typically derived from HPAEC-PAD analysis of oligosaccharides generated by lichenase with a 3-linked reducing end, while all other residues are 4-linked (Chang et al., 2023). Variations exist within the same cereal species due to factors such as genotype and environmental conditions (Jiang & Vasanthan, 2000; Miller, Vincent, Weisz, & Fulcher, 1993; Storsley, Izydorczyk, You, Biliaderis, & Rossnagel, 2003). Chromosome 5U was found to contribute to a reduction in the DP3/DP4 ratio of MLG (Rakszegi et al., 2017). The DP3/DP4 ratio has been used to profile MLG from different sources of cereals because, among all oligosaccharides detectable by HPAEC-PAD, trisaccharide and tetrasaccharide units make up 90–95 % of total oligosaccharides, with longer chains accounting for only 5–10 % (Lante, Canazza, & Tessari, 2023). Barley, wheat, oat, and rice were reported to have DP3/DP4 ratio ranges of 2.3–3.4, 1.5–2.3, 3.0–4.5, and 2.4–2.7, respectively (Lante et al., 2023). In the current study, the DP3/DP4 ratios were determined to be 1.8 for *villosum*, wheat-*villosum* amphyploid, *T. aestivum*, and *T. aestivum-caudata* amphyploid genotypes, 1.9 for the *caudata* line, and 2.5 for the durum wheat line (**Table 1**).

3.4. Monosaccharide and linkage analysis by GC–MS/FID

Whole cell wall polysaccharides of the durum wheat kernel were prepared as de-starched AIR using a well-established procedure that has been proven effective on many higher plant samples (e.g., chickpea kernel, canola meals) (Badhan et al., 2022; Klassen et al., 2023; Li et al., 2022; Li et al., 2019; Low et al., 2020; Wood et al., 2018). DMSO soaking was employed to solubilize all starches, including resistant starches (Sowinski, Gilbert, Lam, & Carpita, 2019). These can be hydrolysed to monosaccharides and oligosaccharides and then further removed by dialysis. It is commonly known that higher plant cell wall polysaccharides consist mainly of hemicellulose, cellulose, and pectins (Pettolino et al., 2012). Cellulose and MLG are both β-D-glucans with linear chains of glucopyranosyl residues, with the distinction that the glucopyranosyl residues are 4-linked only in the former but a mixture of 3- and 4-linked in the latter. The homo-linkage nature of cellulose results in it forming a highly crystalline structure in the cell wall by strong hydrogen bonding formed between neighbouring polysaccharide chains, even though a minor amount of amorphous cellulose also exits naturally in higher plant cell walls. In contrast, the mixed-linkage structural nature of MLG does not form strong crystalline regions because 3-linked residues introduce a kink that prevents the formation of strong interchain hydrogen bonding and thus crystalline region. It is also well-known that weak acid-catalysed hydrolysis methods, such as TFA and methanolysis, cannot break the crystalline structures of cellulose but can hydrolyse non-crystalline cell polysaccharides such as pectins (*e.g.*, homogalacturonan and rhamnogalacturonans) and hemicelluloses (*e.g.*, arabinoxylan, heteroxylan, heteromannans, and MLG) (Bertaud, Sundberg, & Holmbom, 2002; Biswal et al., 2022; Foster, Martin, & Pauly, 2010; Schäfer, Hale, Hoffmann, & Bunzel, 2020; Tingley, Low, Xing, & Abbott, 2021; Willför et al., 2009; Wilson, Deligey, Wang, & Cosgrove, 2021). In the current study, methanolysis (1.25 M HCl in methanol, 80 °C, 16 h) was used to release monosaccharides from non-crystalline cell wall polysaccharides of durum wheat kernels. Results showed that the relative compositions of glucose in the AIRs of *D. villosum* (45.1 %) and *Ae. caudata* (40.1 %) were much higher than those in the *T. durum* (9.7 %), Amphyploid 1 (14.5 %), *T. aestivum* (14.3 %), and Amphyploid 2 (13.5 %) (Table 2**,** Fig. 3). This finding could be due to the presence of other genes responsible for MLG synthesis in wheats, the quantitative nature of the trait and the polygenic control by genes with additive effects (Marcotuli et al. 2016; Houston et al. 2014). Given that crystalline cellulose and MLG are the two major glucose-containing non-starch polysaccharides in the kernel and that the glucose in crystalline cellulose cannot be released by acid methanolysis, the detected glucose was mainly from MLG. Therefore, the result indicated that the relative MLG composition in the cell walls of DV and AC kernels were much higher than in the other samples.

Table 2
Monosaccharide composition (mol%) of non-crystalline cell wall polysaccharides of kernels of wheats, wild relatives, and their hybrid lines.

Monosaccharide	TD	DY	A1	TA	AC	A2
Ara	43.2	19.2	37.0	38.3	21.6	41.4
Gal	6.1	5.4	7.4	7.7	5.5	6.3
GalA	1.6	1.2	1.2	1.2	1.2	1.3
Glc	9.7	45.1	14.5	14.3	40.1	13.5
GlcA	1.7	0.9	1.8	2.1	0.9	1.9
Man	1.8	2.1	3.7	3.8	1.9	3.1
Xyl	35.9	26.1	34.4	32.5	28.7	32.4

Note: Each experiment was conducted in duplicate to generate a mean. **TD**: *T. durum*; **DV**: *D. villosum*; **A1**: *Amphyploid 1*; **TA**: *T. aestivum*; **AC**: *Ae. caudata*; **A2**: Amphyploid 2.

Crystalline cellulose in insoluble cell wall samples can be depolymerized by hydrolysis using water solution of strong acid (*e.g.*, sulphuric acid); however, the hydrolysis condition is so harsh that degradation of released monosaccharides can occur, resulting in a reduced level of detected monosaccharides (Black, Heiss, & Azadi, 2019). Per-*O*-methylation of polysaccharide by methyl iodide in DMSO in the presence of dry sodium hydroxide powder (the Ciucanu Method) has been proved to be effective in completely decrystallising and methylating crystalline cellulose structure (Black et al., 2019; Ciucanu, 2006). The per-*O*-methylated decrystalised cellulose can be depolymerized by relatively weak TFA hydrolysis, and the released methylated monosaccharides can be converted to their partially methylated alditol acetate (PMAA) derivatives for analysis of glycosidic linkage composition using GC–MS/FID (Pettolino et al., 2012). This methylation analysis procedure has been successfully conducted to many whole cell wall samples (Badhan et al., 2022; Klassen et al., 2021; Low et al., 2020; Pham, Kyriacou, et al., 2019; Pham, Schwerdt, et al., 2019; Wood et al., 2018). For the determination of pectin, carboxyl reduction of uronic acids to their corresponding C-6 dideuterium labelled neutral sugar is required before the methylation step by sodium borodeuteride reduction of uronic acid esters generated by carbodiimide activation (Kim & Carpita, 1992) or weak methanolysis treatment (0.5 M methanolic HCl, 80 ℃, 20 min) (Chong et al., 2019; Hosain et al., 2019; Muhidinov et al., 2020). In the current study, the latter carboxy reduction method was used for the whole kernel cell walls, followed by methylation-GC–MS/FID analysis of the carboxyl reduced samples. Results showed that the *D. villosum* and *T. aestivum* cell walls contained much higher levels of 3-linked glucopyranoses than the others (Table 2**,** Fig. 4). The levels of 3-linked glucopyranoses were very similar among *T. durum*, Amphyploid 1, *T. aestivum*, and Amphyploid 2. The results of 3-linked glucopyranoses were in good agreement with the results of total MLG contents. As expected, abundant 4-linked glucopyranoses were detected, originating from the 4-linked residues of MLG and cellulose that were fully decrystallised by methylation. EI-MS spectra of PMAAs from 3-linked glucopyranose and 4-linked glucopyranose from *D. villosum* kernel are shown in Fig. 5.

In addition to the linkages from MLG and cellulose, we also found in the samples many other linkages (Table 2) from various cell wall polysaccharide structures such as arabinoxylan, arabinan, heteroxylan, heteromannan, arabinogalactan, and homogalacturonan according to previous studies on linkage assignment of higher plant cell walls (Badhan et al., 2022; Li et al., 2022; Li et al., 2019; Low et al., 2020; Pettolino et al., 2012; Wood et al., 2018). The weak metanalysis (0.5 M methanolic HCl, 80 ℃, 20 min)-sodium borodeuteride method is a quicker, less expensive carboxyl reduction method that allows larger sample amount (*e.g.*, 10 mg of AIR) and larger throughput compared to the conventional carbodiimide activation-sodium borodeuteride method. This metanalysis-based method has been recently used for

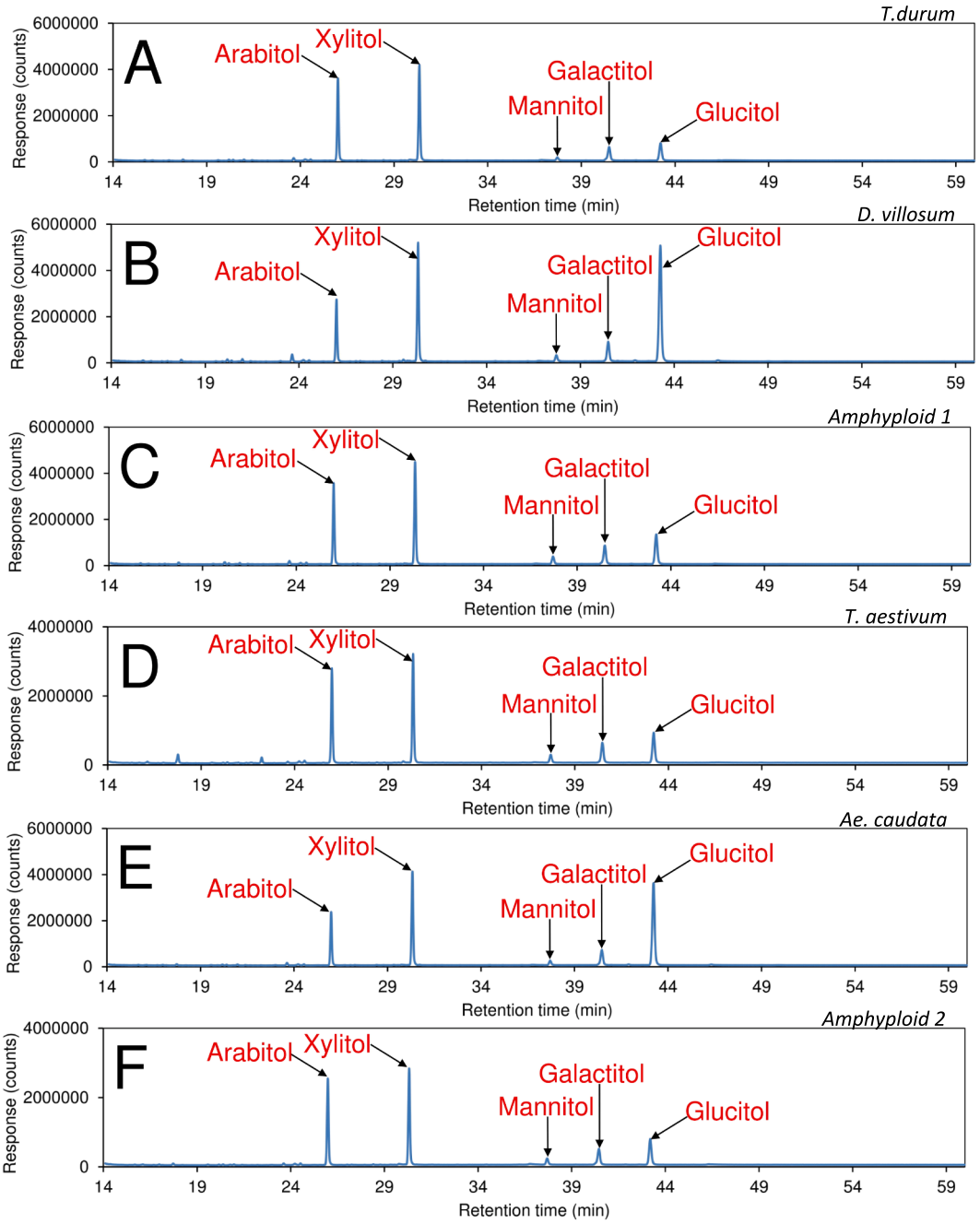


Fig. 3. Total ion current (TIC) chromatograms of alditol acetate derivatives of the six kernel cell wall samples. **TD**: *T. durum*; **DV**: *D. villosum*; **A1**: Amphyploid 1; **TA**: *T. aestivum*; **AC**: *Ae. caudata*; **A2**: Amphyploid 2. Amphyploid 1 refers to the hybrid between *T. durum* and *D. villosum*. Amphyploid 2 denotes the hybrid between *T. aestivum* and *Ae. caudata*.

linkage analysis of some cell wall polysaccharide mixtures and purified fractions containing uronic acids (Chang et al., 2023; Robb et al., 2022; Voiges et al., 2012). This was the first instance of the method being used for carboxyl reduction of kernel cell walls of wheat and its wild relatives. This technique allowed the detection of galacturonic acid linkages in our samples. However, we noted that the method resulted in significant debranching of arabinoxylans. This was evident from the diminished intensity of various arabinofuranose linkages, a reduced presence of branching xylopyranosyl residues, and a pronounced peak of 4-linked xylopyranose residues (the building unit of the debranched xylan backbone). It was apparent that the methyl glycosides of arabinose, resulting from the debranching of arabinoxylan by weak methanolysis, were transformed into highly volatile permethylated methyl glycosides that were lost by repeated evaporation-to-dryness during methylation analysis. Given the minimal amounts of pectins in wheat and its wild relatives, we recommend bypassing the carboxyl reduction for kernel cell walls in future studies unless the primary objective is to extract pectin linkage information. If linkage details for both pectins and arabinoxylans are crucial for future studies, the traditional carbodiimide activation-sodium borodeuteride reduction method should be employed.

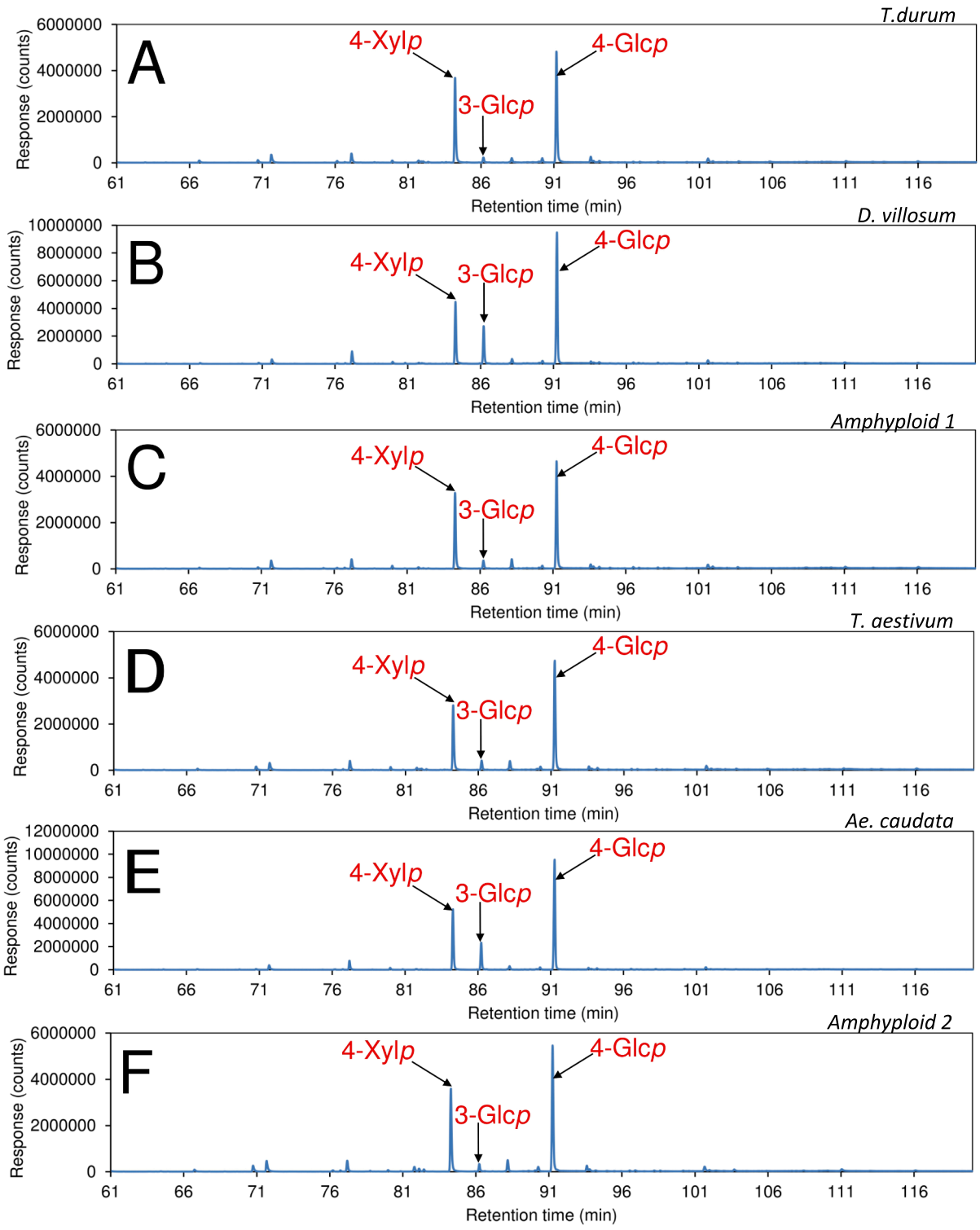


Fig. 4. Total ion current (TIC) chromatograms of partially methylated alditol acetate (PMAA) derivatives of the kernel cell walls of (A) *T. durum*, (B) *D. villosum*, (C) Amphyploid 1, (D) *T. aestivum*; (E) *Ae. Caudata*, and (F) Amphyploid 2. Amphyploid 1 refers to the hybrid between *T. durum* and *D. villosum*. Amphyploid 2 denotes the hybrid between *T. aestivum* and *Ae. caudata*.

3.5. Correlation between CslF6 expression and MLG accumulation and the significance of kernel MLG content enhancement achieved through hybrid breeding

The synthesis of MLG in barley grains is influenced by the expression of *CslF6*, as demonstrated by the significant enhancement in MLG content resulting from endosperm-specific gene overexpression (Burton et al. in 2011). Similar trends have been observed in wheat and rice, where *CslF6* orthologs play analogous roles (Nemeth et al., 2010). We analyzed the correlation between gene expression and MLG accumulation in our genetic material. We found that the correlation between *CslF6* gene expression and MLG accumulation in amphyploid lines was slightly positive but not remarkable ($p = 0.08$, Pearson test). The expression level of *CslF6* in the endosperm (32 DPA) of the durum wheat

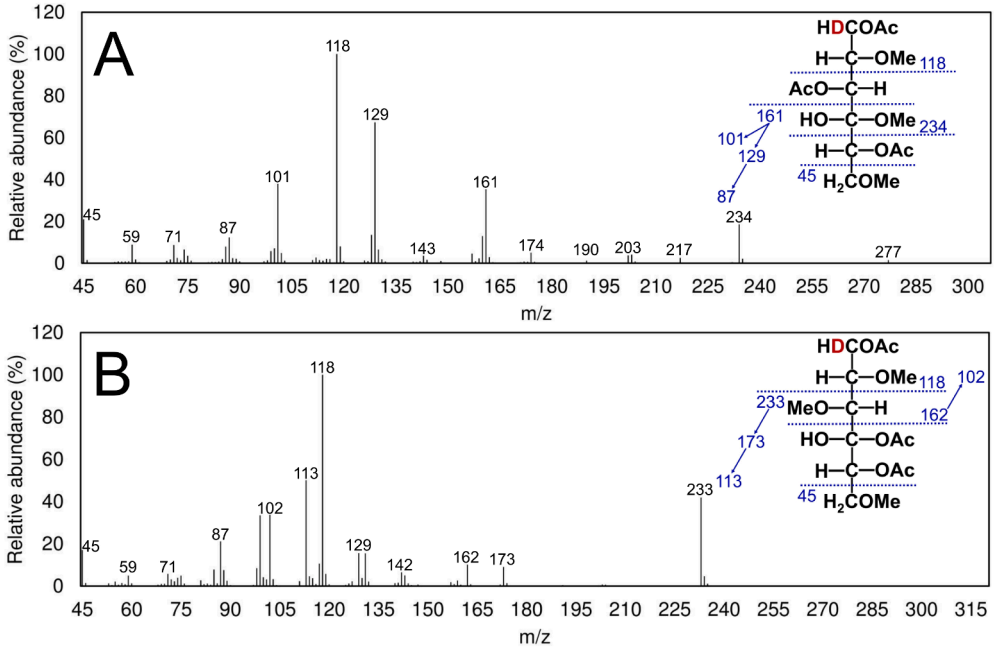


Fig. 5. EI-MS fragmentation patterns of the partially methylated alditol acetates from (A) 1,3-linked glucopyranose and (B) 1,4-linked glucopyranose.

amphyploid was 3 times that of the durum wheat parent line (Fig. 1), while the kernel MLG content was moderately increased from 0.45 % in the parental line to 0.56 % in the hybrid, an enhancement of 24.4 % (Table 1). The kernel MLG content in the hybrid line of *T. aestivum* was 43.3 % higher than in the parent line (Table 1). In contrast, the *CslF6* expression levels in the endosperm harvested at 24 and 32 DPA of the *T. aestivum* parental lines were 5.7 and 31.9 times those of the hybrid lines, respectively, indicating a strong negative correlation between the endosperm *CslF6* expression and kernel MLG content. These observations could be attributed to the presence of a degradation complex, formed as a result of past breeding programs primarily aimed at enhancing wheat productivity. The final products of this complex serve as precursors for the biosynthesis and accumulation of starch in the kernel. This was supported by a previous report of a strong and inverse relationship between MLG and starch biosynthesis (Burton et al., 2008). It was also reported that MLG and amylose levels are controlled by numerous genes interconnected through expression networks, all contributing to the regulation of MLG and amylose metabolism and biosynthesis (Islamovic et al., 2013). The presence of a conserved 'switch motif' at the entrance of the enzyme's transmembrane channel plays a crucial role in the production of 1,3-linked β-D-glucopyranoses, and within this region, a single-point mutation can significantly impact formation of the 1,3-linkage, leading to increased synthesis of cellulosic polysaccharides (Purushotham, Ho, Yu, Fincher, Bulone, & Zimmer, 2022). Therefore, further investigation is needed to clarify the genetic mechanism and to study protein accumulation and structure. See (Table 3).

In the current study, disproving the existence of a remarkable positive correlation between *CslF6* gene expression and MLG content in kernel does not diminish the significance of the MLG content enhancement by 24.4 % for durum wheat and 43.3 % for *T. aestivum* achieved through hybrid breeding. Wheat naturally contain less MLG than oat and barley (Tiefenbacher, 2017). However, people consume wheat products more than those of oat or barley. The inherent properties of MLG as a bioactive polysaccharide and dietary fiber and the omnipresence nature of wheat mean that even a slight increase in MLG content in wheat can significantly enhance the nutritional quality of the human diet. Enhancing MLG synthesis in wheat through innovative breeding techniques could lead to the production of wheat flour with the higher nutritional and functional quality that food processor seeks. This advancement would also elevate the potential of wheat bran, a byproduct of the milling process of wheat grain, as a material for the extraction and production of high-quality purified MLG, a value-added food additive in the industry (Maheshwari et al., 2017).

Table 3
Glycosidic linkage composition (mol%) of different kernel cell walls with pretreatment of weak methanolysis (0.5 M methanolic HCl, 80 °C, 20 min)-sodium borodeuteride reduction.

Linkage	TD	DY	A1	TA	AC	A2
t-Ara*p*	1.5	0.6	1.0	2.0	0.6	3.0
t-Ara*f*	1.1	0.3	0.6	0.8	0.4	1.0
2-Ara*f*	0.8	0.2	0.4	0.4	0.3	0.6
3-Ara*f*	0.1	0.1	0.1	0.1	0.1	0.2
5-Ara*f*	1.0	0.4	0.8	0.9	0.5	2.0
t-GalA*p*	0.1	0.1	0.1	0.1	0.1	0.1
4-GalA*p*	1.1	0.7	0.7	0.8	0.6	1.0
t-Gal*p*	0.9	0.8	1.1	1.2	0.8	0.5
3-Gal*p*	1.3	0.2	1.7	1.7	0.2	0.1
4-Gal*p*	0.5	0.3	0.4	0.5	0.3	0.6
6-Gal*p*	0.8	0.5	0.6	0.7	0.6	0.4
3,6-Gal*p*	0.2	0.2	0.5	0.6	0.2	0.5
t-GlcA*p*	2.3	1.0	1.8	2.0	1.0	1.9
t-Glc*p*	1.5	3.7	2.5	2.0	3.4	1.7
3-Glc*p*	2.2	12.8	3.1	3.6	10.4	2.4
4-Glc*p*	40.5	45.6	42.0	42.0	45.9	40.1
6-Glc*p*	0.4	0.2	0.3	0.4	0.3	0.3
2,4-Glc*p*	0.3	0.3	0.4	0.3	0.3	0.3
3,4-Glc*p*	0.3	0.3	0.6	0.3	0.4	0.2
3,6-Glc*p*	0.1	0.3	0.4	0.2	0.3	0.2
4,6-Glc*p*	0.5	0.2	0.6	0.4	0.2	0.5
t-Man*p*	0.3	0.2	0.5	0.4	0.3	0.4
4-Man*p*	1.0	1.7	1.8	1.9	1.4	3.7
t-Xyl*p*	4.1	2.3	4.1	4.4	2.9	4.9
2-Xyl*p*	2.5	1.3	2.2	2.5	1.5	2.8
4-Xyl*p*	30.8	23.5	28.4	26.6	24.9	26.4
2,4-Xyl*p*	1.7	0.6	1.3	1.2	0.6	1.6
3,4-Xyl*p*	0.7	0.3	0.8	0.5	0.4	0.8
2,3,4-Xyl*p*	1.5	1.1	1.3	1.6	1.1	1.6

Note: Each experiment was conducted in duplicate to generate a mean. **TD**: *T. durum*; **DV**: *D. villosum*; **A1**: Amphyploid 1; **TA**: *T. aestivum*; **AC**: *Ae. caudata*; **A2**: Amphyploid 2.

4. Conclusions

The results presented in this study shed light on the complex relationship between gene expression and MLG accumulation in wheat and its related species at various developmental stages and among different genotypes. While the expression of *CslF6* has been previously associated with MLG content in grains, our findings suggested a relationship between *CslF6* gene expression and MLG accumulation across different lines. Although the correlation was not strongly pronounced in amphyploid lines, the noteworthy increase in MLG content by 24.4 % in durum wheat and 43.3 % in *T. aestivum* through hybrid breeding presents promising opportunities for enhancing the nutritional profile of these staple grains. Given the global prominence of wheat in human diets, even slight increases in MLG content can lead to significant advancements in dietary quality. Our research paints an optimistic picture for the future of high MLG wheat breeding by introducing genes from wheat wild relatives. Moreover, the study offers insights into the composition and structural features of MLG and other non-starch cell wall polysaccharides in the kernels of wheats, their wild relatives, and hybrid lines between them, providing valuable data for upcoming research in this domain. The study also demonstrated that HPAEC-PAD and GC–MS-based glycomics can be potent tools for plant breeders to enhance the efficiency and accuracy of selecting high MLG wheat relatives and hybrid lines.

CRediT authorship contribution statement

Ilaria Marcotuli: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Xiaohui Xing:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation. **Davide Caranfa:** Formal analysis. **Stefania L. Giove:** Formal analysis. **Yves S.Y. Hsieh:** Writing – review & editing, Data curation. **Shu-Chieh Chang:** Formal analysis. **D. Wade Abbott:** Writing – review & editing, Writing – original draft, Data curation. **Agata Gadaleta:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Agata Gadaleta reports article publishing charges and equipment, drugs, or supplies were provided by university of Bari. Agata Gadaleta reports a relationship with University of Bari that includes: funding grants. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Andersen, C. L., Jensen, J. L., & Ørntoft, T. F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research, 64*(15), 5245–5250. https://doi.org/10.1158/0008-5472.Can-04-0496

Bacic, A., Moody, S. F., & Clarke, A. E. (1986). Structural analysis of secreted root slime from maize (*Zea mays* L.). *Plant Physiology, 80*(3), 771–777. https://doi.org/10.1104/pp.80.3.771

Badhan, A., Low, K. E., Jones, D. R., Xing, X., Milani, M. R. M., Polo, R. O., & McAllister, T. A. (2022). Mechanistic insights into the digestion of complex dietary fibre by the rumen microbiota using combinatorial high-resolution glycomics and transcriptomic analyses. *Computational and Structural Biotechnology Journal, 20*, 148–164. https://www.sciencedirect.com/science/article/pii/S2001037021005171.

Bain, M., van de Meene, A., Costa, R., & Doblin, M. S. Characterisation of Cellulose Synthase Like F6 (CslF6) Mutants Shows Altered Carbon Metabolism in β-D-(1,3;1,4)-Glucan Deficient Grain in Brachypodium distachyon. (1664-462X (Print)).

Bain, M., van de Meene, A., Costa, R., & Doblin, M. S. (2020). Characterisation of cellulose synthase like F6 (CslF6) mutants shows altered carbon metabolism in beta-D-(1,3;1,4)-glucan deficient grain in brachypodium distachyon. *Front Plant Sci, 11*, Article 602850. https://doi.org/10.3389/fpls.2020.602850

Beres, B. L., Rahmani, E., Clarke, J. M., Grassini, P., Pozniak, C. J., Geddes, C. M., & Ransom, J. K. (2020). A systematic review of durum wheat: Enhancing production systems by exploring genotype, environment, and management (G × E × M) synergies. *Frontiers in Plant Science, 11*. https://doi.org/10.3389/fpls.2020.568657

Beresford, G., & Stone, B. A. (1983). (1→3), (1→4)-β-D-glucan content of Triticum grains. *Journal of Cereal Science, 1*(2), 111–114. https://doi.org/10.1016/S0733-5210(83)80028-9

Bertaud, F., Sundberg, A., & Holmbom, B. (2002). Evaluation of acid methanolysis for analysis of wood hemicelluloses and pectins. *Carbohydrate Polymers, 48*(3), 319–324. https://www.sciencedirect.com/science/article/pii/S0144861701002491.

Bhoite, R., Satyavrat, V., & Premasudha Sadananda, M. Clinical benefits of β-glucan supplementation in children: a review. LID - 37. (2731-4286 (Electronic)).

Biswal, A. K., Hengge, N. N., Black, I. M., Atmodjo, M. A., Mohanty, S. S., Ryno, D., & Mohnen, D. (2022). Composition and yield of non-cellulosic and cellulosic sugars in soluble and particulate fractions during consolidated bioprocessing of poplar biomass by *Clostridium thermocellum*. *Biotechnology for Biofuels and Bioproducts, 15*(1), 23. https://doi.org/10.1186/s13068-022-02119-9

Black, I., Heiss, C., & Azadi, P. (2019). Comprehensive monosaccharide composition analysis of insoluble polysaccharides by permethylation to produce methyl alditol derivatives for gas chromatography/mass spectrometry. *Analytical Chemistry, 91*(21), 13787–13793. https://doi.org/10.1021/acs.analchem.9b03239

Burton, R. A., Collins, H. M., Kibble, N. A. J., Smith, J. A., Shirley, N. J., Jobling, S. A., & Fincher, G. B. (2011). Over-expression of specific HvCslF cellulose synthase-like genes in transgenic barley increases the levels of cell wall (1,3;1,4)-β-d-glucans and alters their fine structure. *Plant Biotechnology Journal, 9*(2), 117–135. https://doi.org/10.1111/j.1467-7652.2010.00532.x

Burton, R. A., Jobling, S. A., Harvey, A. J., Shirley, N. J., Mather, D. E., Bacic, A., & Fincher, G. B. (2008). The genetics and transcriptional profiles of the cellulose synthase-like *HvCslF* gene family in barley. *Plant Physiology, 146*(4), 1821–1833. https://doi.org/10.1104/pp.107.114694

Burton, R. A., Wilson, S. M., Hrmova, M., Harvey, A. J., Shirley, N. J., Medhurst, A., & Fincher, G. B. (2006). Cellulose synthase-like *CslF* genes mediate the synthesis of cell wall (1,3;1,4)-β-D-glucans. *Science, 311*(5769), 1940–1942. https://doi.org/10.1126/science.1122975

Carpita, N. C., & Shea, E. M. (1989). Linkage structure of carbohydrates by gas chromatography-mass spectrometry (GC-MS) of partially methylated alditol acetates. In B. C.J. & G. D. McGinnis (Eds.), *Analysis of Carbohydrates by GLC and MS* (pp. 157-216). Boca Raton, Florida: CRC Press, Inc.

Casillo, A., D'Angelo, C., Parrilli, E., Tutino, M. L., & Corsaro, M. M. (2022). Membrane and extracellular matrix glycopolymers of *Colwellia psychrerythraea* 34H: Structural changes at different growth temperatures. *Frontiers in Microbiology, 13*. https://doi.org/10.3389/fmicb.2022.820714

Chang, S.-C., Kao, M.-R., Saldivar, R. K., Díaz-Moreno, S. M., Xing, X., Furlanetto, V., & Hsieh, Y. S. Y. (2023). The Gram-positive bacterium *Romboutsia ilealis* harbors a polysaccharide synthase that can produce (1,3;1,4)-β-D-glucans. *Nature Communications, 14*(1), 4526. https://doi.org/10.1038/s41467-023-40214-z

Chong, H. H., Cleary, M. T., Dokoozlian, N., Ford, C. M., & Fincher, G. B. (2019). Soluble cell wall carbohydrates and their relationship with sensory attributes in *Cabernet Sauvignon* wine. *Food Chemistry, 298*, Article 124745. https://www.sciencedirect.com/science/article/pii/S0308814619308155.

Ciucanu, I. (2006). Per-*O*-methylation reaction for structural analysis of carbohydrates by mass spectrometry. *Analytica Chimica Acta, 576*(2), 147–155. http://www.sciencedirect.com/science/article/pii/S0003267006012372.

Collins, H. M., Burton, R. A., Topping, D. L., Liao, M.-L., Bacic, A., & Fincher, G. B. (2010). Variability in fine structures of noncellulosic cell wall polysaccharides from cereal grains: Potential importance in human health and nutrition. *Cereal Chemistry., 87*(4), 272–282. https://doi.org/10.1094/cchem-87-4-0272

Cory, A. T., Båga, M., Anyia, A., Rossnagel, B. G., & Chibbar, R. N. (2012). Genetic markers for *CslF6* gene associated with (1,3;1,4)-β-glucan concentration in barley grain. *Journal of Cereal Science, 56*(2), 332–339. https://www.sciencedirect.com/science/article/pii/S0733521012000446.

De Pace, C., Bizzarri, M., Vittori, D., Vaccino, P., Caceres, M. E., Ceccarelli, M., . . . Vida, G. (2014). The progeny from the [(*T. turgidum* X *Dasypyrum villosum*) amphiploid X *Triticum aestivum*] hybridization is an effective source of new durum wheat inbred lines. In E. Porceddu, C. O. Qualset & A. B. Damania (Eds.), *Proceedings of the International Symposium on Genetics and breeding of durum wheat* (pp. 189-200): Bari : CIHEAM.

Doblin, M. S., Pettolino, F. A., Wilson, S. M., Campbell, R., Burton, R. A., Fincher, G. B., & Bacic, A. (2009). A barley cellulose *synthase-like CSLH* gene mediates (1,3;1,4)-β-D-glucan synthesis in transgenic *Arabidopsis*. *Proceedings of the National Academy of Sciences, 106*(14), 5996–6001. https://doi.org/10.1073/pnas.0902019106

Farrokhi, N., Burton, R. A., Brownfield, L., Hrmova, M., Wilson, S. M., Bacic, A., & Fincher, G. B. (2006). Plant cell wall biosynthesis: Genetic, biochemical and functional genomics approaches to the identification of key genes. *Plant Biotechnology Journal, 4*(2), 145–167. https://doi.org/10.1111/j.1467-7652.2005.00169.x

Feuillet, C., Langridge, P., & Waugh, R. (2008). Cereal breeding takes a walk on the wild side. *Trends in Genetics, 24*(1), 24–32. https://www.sciencedirect.com/science/article/pii/S0168952507003484.

Fincher, G. B., & Stone, B. A. (2004). Chemistry of nonstarch polysaccharides. In C. Wrigley (Ed.), *Encyclopedia of Grain Science* (pp. 206–223). Oxford: Elsevier.
Foster, C. E., Martin, T. M., & Pauly, M. (2010). Comprehensive compositional analysis of plant cell walls (lignocellulosic biomass) part II: carbohydrates. *Journal of Visual Experiments, 37*, Article e1837. https://doi.org/10.3791/1837
Francin-Allami, M., Bouder, A., Geairon, A., Alvarado, C., Le-Bot, L., Daniel, S., & Sibout, R. (2023). Mixed-linkage glucan is the main carbohydrate source and starch is an alternative source during *Brachypodium* grain germination. *International Journal of Molecular Sciences, 24*(7). https://doi.org/10.3390/ijms24076821
Gottstein, V., Bernhardt, M., Dilger, E., Keller, J., Breitling-Utzmann, C. M., Schwarz, S., & Bunzel, M. (2021). Coffee silver skin: Chemical characterization with special consideration of dietary fiber and heat-induced contaminants. *Foods, 10*(8), 1705. https://www.mdpi.com/2304-8158/10/8/1705.
Grant, C. A., Cubadda, F., Carcea, M., Pogna, N. E., & Gazza, L. (2012). Vitamins, minerals, and nutritional value of durum wheat. In M. Sissons, J. Abécassis, B. Marchylo, & M. Carcea (Eds.), *Durum Wheat: Chemistry and Technology* (2nd edition:). AACC.
Grewal, S., Othmeni, M., Walker, J., Hubbart-Edwards, S., Yang, C.-Y., Scholefield, D., & King, J. (2020). Development of wheat-*Aegilops caudata* introgression lines and their characterization using genome-specific KASP markers. *Frontiers in Plant Science, 11*. https://doi.org/10.3389/fpls.2020.00606
Hosain, N. A., Ghosh, R., Bryant, D. L., Arivett, B. A., Farone, A. L., & Kline, P. C. (2019). Isolation, structure elucidation, and immunostimulatory activity of polysaccharide fractions from *Boswellia carterii* frankincense resin. *International Journal of Biological Macromolecules, 133*, 76–85. https://www.sciencedirect.com/science/article/pii/S0141813019301837.
Houston, K., Russell, J., Schreiber, M., Halpin, C., Oakey, H., Washington, J. M., Booth, A., Shirley, N., Burton, R. A., Fincher, G. B., & Waugh, R. (2014). A genome wide association scan for (1,3;1,4)-β- glucan content in the grain of contemporary 2-row Spring and Winter barleys. *BMC Genomics, 15*(1), 907. https://doi.org/10.1186/1471-2164-15-907
Hu, G., Burton, C., Hong, Z., & Jackson, E. (2014). A mutation of the cellulose-synthase-like (CslF6) gene in barley (Hordeum vulgare L.) partially affects the β-glucan content in grains. *Journal of Cereal Science, 59*(2), 189–195. https://doi.org/10.1016/j.jcs.2013.12.009
Islamovic, E., Obert, D. E., Oliver, R. E., Marshall, J. M., Miclaus, K. J., Hang, A., & Jackson, E. W. (2013). A new genetic linkage map of barley (*Hordeum vulgare* L.) facilitates genetic dissection of height and spike length and angle. *Field Crops Research, 154*, 91–99. https://www.sciencedirect.com/science/article/pii/S0378429013002104.
Ivanizs, L., Marcotuli, I., Rakszegi, M., Kalapos, B., Szőke-Pázsi, K., Farkas, A., & Molnár, I. (2022). Identification of New QTLs for Dietary Fiber Content in Aegilops biuncialis. *International Journal of Molecular Sciences, 23*(7). https://doi.org/10.3390/ijms23073821
Jiang, G., & Vasanthan, T. (2000). MALDI-MS and HPLC quantification of oligosaccharides of lichenase-hydrolyzed water-soluble β-glucan from ten barley varieties. *Journal of Agricultural and Food Chemistry, 48*(8), 3305–3310. https://doi.org/10.1021/jf0001278
Jones, D. R., Xing, X., Tingley, J. P., Klassen, L., King, M. L., Alexander, T. W., & Abbott, D. W. (2020). Analysis of active site architecture and reaction product linkage chemistry reveals a conserved cleavage substrate for an endo-alpha-mannanase within diverse yeast mannans. *Journal of Molecular Biology, 432*(4), 1083–1097. http://www.sciencedirect.com/science/article/pii/S002228362030034 6.
Kim, J. B., & Carpita, N. C. (1992). Changes in esterification of the uronic acid groups of cell wall polysaccharides during elongation of maize coleoptiles. *Plant Physiology, 98*(2), 646–653. https://doi.org/10.1104/pp.98.2.646
Kishii, M. (2019). An update of recent use of *Aegilops* species in wheat breeding. *Frontiers in Plant Science, 10*. https://doi.org/10.3389/fpls.2019.00585
Klassen, L., Reintjes, G., Li, M., Jin, L., Amundsen, C., Xing, X., & Abbott, D. W. (2023). Fluorescence activated cell sorting and fermentation analysis to study rumen microbiome responses to administered live microbials and yeast cell wall derived prebiotics. *Frontiers in Microbiology, 13*. https://doi.org/10.3389/fmicb.2022.1020250
Klassen, L., Xing, X., Tingley, J. P., Low, K. E., King, M. L., Reintjes, G., & Abbott, D. W. (2021). Approaches to investigate selective dietary polysaccharide utilization by human gut microbiota at a functional level. *Frontiers in Microbiology, 12*(308). https://doi.org/10.3389/fmicb.2021.632684
Lante, A., Canazza, E., & Tessari, P. (2023). Beta-glucans of cereals: Functional and technological properties. *Nutrients, 15*(9), 2124. https://www.mdpi.com/2072-6643/15/9/2124.
Li, J., Hsiung, S.-Y., Kao, M.-R., Xing, X., Chang, S.-C., Wang, D., & Hsieh, Y. S. Y. (2022). Structural compositions and biological activities of cell wall polysaccharides in the rhizome, stem, and leaf of Polygonatum odoratum (Mill.) Druce. *Carbohydrate Research, 521*, Article 108662. https://www.sciencedirect.com/science/article/pii/S000862152200163X.
Li, J., Wang, D., Xing, X., Cheng, T.-J.-R., Liang, P.-H., Bulone, V., & Hsieh, Y. S. Y. (2019). Structural analysis and biological activity of cell wall polysaccharides extracted from Panax ginseng marc. *International Journal of Biological Macromolecules, 135*, 29–37. http://www.sciencedirect.com/science/article/pii/S014181301931150X.
Little, A., Lahnstein, J., Jeffery, D. W., Khor, S. F., Schwerdt, J. G., Shirley, N. J., & Bulone, V. (2019). A novel (1,4)-β-linked glucoxylan is synthesized by members of the *cellulose synthase-like F* gene family in land plants. *ACS Central Science, 5*(1), 73–84. https://doi.org/10.1021/acscentsci.8b00568
Low, K. E., Xing, X., Moote, P. E., Inglis, G. D., Venketachalam, S., Hahn, M. G., & Abbott, D. W. (2020). Combinatorial glycomic analyses to direct CAZyme discovery for the tailored degradation of canola meal non-starch dietary polysaccharides. *Microorganisms, 8*(12), 1888. https://www.mdpi.com/2076-2607/8/12/1888.
Maheshwari, G., Sowrirajan, S., & Joseph, B. (2017). Extraction and isolation of β-glucan from grain sources—A review. *Journal of Food Science, 82*(7), 1535–1545. https://doi.org/10.1111/1750-3841.13765
Marcotuli, I., Houston, K., Schwerdt, J. G., Waugh, R., Fincher, G. B., Burton, R. A., Blanco, A., & Gadaleta, A. (2016). Genetic diversity and genome wide association study of β-glucan content in tetraploid wheat grains. *PLoS ONE, 11*(4), e0152590.
Marcotuli, I., Colasuonno, P., Blanco, A., & Gadaleta, A. (2018). Expression analysis of cellulose synthase-like genes in durum wheat. *Scientific Reports, 8*(1), 15675. https://doi.org/10.1038/s41598-018-34013-6
Marcotuli, I., Colasuonno, P., Hsieh, Y. S. Y., Fincher, G. B., & Gadaleta, A. (2020). Non-starch polysaccharides in durum wheat: A review. *International Journal of Molecular Sciences, 21*(8), 2933. https://www.mdpi.com/1422-0067/21/8/2933.
Marcotuli, I., Houston, K., Waugh, R., Fincher, G. B., Burton, R. A., Blanco, A., & Gadaleta, A. (2015). Genome wide association mapping for arabinoxylan content in a collection of tetraploid wheats. *PLoS ONE, 10*(7), e0132787.
Marcotuli, I., Hsieh, Y. S. Y., Lahnstein, J., Yap, K., Burton, R. A., Blanco, A., & Gadaleta, A. (2016). Structural variation and vontent of arabinoxylans in endosperm and bran of durum wheat (*Triticum turgidum* L.). *Journal of Agricultural and Food Chemistry, 64*(14), 2883–2892. https://doi.org/10.1021/acs.jafc.6b00103
Marcotuli, I., Soriano, J. M., & Gadaleta, A. (2022). A consensus map for quality traits in durum wheat based on genome-wide association studies and detection of ortho-meta QTL across cereal species. *Frontiers in Genetics, 13*. https://doi.org/10.3389/fgene.2022.982418
Mathews, R., Kamil, A., & Chu, Y. (2020). Global review of heart health claims for oat beta-glucan products. *Nutrition Reviews, 78*(Supplement_1), 78–97. https://doi.org/10.1093/nutrit/nuz069
McCleary, B. V., & Codd, R. (1991). 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, 55*(2), 303–312. https://doi.org/10.1002/jsfa.2740550215
Miller, S. S., Vincent, D. J., Weisz, J., & Fulcher, R. G. (1993). Oat β-glucans: An evaluation of eastern Canadian cultivars and unregistered lines. *Canadian Journal of Plant Science, 73*, 429–436.
Muhidinov, Z. K., Bobokalonov, J. T., Ismoilov, I. B., Strahan, G. D., Chau, H. K., Hotchkiss, A. T., & Liu, L. (2020). Characterization of two types of polysaccharides from *Eremurus hissaricus* roots growing in Tajikistan. *Food Hydrocolloids, 105*, Article 105768. https://www.sciencedirect.com/science/article/pii/S0268005X19316777.
Nemeth, C., Freeman, J., Jones, H. D., Sparks, C., Pellny, T. K., Wilkinson, M. D., & Shewry, P. R. (2010). Down-regulation of the *CSLF6* gene results in decreased (1,3;1,4)-β-D-glucan in endosperm of wheat. *Plant Physiology, 152*(3), 1209–1218. https://doi.org/10.1104/pp.109.151712
Palmer, R., Cornuault, V., Marcus, S. E., Knox, J. P., Shewry, P. R., & Tosi, P. (2015). Comparative in situ analyses of cell wall matrix polysaccharide dynamics in developing rice and wheat grain. *Planta, 241*(3), 669–685. https://doi.org/10.1007/s00425-014-2201-4
Pettolino, F. A., Walsh, C., Fincher, G. B., & Bacic, A. (2012). Determining the polysaccharide composition of plant cell walls. *Nature Protocols, 7*(9), 1590–1607. https://doi.org/10.1038/nprot.2012.081
Pham, T. A. T., Kyriacou, B. A., Schwerdt, J. G., Shirley, N. J., Xing, X., Bulone, V., & Little, A. (2019). Composition and biosynthetic machinery of the *Blumeria graminis* f. sp. hordei conidia cell wall. *The Cell Surface, 5*, Article 100029. http://www.sciencedirect.com/science/article/pii/S2468233019300076.
Pham, T. A. T., Schwerdt, J. G., Shirley, N. J., Xing, X., Hsieh, Y. S. Y., Srivastava, V., & Little, A. (2019). Analysis of cell wall synthesis and metabolism during early germination of *Blumeria graminis* f. sp. hordei conidial cells induced *in vitro*. *The Cell Surface, 5*, Article 100030. http://www.sciencedirect.com/science/article/pii/S2468233019300088.
Purushotham, P., Ho, R., Yu, L., Fincher, G. B., Bulone, V., & Zimmer, J. (2022). Mechanism of mixed-linkage glucan biosynthesis by barley *cellulose synthase–like CslF6* (1,3;1,4)-β-glucan synthase. *Science Advances, 8*(45), Article eadd1596. https://doi.org/10.1126/sciadv.add1596
Rakszegi, M., Molnár, I., Lovegrove, A., Darkó, É., Farkas, A., Láng, L., & Shewry, P. (2017). Addition of *Aegilops* U and M chromosomes affects protein and dietary fiber content of wholemeal wheat flour. *Frontiers in Plant Science, 8*. https://doi.org/10.3389/fpls.2017.01529
Robb, C. S., Hobbs, J. K., Pluvinage, B., Reintjes, G., Klassen, L., Monteith, S., & Boraston, A. B. (2022). Metabolism of a hybrid algal galactan by members of the human gut microbiome. *Nature Chemical Biology, 18*(5), 501–510. https://doi.org/10.1038/s41589-022-00983-y
Schäfer, J., Hale, J., Hoffmann, C. M., & Bunzel, M. (2020). Mechanical properties and compositional characteristics of beet (*Beta vulgaris* L.) varieties and their response to nitrogen application. *European Food Research and Technology, 246*(10), 2135–2146. https://doi.org/10.1007/s00217-020-03562-4
Schneider, A., Molnár, I., & Molnár-Láng, M. (2008). Utilisation of Aegilops (goatgrass) species to widen the genetic diversity of cultivated wheat. *Euphytica, 163*(1), 1–19. https://doi.org/10.1007/s10681-007-9624-y
Smith, P. J., O'Neill, M. A., Backe, J., York, W. S., Peña, M. J., & Urbanowicz, B. R. (2020). Analytical techniques for determining the role of domain of unknown function 579 proteins in the synthesis of *O*-methylated plant polysaccharides. *SLAS Technology: Translating Life Sciences Innovation, 25*(4), 345–355. https://doi.org/10.1177/2472630320912692

Sopiwnyk, E. (2018). Durum production and consumption, a global perspective. In *Sustainable Production of Durum Wheat in Canada* (pp. 5–9). Bannockburn, IL: Barilla America Inc.

Sowinski, E. E., Gilbert, S., Lam, E., & Carpita, N. C. (2019). Linkage structure of cell-wall polysaccharides from three duckweed species. *Carbohydrate Polymers, 223*, Article 115119. http://www.sciencedirect.com/science/article/pii/S0144861719307866.

Storsley, J. M., Izydorczyk, M. S., You, S., Biliaderis, C. G., & Rossnagel, B. (2003). Structure and physicochemical properties of β-glucans and arabinoxylans isolated from hull-less barley. *Food Hydrocolloids, 17*(6), 831–844. https://www.sciencedirect.com/science/article/pii/S0268005X03001048.

Sun, L., Hu, M., Zhao, J., Lv, L., Zhang, Y., Liu, Q., & Zhang, Y. (2021). Molecular characteristics, synthase, and food application of cereal β-glucan. *Journal of Food Quality, 2021*, 6682014. https://doi.org/10.1155/2021/6682014

Sushytskyi, L., Synytsya, A., Čopíková, J., Lukáč, P., Rajsiglová, L., Tenti, P., & Vannucci, L. E. (2023). Perspectives in the application of high, medium, and low molecular weight oat β-D-glucans in dietary nutrition and food technology-A short overview. *Foods, 12*(6), 1121. https://www.mdpi.com/2304-8158/12/6/1121.

Taketa, S., Yuo, T., Tonooka, T., Tsumuraya, Y., Inagaki, Y., Haruyama, N., & Jobling, S. A. (2012). 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, 63*(1), 381–392. https://doi.org/10.1093/jxb/err285

Tiefenbacher, K. F. (2017). Technology of main Ingredients-water and flours. In K. F. Tiefenbacher (Ed.), *Wafer and Waffle* (pp. 15–121). Academic Press.

Tingley, J. P., Low, K. E., Xing, X., & Abbott, D. W. (2021). Combined whole cell wall analysis and streamlined in silico carbohydrate-active enzyme discovery to improve biocatalytic conversion of agricultural crop residues. *Biotechnology for Biofuels, 14*(1), 16. https://doi.org/10.1186/s13068-020-01869-8

Tonooka, T., Aoki, E., Yoshioka, T., & Taketa, S. (2009). A novel mutant gene for (1–3, 1–4)-β-D-glucanless grain on barley (*Hordeum vulgare* L.) chromosome 7H. *Breeding Science, 59*(1), 47–54. https://doi.org/10.1270/jsbbs.59.47

Vega-Sanchez, M. E., Verhertbruggen, Y., Christensen, U., Chen, X., Sharma, V., Varanasi, P., & Ronald, P. C. (2012). 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, 159*(1), 56–69. https://doi.org/10.1104/pp.112.195495

Voiges, K., Adden, R., Rinken, M., & Mischnick, P. (2012). Critical re-investigation of the alditol acetate method for analysis of substituent distribution in methyl cellulose. *Cellulose, 19*(3), 993–1004. https://doi.org/10.1007/s10570-012-9663-y

Willför, S., Pranovich, A., Tamminen, T., Puls, J., Laine, C., Suurnäkki, A., & Holmbom, B. (2009). Carbohydrate analysis of plant materials with uronic acid-containing polysaccharides–A comparison between different hydrolysis and subsequent chromatographic analytical techniques. *Industrial Crops and Products, 29*(2), 571–580. https://www.sciencedirect.com/science/article/pii/S0926669008002082.

Wilson, L. A., Deligey, F., Wang, T., & Cosgrove, D. J. (2021). Saccharide analysis of onion outer epidermal walls. *Biotechnology for Biofuels, 14*(1), 66. https://doi.org/10.1186/s13068-021-01923-z

Wong, S. C., Shirley, N. J., Little, A., Khoo, K. H. P., Schwerdt, J., Fincher, G. B., & Mather, D. E. (2015). Differential expression of the *HvCslF6* gene late in grain development may explain quantitative differences in (1,3;1,4)-β-glucan concentration in barley. *Molecular Breeding, 35*(1), 20. https://doi.org/10.1007/s11032-015-0208-6

Wood, J. A., Tan, H. T., Collins, H. M., Yap, K., Khor, S. F., Lim, W. L., & Tucker, M. R. (2018). Genetic and environmental factors contribute to variation in cell wall composition in mature desi chickpea (*Cicer arietinum* L.) cotyledons. *Plant, Cell & Environment, 41*(9), 2195–2208. https://doi.org/10.1111/pce.13196

Yu, L., Yakubov, G. E., Zeng, W., Xing, X., Stenson, J., Bulone, V., & Stokes, J. R. (2017). Multi-layer mucilage of *Plantago ovata* seeds: Rheological differences arise from variations in arabinoxylan side chains. *Carbohydrate Polymers, 165*, 132–141. http://www.sciencedirect.com/science/article/pii/S0144861717301583.

CHAPTER III

Genome-Wide Association Study for Protein and Color Content in a Tetraploid Wheat Collection

Background: Grain protein content (GPC) and yellow index (YI) are critical quality traits for assessing the quality of tetraploid wheat. The evaluation of these traits is essential for the development of wheat varieties with improved quality. Recent genetic advancements have opened new possibilities, with the identification of genes that regulate nitrogen uptake and amino acid metabolism, which are critical for nitrogen transfer to the grain. Genes and enzymes such as *nitrate reductase* and *glutamine synthetase* have been linked to QTLs for GPC, while other regulatory genes contribute to nitrogen remobilization, enhancing the nutritional efficiency in plants. The carotenoid biosynthesis in wheat begins with *phytoene synthase* (PSY) and involves enzymes such as *phytoene desaturase* (PDS) and *zeta-carotene desaturase* (ZDS), leading to the production of lycopene, which can be converted into α-carotene, β-carotene, and xanthophylls. These carotenoids are crucial for plant growth, stress responses, and the coloration of wheat-derived products, while also serving as precursors of vitamin A and antioxidants.

AIM: The primary objective of this study was to characterize GPC and YI across the 144 tetraploid wheat accessions, aiming to identify novel superior alleles that could be useful for wheat improvement. A total of 3,924 SNP markers suitable for a genome-wide association study (GWAS) were employed.

Results: The GWAS, conducted using the Mixed Linear Model (Q+K), identified 22 significant marker-trait associations for both GPC and YI, corresponding to 10 and 12 genomic regions associated with these traits, respectively. QTL for GPC were detected on chromosomes 2A, 3A, 5A, and 7A, while QTL for YI were localized on chromosomes 1B, 2B, 3A, 4B, 6A, 6B, and 7A, reflecting the complex genetic architecture of both traits. The phenotypic variation (R^2) explained ranged from 8.1% to 8.6% for GPC and from 8.2% to 11.6% for YI. Additionally, two candidate genes were identified: *glutamine synthetase 2* (*GS2*) on chromosome 2A, involved in nitrogen metabolism, and *phytoene synthase 1* (*Psy-A1*) on chromosome 7A, the key gene involved in carotenoid pigment synthesis. The identification of novel haplotypes enhances the understanding of the genetic mechanisms governing protein and color traits in tetraploid wheat and facilitates the future development of superior wheat varieties.

Genome-Wide Association Study for Protein and Color Content in a Tetraploid Wheat Collection

Ilaria Marcotuli[1*], Patricia Cabas-Lühmann[2], Davide Caranfa[1], Antonia Mores[1], Stefania Lucia Giove[1], Pasqualina Colasuonno[1], Sara Muciaccia[1], Maria Simone[1], Andrés R. Schwember[2], Agata Gadaleta[1]

[1]Department of Soil, Plant and Food Sciences, University of Bari Aldo Moro, Bari, Italy
[2]Departamento de Ciencias Vegetales, Facultad de Agronomía y Sistemas Naturales, Pontificia Universidad Católica de Chile, Santiago, Chile

*** Correspondence:**
Ilaria Marcotuli
ilaria.marcotuli@uniba.it

Keywords: tetraploid wheats, grain protein content, yellow index, QTL, candidate genes

Abstract

Grain protein content (GPC) and yellow index (YI) are the most important quality traits to measure the quality of tetraploid wheats that benefit the pasta consumers. The main objective of this study was to characterize GPC and YI using 144 different accessions of tetraploid wheats (*Triticum turgidum* L., 2n = 4x = 28; AABB genome) constituted by several subspecies and evaluated in southern Italy for 2 years (2019 and 2022) and in central Chile during 2021 in order to identify new useful superior alleles. SNPs marker, suitable for genome-wide association study (GWAS) were 3,924. GWAS was carried out using the Mixed Linear Model (Q+K) which identified 22 marker-trait associations for both traits analyzed, representing 10 and 12 regions associated with grain protein content and yellow index, respectively. QTL linked to GPC were detected on chromosomes 2A, 3A, 5A, and 7A, while QTL for YI were localized on chromosomes 1B, 2B, 3A, 4B, 6A, 6B, 7A, reflecting their complex genetic nature. The phenotypic variation (R^2) explained, ranking from 8.1 to 8.6% for GPC and from 8.2 to 11.6% for YI content. In addition, one gene involved in nitrogen metabolism, the *glutamine synthetase 2* (*GS2*) on chromosome 2A, and the main gene involved in the synthesis of carotenoid pigments, the *phytoene synthase 1* (*Psy-A1*) on chromosome 7A were identified as key candidate genes regulating these two quality traits. Finally, the identification of new haplotypes improves our understanding of the genetic mechanisms controlling protein and color traits in tetraploid wheat and facilitate the development of superior wheat varieties in the future.

1 Introduction

Wheat is a staple crop of global significance, providing essential nutrients and energy to a substantial portion of the world's population. Among wheat species, tetraploid wheat (*Triticum turgidum*) hold a crucial place due to their large genetic diversity and agronomic importance, as well as by the nutritional properties of their kernels. Grain protein content (GPC) is a central trait in both bread (*Triticum aestivum* L.) and durum wheat (*Triticum turgidum* L. ssp. *durum*), influencing their nutritional value, baking quality, and pasta-making properties. Compared to grain legumes, mature wheat grains have a relatively low storage protein content, generally between 8 and 15 % (Shewry, 2009). Other more ancient tetraploid wheat species, such as emmer (*Triticum turgidum* ssp. *dicoccon* Thell) and wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*), exhibit variations in GPC ranging from 11% to 24% (Biradar et al., 2022; Cabas-Lühmann et al., 2023). Over the 20th century, it has been observed that newer wheat cultivars tend to have lower GPC compared to older ones, primarily due to breeding techniques that have focused on increasing yields, a trait that is negatively correlated with GPC (Subira et al., 2014). As a result, the development of new wheat varieties with higher GPC has become a breeding priority, often involving the consideration of alternative wheat species.

Quantitative trait loci (QTLs) associated with GPC have been successfully introgressed into modern wheat cultivars, with some located on the chromosomes of ancient tetraploid wheat species (Cabas-Lühmann et al., 2023). Among these, the most well-known is *Gpc-B1* on chromosome 6B in wild emmer wheat lines 'FA-15-3' and 'F-28-8-3' from Israel (Joppa and Cantrel, 1990). This QTL was shown to significantly increase GPC compared to the recurrent durum wheat parent 'Langdon' [LDN] (Joppa et al., 1991, 1997; Olmos et al., 2003). *Gpc-B1* encodes for a NAC transcription factor that accelerates leaf senescence and enhances nutrient remobilization from leaves to developing grains, thereby increasing the final GPC (Uauy et al., 2006; Avni et al., 2014). Additionally, other GPC-related QTLs located on chromosomes 1A, 2A, 3A, 4A, 4B, 5A, 6B, and 7B of wild emmer line 'Y12-3' explained between 0.6% and 24.4% of the phenotypic variation in GPC (Fatiukha et al., 2020).

Over the past 30 years, increases in GPC have mainly been achieved by increasing nitrogen (N) fertilization.Recent advances in genetics have opened new possibilities for improving GPC. A key strategy to reach this aim is the identification of genes involved in N uptake and those that control enzymes responsible for amino acid metabolism, which are crucial for transferring N to the grain (Wang et al., 2024). Nitrogen metabolism is a complex process that includes nitrate uptake, reduction, assimilation into amino acids, and translocation throughout the plant (Fortunato et al., 2023). Various genes and transcription factors are involved in these processes (Balyan et al., 2016). Nitrate transporters, particularly the NRT2 family, play a central role in nitrate uptake, acting as high-affinity transporters in roots to efficiently use soil nitrate even when it is scarce, as reported in the literature in *Arabidopsis* (Garnett et al., 2009). Nitrate absorbed by the plant is first reduced by cytosolic nitrate reductase (NR) to nitrite, which is then transported into the chloroplast and further reduced by nitrite reductase (NIR) to ammonium. This ammonium is incorporated into organic molecules via the glutamine synthetase (GS) and glutamate synthetase (GOGAT) pathways (Lea and Azevedo 2007; Fortunato et al. 2023). Studies have linked these genes to GPC QTL in durum wheat (Gadaleta et al., 2011, 2014; Nigro et al., 2013, 2014, 2016; Arriagada et al., 2022).

Beyond these key enzymes, others involved in N utilization include glutamate dehydrogenase, alanine aminotransferase, and asparagine synthetase, which are all part of amino acid metabolism. Glutamate dehydrogenase (GDH) facilitates the interconversion between glutamate and 2-oxoglutarate (Habash et al., 2007; Quarrie et al., 2005; Fontaine et al., 2009; Xu et al., 2014), while alanine aminotransferase

(AlaAT) is involved in the synthesis and degradation of alanine (Shrawat et al., 2008; Lu et al., 2016). Asparagine synthetase (ASN) catalyzes the formation of asparagine (Asn) and glutamate from glutamine (Gln) and aspartate (Lu et al., 2016). Several ASN genes have been identified in durum wheat, with ASN1 being upregulated in response to certain mineral deficiencies (Gao et al., 2016). Besides these genes directly involved in nitrogen metabolism, other genes act as regulatory and signaling factors under specific conditions (Xu et al., 2012). For instance, NIN-like protein (NLP7) is a transcription factor crucial for nitrate assimilation and signaling in higher plants (Castaings et al. 2009). Genes related to nitrogen remobilization also play a significant role in determining N content and final yield. Overexpression of the *cytosolic pyruvate orthophosphate dikinase* (*PPDK*) gene has increased seed weight in *Arabidopsis* by accelerating N remobilization from leaves, thereby boosting growth rate, seed weight, and N content (Taylor et al., 2010). These genes and enzyme activities represent potential targets for improving nitrogen use efficiency (NUE) in crop plants that are highly valuable for wheat breeding programs.

A second important quality trait is the carotenoid content, which is the organic pigment found widely in plants, photosynthetic algae, and some fungi and bacteria (Hirschberg et al., 2001). These pigments are commonly associated with the thylakoid membranes of chloroplasts, and they are responsible for the yellow, orange, and red colors observed in many flowers, fruits, and roots (Cazzonelli et al., 2010). In plants, carotenoids play a crucial role in photosynthesis and photo-oxidative protection, and they are also precursors for apocarotenoid hormones like abscisic acid and strigolactones (Seo et al., 2002; Xie et al., 2010). The relationship between carotenoids and human health has been extensively studied, with evidence suggesting that they help to protect against conditions mediated by reactive oxygen species (ROS), including cardiovascular diseases, certain cancers, and various neurological and eye-related disorders (Fiedor et al., 2014).

Carotenoids are categorized into two main types: carotenes, which are hydrocarbons, and xanthophylls, which contain oxygen (Moise et al., 2014). The biosynthesis of carotenoids, primarily studied in *Arabidopsis thaliana*, rice, maize, and some ornamental plants, begins with the enzyme phytoene synthase (PSY) catalyzing the formation of phytoene from geranylgeranyl diphosphate. This phytoene undergoes a series of desaturation reactions involving enzymes like phytoene desaturase (PDS) and zeta-carotene desaturase (ZDS), eventually producing lycopene. Lycopene can then be converted into α-carotene or β-carotene, which are further modified into various xanthophylls. These processes lead to the production of abscisic acid and strigolactones, which are essential for plant growth and stress responses (Moise et al., 2014; Ruiz-Sola et al., 2012).

Carotenoids in wheat not only serve as antioxidants and vitamin A precursors but also influence the color of wheat products (Lap et al., 2021). Consumers generally prefer white bread from common wheat and yellow pasta from durum wheat. The yellow color in durum wheat is primarily attributed to lutein, a type of carotenoid (xanthophyll) that constitutes approximately 80% of the grain's endosperm pigments (Ramachandran et al., 2010). The total yellow pigment content (YPC) is typically quantified using the AACCI Approved Method 14-50.01 (AACC International, 2010). However, this method is time-consuming due to the extended period required for pigment extraction with butyl alcohol (Sandmann, 2001). A faster and more convenient alternative is measuring the yellow index (YI), which is based on the CIE *b** color determination using light reflectance through chromatography. This method provides a quick assessment of yellow color in durum wheat (Sissons et al., 2020). Among the tools available for this purpose, the Minolta colorimeter (Konica Minolta Pty Ltd, Macquarie Park, NSW) is the most commonly used instrument for obtaining these measurements. High YI is a desirable trait for pasta production and an objective in durum breeding programs (Pozniak et al. 2007; Zhang and Dubcovsky 2008) because of the general demand of the

industry for pasta to be bright yellow (Fu et al. 2011). However, the final color can be affected by storage and processing, where carotenoid degradation occurs due to enzymes like polyphenol oxidase and lipoxygenase (Mares et al., 2001).

The color of wheat flour and semolina is a quantitative trait controlled by multiple genes and influenced by environmental factors (Clarke et al., 2008; Schulthess et al. 2013). Studies mapping yellow pigment content (YPC) and yellow index (YI) in wheat have identified QTL on all chromosomes, with a major QTL on chromosome 7A linked to variations in the *phytoene synthase* (*Psy-A1*) gene (He et al., 2008; He et al., 2009; Zhang and Dubcovsky 2008; Patil et al. 2018). Minor QTLs have been reported in chromosomes 2A and 6A in durum wheat. These QTLs explain 2.3 to 15% of the YPC varietions and 5 to 21 % YI varietions across different populations, respectively (see review Colasuonno et al., 2019). Despite advances in understanding carotenoid regulation, only a few carotenoids biosynthetic genes have been identified and cloned in wheat, such as *phytoene synthase* (*PSY*) (He et al., 2008; Pozniak et al., 2007; Dibari et al., 2012), *lycopene ε-cyclase* (*LYCE*) (Howitt et al., 2009), and *carotene desaturase* (*PDS*) (Cong et al., 2010).

Quantitative studies offer a robust approach to dissect the genetic architecture of complex traits by leveraging natural genetic variation within a diverse collection of genotypes (Rufo et al., 2021).

Additionally, the enhancement of quality traits in modern wheat cultivars through the introgression of genes from diverse collections has been significantly advanced by marker-assisted selection (MAS) and other molecular breeding tools (Subedi et al., 2023). These approaches leverage genomic resources to predict quality attributes with high accuracy and efficiency, accelerating crop improvement and cultivar development by enabling the selection of genotypes with superior end-use quality traits (Subedi et al., 2023).

The creation of core marker sets has been instrumental in effective genomic-assisted breeding. For wheat, a set of markers distributed across the genome has been developed, facilitating the detection of genetic variations associated with quality traits (Ishikawa et al., 2022). This resource enhances the precision of selection in breeding programs and facilitate the efficient incorporation of beneficial genes from diverse genetic resources, leading to wheat varieties with enhanced end-use quality (Ishikawa et al., 2022).

In this study, we performed GWAS on a diverse tetraploid wheat collection to identify genomic regions associated with protein content and yellow index. Additionally, we employed Quantitative Trait Loci (QTL) mapping in a collection of tetraploid wheats through GWAS analysis to pinpoint specific loci that significantly contribute to the phenotypic variation observed in these traits. The identification of these loci will enhance our understanding of the genetic mechanisms governing protein and color traits in tetraploid wheat active in different environment condition and facilitate the development of superior wheat varieties through marker-assisted selection.

2 Material and methods

2.1 Plant material and field trial

A total of 144 accessions of tetraploid wheats (*Triticum turgidum* L., 2n = 4x = 28; AABB genome) were grown in southern Italy at the experimental fields at Valenzano (41°01'13.1"N, 16°54'12.9"W) for 2 years (2019 and 2022). The rainfall during the growing season in Italy (2019/2020 – November/June) was 43 mm with a minimum average temperature of 0.7 °C and maximum average temperature of 30 °C and (2021/2022 – November/June) 63 mm with a minimum average temperature

of 1.6 °C and maximum average temperature of 34 °C (data from https://www.visualcrossing.com/weather-data). In central Chile the experimental field station was located in Pirque (33° 40'00''S, 70° 35'00''W - 654 m above the sea level) during 2021. The rainfall during the growing season in Chile (2021- August/December) was 56 mm with average maximum temperature of 22.7 °C during December and a minimum average temperature of 3.6 °C in August (data from the climatological station at Pirque experimental station). The collection included the accessions of several *T. turgidum* subspecies: *dicoccum* (125 accessions), *paleocolchicum* (2), *polonicum* (8), and *turanicum* (9) (Supplementary Table 1). A randomized complete block design with three replications was employed in the three field experiments. Each plot consisted of 1-meter rows spaced 30 cm apart, with 80 germinating seeds per plot. In Italy, fertilizers were applied at the following rates before sowing: 68 kg N/ha and 46 kg P_2O_5/ha. Additionally, 32 kg N/ha was applied during stem elongation. The plots were irrigated only when there was no rain during the developmental stages (emergency irrigation). When the kernels were in waxy stage, the irrigation stop even if it's drought. Plots were hand-harvested at maturity. In Chile, fertilizers were applied at the following rates before sowing: 51 kg N/ha, 69 kg P_2O_5/ha, and 53 kg K_2O/ha. Additionally, 184 kg N/ha was applied during the tillering stage. The plots were irrigated to prevent drought stress, and weeds were controlled using chemical methods. At maturity, plots were hand-harvested and threshed.

2.2 Grain Composition and statistical analysis

Grain samples (40 g) were milled into whole wheat meal using a UDY cyclone sample mill (UDY Corp., Fort Collin). Yellow index was determined by Minolta CR410 colorimeter (Konica Minolta, Ramsey, NJ, USA) configured to measure Commission Internationale d'Eclairage (CIE) L*(brightness), *a** (redness), and *b**(yellowness) color values. The ground samples were placed in a black cell that was 1 cm deep and covered with a quartz glass. The grain protein content was measured using the Grain Analyser Infratec™ NOVA (FOSS).

Each environment (location-year) included 144 genotypes. Genotypes were considered fixed effects, while blocks were treated as random within each environment. All collected data were subjected to analysis of variance (ANOVA) at a 95% confidence level (F tests: $P \leq 0.05$) using the PROC GLM procedure of SAS® (SAS Institute Inc., Cary, NC, USA). The least square means were calculated using the "LSmeans" procedure in SAS and separated by Fisher's-protected LSD at P=0.05. This method was chosen to compare group means, particularly because some of the data were unbalanced. By using this approach, the best-fit means for the statistical model were obtained, allowing for the evaluation of statistical differences between the groups. Broad-sense heritability (H^2) was estimated as the proportion of genetic variance ($\sigma^2 g$) to phenotypic variance using the Rstudio® package variability. Heritability values were categorized as follows: <0.5 indicating low, 0.5 to 0.75 indicating moderate, 0.75 to 0.9 indicating high, and >0.9 indicating very high broad-sense heritability.

2.3 QTL and Candidate Gene Detection

The 7K iSelect array, developed by Illumina CSPro® (San Diego, CA, USA) (Wang et al., 2014), was employed to analyze 6,731 SNP sequences across the collection. Genotyping was conducted on 1 µg of genomic DNA at TraitGenetics GmbH (Gatersleben, Germany), following the manufacturer's protocols as outlined by Akhunov et al. (2009). These genotyping assays were executed using the Illumina iScan reader and processed with GenomeStudio software v2011.1 (Illumina CSPro®).

Prior to conducting the genome-wide association study (GWAS), markers with a minor allele frequency below 10% and those with over 5% missing data were excluded from the dataset using GenAlEx software. This filtering process yielded 3,942 SNPs for the population structure analysis.

To analyze the population structure, three methods were employed: principal coordinate analysis (PCoA) using GenAlEx, Bayesian clustering with STRUCTURE version 2.3.4, and an unrooted Bayesian tree generated with TASSEL 5.2.70. The STRUCTURE analysis utilized an admixture model with correlated allele frequencies, testing subpopulation numbers (K) from 1 to 10, with a burn-in period of 10,000 iterations followed by 10,000 Markov chain Monte Carlo (MCMC) iterations. The GAPIT software packages from R facilitated the identification of significant associations with grain quality traits.

The GWAS was conducted using two methods (MLM and BLINK) (Huang et al., 2019) to evaluate the trait-SNP association for protein a yellow pigment traits using the Genomic Association and Prediction Integrated Tool (GAPIT) (version 3) (Wang et al., 2021), incorporating both the Q matrix and the kinship matrix (K). These models assessed the correlation coefficient (R^2) and marker effect of each SNP in relation to protein content and yellow index. The optimal –Log10(p-value) threshold was determined from the F-test for testing H0: No association between the SNP and trait. The QTL identified with the MLM and BLINK models provided important information for understanding the genetic architecture of protein and color traits.

Using the SNP sequences available on CerealDB website (https://www.cerealsdb.uk.net/cerealgenomics/CerealsDB/indexNEW.php) and SVEVO genome sequence available on GrainGenes database (https://wheat.pw.usda.gov/GG3/), a bioinformatics analysis was conducted to identify genes involved in the biosynthesis of protein and yellow pigments through BLAST analysis. In details, SNPs located on the QTL regions were used as a query for the databanks and gene sequences with the higher homology were picked.

3 Results

3.1 Protein content and yellow index in tetraploid wheats

A wide and significant variability for GPC was observed in the collection, confirming the quantitative nature of the traits (**Figure 1**), the multi-environment ANOVA determined that the genotype x environment interaction for GPC and YI was statistically significant ($P \leq 0.001$) (Supplementary Table 2). **Table 1** summarizes the mean, standard deviation (SD), median, range, and broad-sense heritability (h^2) for GPC and YI per environment (n = 144). Grain protein content ranged from 10.6 % to 27.4 % with an average of 16.0 %. In particular, the environment "Bari19" exhibited the highest GPC, averaging about 2% higher than the other two environments. Bari19 also had a range of about 14% difference between the lowest and highest values, followed by Chile21 with a 12% difference, and Bari22 with a 10% difference. Across all the environments, the genotype with the lowest GPC was PI387744 with 13.8%, while the genotype with the highest GPC was PI352350 with 18.7%; both were *Triticum dicoccum* (Supplementary Table 3).

The YI values ranged from 8.21 to 16.37 with an average of 12.01 (**Figure 1**). In Bari19, YI values ranged from 8.21 to 13.54, with a median of 9.78, followed by Chile21, which ranged from 10.80 to 16.60, and Bari22, which ranged from 10.73 to 16.37. Averaged across Bari19 and Chile21, the

genotype with the lowest YI was PI349050 (*Triticum turgidum* subsp. *paleocolchicum*) with 10.15, while the genotype with the highest YI was PI211691 (*Triticum turgidum* subsp. *turanicum*) with 13.76 (Supplementary Table 3).

Broad-sense heritability (H^2) was estimated as the proportion of variance attributed to the genotype, considering the interaction between genotype and environment. Traits with higher H^2 values are more responsive to genotype and thus highly heritable (>0.75). For GPC, H^2 ranged from low (0.44 in Chile21) to very high (0.90 in Bari22) with an overall of 0.74. For YI, h^2 ranged from low (0.40 in Chile21) to high (0.88 in Bari19) with an overall of 0.65.

3.2 Molecular data and population structure

A total of 3,924 SNPs (58.6% of the original Wheat 7K iSelect array), obtained after removing the markers with a minor allele frequency below 10% and those with over 5% missing data, were informative for a genome-wide association study. The high-resolution consensus linkage map of durum wheat (Maccaferri et al., 2014) was used in the current study as reference map for SNP markers, while the axiom markers, which are not yet mapped into the durum wheat consensus map, but recently added to the SGS SNP array, were located based on the reference Chinese Spring genome assembly (IWGSC RefSeq v2.1).

All the SNPs that failed to amplify, monomorphic, located on chromosome D and/or did not meet the criteria for missing data and allele frequency outlined in the methods section, were removed. The B genome contained the highest number of SNPs 2,028, in contrast to the A genome, which had 1,896 SNPs.

Principal coordinate analysis (PCoA) (**Figure 2A**) was conducted on a dataset consisting of 144 genotypes and 3,942 markers to uncover the genetic structure. The first three coordinates explained 22.86%, 10.55% and 9.83% of the variation, respectively, accounting for a total of 43.24%. An unrooted Bayesian tree, constructed using the Neighbor joining method, grouped accessions into clades based on the genetic distances from the 7K iSelect markers. This phylogenetic approach validated the subgroup divisions observed in the PCoA (**Figure 2B**).

To classify individuals into subpopulations based on genetic similarity, a Bayesian approach was used implemented in STRUCTURE. Following the methodology described by Evanno et al. (2005), ΔK values were plotted against the number of sub-groups (K). Adhering to the procedure outlined in Evanno et al. (2005), ΔK against the K values of the sub-groups were plotted (Supplementary file 1). Based on the STRUCTURE results, the most probable number of subpopulations was 2. Consequently, with K set to 2, the collection was divided into two sub-groups (group 1 and group 2) comprising each 72 accessions, based on SNP data (**Figure 2C**). Specifically, the *dicoccum* genotypes were assigned both to cluster 1 and cluster 2, with a mean Q1 membership of 0.93, anda mean Q2 membership of 0.73, respectively. In detail, the Q1 cluster was composed by the 67 *dicoccum* genotypes, 2 *paleocolchicum*, 2 *turanicum* and 1 *polonicum*, while the Q2 cluster by 58 *dicoccum* genotypes, 7 *turanicum* and 7 *polonicum*.

3.3 QTL detection and validation through marker-trait associations analysis

The GWAS carried out using the MLM (Q+K), which integrated both the PCA and relatedness matrices to minimize the false-positive rate and the BLINK model (Q+K), which employs the Bayesian Information Criterion within a fixed-effect framework to approximate maximum likelihood,

thereby reducing computational complexity. Specifically, were identified 22 marker-trait associations for both the traits analyzed representing 10 and 12 regions associated with grain protein content and yellow index, respectively (**Table 2**). Specifically, ten genomic regions were associated with GPC, located on chromosomes 1B, 2A, 3A (two QTL), 3B, 4B, 5A, 6B and 7A (two QTL), respectively. For YI, 12 regions were identified on chromosomes 1B, 2A, 2B, 3A, 4A, 4B (two QTL), 5A, 6A, 6B, 7A (two QTL). A QTL was considered significant when one or more markers were associated with the traits at –Log10(p-value)>3 (LOD>3) and – log10(p-value)>5 for MLM analysis in Chile 2021 for both the traits protein contend and yellow index (Akond et al. 2019).

QTL linked to grain protein content in grain were found on chromosomes 1B, 2A, 3A (two QTL), 3B, 4B, 5A, 6B and 7A (two QTL) (**Figure 3, Supplementary Figures 2 and 3**), while QTL for yellow index were identified on chromosomes 1B, 2A, 2B, 3A, 4A, 4B (two QTL), 5A, 6A, 6B, 7A (**Figure 4, Supplementary Figures 4 and 5**). Some of the QTL were detected only using the BLINK method while some other only with the MLM method. The QTL identified with both methods were *QYI.bc.1B*, *QYI.bc.2B*, *QYI.bc.4B.1*, *QYI.bc.6A*, *QYI.bc.6B.1*, *QYI.bc.7A.2* (**Table 2**). Additionally, many QTL were identified in more than one environment such as *QGpc.bc.1B*, *QGpc.bc.5A*, *QYI.bc.2A*, *QYI.bc.2B*, *QYI.bc.3A* and *QYI.bc.7A.2*. The phenotypic variation (R^2) explained by each of these markers varied from 8.1 to 12.5% for protein content and from 8.2 to 15.5% for yellow index content. The estimated allelic effects for each marker ranged from −4.39 to 4.26 units for protein content, and from -0.74 to 3.58 for yellow color (**Table 2**).

3.4 Candidate genes search

Using the annotated gene sequences and the SNP sequence available we conducted a bioinformatic analysis in order to find genes involved in protein and pigment synthesis and degradation. The analysis allowed the identification of one of the gene involved in nitrogen metabolism, the *glutamine synthetase 2* (*GS2*) on chromosome 2A, and the main gene involved in the synthesis of carotenoid pigments, the *phytoene synthase 1* (*Psy-A1*) on chromosome 7A (**Table 2**).

4 Discussion

Nitrogen metabolism involves a multifaceted pathway that includes nitrogen (N) uptake, reduction, assimilation, and translocation, posing significant challenges for researchers and breeders due to its complexity and the influence of both genetic and environmental factors on Nitrogen Use Efficiency (NUE) and Grain Protein Content (GPC) (Ortiz-Monasterio et al., 1997; Foulkes et al., 2009). Traditionally, improving GPC has been difficult due to its negative correlation with grain yield (GY) in wheat grown under the same N availability conditions (Bogard et al., 2010). Physiological reasons for this process include the dilution effect of grain nitrogen by carbohydrate compounds and the bioenergetic requirements for synthesizing carbohydrates and proteins (Kastori and Petrovic, 2004).

Several studies have investigated the effects of introgressing the *Gpc-B1* gene on wheat yield and grain protein content (GPC), with mixed outcomes. Vishwakarma et al. (2014) reported that the introgression of *Gpc-B1* improved GPC without a yield penalty in certain wheat lines as well as Tabbita et al. (2017), which summarized the effects of *Gpc-B1* on yield and quality traits across 25 studies, noting that while *Gpc-B1* can enhance GPC, its impact on yield varies depending on genetic background and environmental conditions. Conversely, a negative correlation between grain yield

and GPC was observed in Near-isogenic lines (NILs, indicating that the inclusion of the *Gpc-B1* allele significantly decreased grain yield potential (Kuhn et al. 2016).

However, there have been successful cases of selecting wheat lines that exhibit increases in both GY and GPC, suggesting the possibility of simultaneous improvement when appropriate genetic sources are utilized (Brevis and Dubcovsky, 2010).

The study of genes involved in carotenoid biosynthesis, such as *PSY1*, *PSY2*, and others, has revealed their importance in determining yellow pigment content in wheat, which is crucial for the nutritional quality of wheat products (Campos et al., 2016). Manipulating these genes could potentially enhance the carotenoid content in wheat, improving its nutritional value, particularly in terms of provitamin A content (Richaud et al., 2018).

Here we report the phenotypic variation in GPC and YI contents in a collection of 144 tetraploid wheat lines. In this study, the broad-sense heritability (H2) values from the Italian environments indicate that most of the variability in GPC and YI can be attributed to genetic factors, which is reflected in the high heritability values (Table 1), suggesting that these traits are largely determined by genotype in these settings. In contrast, the medium-low heritability values observed in the Chilean environment for both traits suggest that environmental factors have a more significant impact there. Despite the strong genetic influence on YI, environmental conditions can still affect the final trait expression, as reported by Schulthess et al. (2013).

Grain protein content is a critical trait for product manufacturing in *Triticum durum* and *Triticum aestivum*.

Grain protein content is a critical trait for product manufacturing in *Triticum durum* and *Triticum aestivum*. For high-quality pasta production, a GPC exceeding 13% is desirable, as levels below 11% result in poor-quality pasta regardless of drying conditions during processing (Delcour and Hoseney, 2010). Among the genotypes analyzed, the tetraploid species *Triticum dicoccon* (emmer) showed both the lowest and highest GPC values. Previous studies have reported that emmer typically has a higher GPC than durum wheat, with values reaching up to 22% (Biradar et al., 2022; Cabas-Lühmann et al., 2023) compared to 15% in durum wheat (Shewry, 2009). These findings are consistent with the results of this study, where average GPC values for emmer genotypes ranged from 13.8% to 18.7%.

In species like durum wheat, a high yellow color is desirable because it is crucial to meeting the needs of stakeholders, from consumers to pasta manufacturers, due to its association with high antioxidant capacity (Troccoli et al., 2000; Alzuwaid et al., 2021). In this study, the YI values ranged from low to medium, averaging 10.15 to 13.76 across different environments. Among the species studied, *Triticum turgidum* subsp. *paleocolchicum* had the lowest YI, while *Triticum turgidum* subsp. *turanicum* had the highest. The YI values observed in this study are below those typically reported for durum wheat. For example, in a study by Digesù et al. (2009) comparing the YI of old tetraploid wheat species with durum wheat released before and after 20 years, it was found that newer durum wheat cultivars had a mean YI of 17, compared to 16.1 for older cultivars. The mean YI for subsp. *turanicum* was 15.5 (ranging from 14.1 to 18.7 in 2006), while for emmer, it was 13.3 (ranging from 11.5 to 15.0 in 2007), similar to old durum varieties released before the 1970s. This suggests that selection for yellow color has improved over the last two decades (Ficco et al., 2014).

The dense genotyping platform used to gather genotypic data for our analysis, along with the growing availability of annotated genomes (Wang et al., 2014), enhance the prediction of candidate genes associated with the trait of interest through association analysis. In our study, the initial set of single nucleotide polymorphisms (SNPs) was filtered to ensure data quality and relevance, resulting in a subset of 3,924 SNPs suitable for genome-wide association study (GWAS). This curation achieved

an average marker density of 6 SNPs per centimorgan (cM), with a notable disparity in SNP distribution between the B genome (2,028 SNPs) and the A genome (1,896 SNPs). This imbalance may reflect underlying differences in genetic diversity or historical selection pressures within these genomic regions (Leigh et al., 2022).

The principal coordinate analysis (PCoA) of 144 genotypes using 3,942 markers revealed significant genetic structuring within the population further point to a non-random distribution of genetic diversity, likely influenced by both historical events and selection pressures. Similar data were already found in a tetraploid wheat collection composed by 230 accessions of different *Triticum* subspecies (Marcotuli et al., 2016). The first three principal coordinates accounted for 43.24% of the total genetic variation, indicating a genetic architecture in the population of study. The subsequent construction of an unrooted Bayesian tree using the Neighbor joining method further corroborated the genetic clustering observed in the PCoA, suggesting distinct genetic clades within the dataset.

The principal coordinate analysis (PCoA) and the Bayesian tree (Figures 2A and 2B) revealed significant genetic structuring within the population, which may be attributed to the shared ancestry or geographic and ecological differentiation among the genotypes. The clustering of genotypes into distinct clades suggests the presence of subpopulations with unique evolutionary trajectories (Marcotuli et al., 2016; Liu et al., 2024).

To refine our understanding of the population structure, a Bayesian approach via the STRUCTURE software was employed. The Evanno method indicated that the optimal number of subpopulations (K) was two, as shown in the K plot (Supplementary file 1), and resulting in two distinct sub-groups with equal representation (72 accessions each). The clustering revealed by STRUCTURE analysis reflected the PCoA distribution considering the Coordinate 2 . This clear delineation of subpopulations based on genetic similarity highlights the utility of STRUCTURE analysis in identifying and validating genetic subgroups within a diverse germplasm collection.

The observed genotypic similarities and distinctions in our study highlight the complex interplay of evolutionary history, natural selection, and human-driven selection in shaping genetic diversity within the population. The origin of the different genotypes within the population may stem from a combination of ancestral divergence and adaptive evolution. The clustering observed in both PCoA and the Bayesian tree suggests that these genotypes could have originated from geographically distinct populations or breeding programs that introduced novel alleles justifying the structuring of the collection. The genetic structuring also hints at historical isolation or selective breeding for specific traits, which may have contributed to the divergence among clades.

The GWAS identified 22 significant marker-trait associations for the traits under study: grain protein content (GPC) and yellow index (YI). These findings underscore the polygenic nature of both traits and the involvement of multiple chromosomal regions in their expression (Blanco et al., 2002; Parada et al., 2020).

The identification of QTL associated with GPC is aligned with previous studies that have highlighted the importance of nitrogen uptake and assimilation in determining protein content in wheat grains (Zhao et al., 2023). The presence of QTL on chromosome 1B as well as on chromosomes 2A, 3A has already been reported in the literature (Marcotuli et al., 2022; Nigro et al., 2019), underscoring their potential in breeding programs aimed at improving GPC while maintaining NUE (Habash et al., 2007). Additionally, we confirmed the presence of the *glutamine synthetase* gene (*GS2*) associated

with the QTL on chromosome 2A, as previously reported by other researchers (Nigro et al., 2020; Fortunato et al., 2023). Recent genome-wide association studies (GWAS) have identified novel marker-trait associations (MTAs) linked to grain protein content (GPC) across various wheat chromosomes (Shahi et al., 2024). For instance, a meta-analysis consolidated 459 GPC-related quantitative trait loci (QTLs) from 48 studies, pinpointing 57 meta-QTLs and seven QTL hotspots distributed across all wheat chromosomes except 1D and 4D (Saini et al., 2022). Additionally, a GWAS involving 266 soft red winter wheat genotypes uncovered 80 significant MTAs for ten end-use quality traits, including kernel and flour protein content. Among these, 13 major-effect QTLs explained over 10% of phenotypic variance, with 12 considered novel (Subedi et al., 2024). These findings underscore the complex genetic architecture of GPC and highlight the potential for utilizing diverse genomic regions in breeding programs aimed at enhancing wheat protein content without compromising yield.

The GWAS analysis allowed also the identification of a QTL on chromosome 3B, which was already detected by Nigro et al. 2019. In the same region the *NADH-dependent glutamate synthase* (*NADH-GOGAT*) gene was identified and previously reported in literature (Nigro et al. 2019; Fortunato et al., 2023). Another QTL already present in literature was found on chromosome 4B, with a slight difference in the chromosome position. Our QTL was found at 62.9 cM, while the previous one on 70.4 cM and we could not find the association with the *glutathione reductase 1* (*gsr1*) gene as reported in literature (Nigro et al., 2019). Others QTL were detected on chromosomes 5A, 6B and two on 7A confirming data on literature (Blanco et al. 2012; Cabas-Lühmann et al, 2024).

From literature, YP content is primarily influenced by genes located on homoeologous group 7 chromosomes. Parker et al. (1998), used a collection of 150 SSD lines from Schomburgk/Yarralinka, to identify two significant QTLs on chromosomes 7A and 3A, which accounted for 60% and 13% of the phenotypic variance, respectively. In a Sunco/Tasman mapping population, two QTL were detected on chromosomes 7A and 3B that explained 27% and 20% of the phenotypic variance, respectively (Mares and Campbell 2001), while a major QTL for flour yellowness *b** was mapped on chromosome 7B, accounting for 48% and 61% of the phenotypic variance in different growing seasons (Kuchel et al., 2006). Zhang et al. (2006) identified a major QTL on chromosome 7A that influenced kernel YP content and flour yellowness *b**, explaining between 12.9% and 37.6% of the phenotypic variance across five environments. Additional QTL for YP content were detected on homoeologous group 1 chromosomes (Ma et al., 1999), and on chromosomes 4A and 5A (Hessler et al. 2002), 2D and 4D (Zhang et al., 2006), and 2A, 4B, and 6B (Pozniak et al., 2007). Similarly, in durum wheat, major QTL for YP content were found on chromosomes 7A and 7B (Elouafi et al., 2001; Pozniak et al., 2007). This indicates that YP content in wheat grain is controlled by multiple genes, in addition to the primary genes on homoeologous group 7 chromosomes.

Our study allowed the identification of QTL associated with YPC highlight the genetic control of carotenoid biosynthesis in wheat. The significant QTL identified on chromosomes 1B, 2A, 2B, 3A, 4A, 4B (two QTL), 5A, 6A, 6B, and 7A (two QTL) were already reported in literature and suggested the involvement of biosynthetic genes in determining YPC, providing targets for enhancing the nutritional quality of durum wheat (Colasuonno et al. 2017; He et al., 2008; Pozniak et al., 2007; Dibari et al., 2012, N'Diaye et al., 2018).

The QTL reported on chromosome 2B was already reported by Colasuonno et al. (2017), which was co-located with the *violaxanthin de-epoxidase* gene; parallelly, the QTL on chromosome 4B on 69.3

cM was identified at 83.1 cM and co-segregant with the *phytoene desaturase* gene (Colasuonno et al., 2017). Finally, the QTL QYI.bc.6B.1 was reported already by N'Diaye et al. (2018), confirming the presence of a locus controlling the yellow color of the semolina on chromosome 6B.

In our report, we identified the *phytoene synthase 1 (Psy-A1)* gene associated with the QTL on chromosome 7A as reported in literature repeatedly and responsible for the regulation of the yellow color of semolina and pasta (Elouafi et al., 2001; Pozniak et al., 2007; Patil et al., 2018). Notably, we identified a key gene implicated in the synthesis of carotenoid pigments in durum wheat, which regulate the yellow coloration of semolina and pasta. This gene is crucial targets for breeding programs aimed at enhancing the nutritional and aesthetic qualities of wheat products. The involvement of the *Psy1* gene in carotenoid accumulation is reported also in other species, such as *Zea mays* and *Oryza sativa*, and its localization to chromosome 6 in maize and rice, which share synteny with wheat group 7 chromosomes, underscores its potential role in wheat YP content. There is evidence that *Psy1* is significant associated with endosperm yellowness in maize (Palaisa et al., 2003) and correlated with carotenoid content in maize endosperm (Gallagher et al. 2004), highlighting its importance in the carotenoid biosynthetic pathway. This connection suggests that the genes responsible for kernel YP content on wheat homoeologous group 7 chromosomes are likely *Psy1*. Supporting this hypothesis, several researchers (Colasuonno et al., 2014; Pozniak et al., 2007) mapped *Psy1* to the linkage group 7 chromosomes in durum wheat.

In this study, we identified key genomic regions associated with traits of interest, which can be converted into KASP markers. The identified SNPs were selected based on their strong association with the target traits, ensuring their relevance in marker-assisted selection (MAS). Before implementation, selected SNPs would need to be validated for their polymorphism and robustness in a diverse germplasm panel. The development process would involve optimizing allele-specific primer design to ensure high genotyping efficiency and minimal amplification errors.

If successfully developed, these KASP markers could provide breeders with a reliable tool for rapid screening of desirable alleles, thus improving selection accuracy and accelerating genetic gain. Additionally, testing the markers across different genetic backgrounds would be essential to confirm their transferability and practical utility in breeding programs.

Future research should focus on further validating these potential markers in breeding populations, assessing their effectiveness in genomic selection models, and exploring their application in marker-assisted backcrossing for the introgression of favorable alleles into elite cultivars. The integration of KASP markers into breeding pipelines could ultimately enhance selection efficiency and contribute to the rapid development of improved varieties.

5 Conclusion

The genetic improvement of GPC and carotenoid content in wheat is crucial for enhancing both the nutritional quality and marketability of wheat products. Advances in genetic and genomic tools, coupled with a deeper understanding of the biochemical pathways involved in N metabolism and carotenoid biosynthesis, offer promising strategies for developing wheat varieties that meet the dual challenges of high yield and quality under sustainable agricultural practices. To the best of our knowledge, this is the first study that evaluates both GPC and GY using a diverse population of tetraploid wheats, which confers scientific novelty to this work. Future research should focus on

validating these QTL in diverse genetic backgrounds and under varying environmental conditions to ensure their stability and effectiveness in breeding programs. Additionally, the functional characterization of the identified genes through gene editing and transcriptomic analyses could provide deeper insights into the molecular mechanisms governing GPC and YI, covering the way for more precise genetic interventions, which could in turn benefit the farmers and the consumers.

Overall, this study underscores the potential of genomic tools and approaches in advancing wheat breeding, ultimately contributing to improved crop quality and sustainability in wheat production systems.

6 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author Contributions

IM: Conceptualization, Data curation, Formal Analysis, Investigation, Writing – original draft, Writing – review & editing, Funding acquisition. PCL: Conceptualization, Data curation, Formal Analysis, Investigation, Writing – original draft, Writing – review & editing. DC: Formal Analysis. AM: Review & editing. SLG: Formal Analysis. PC: Writing – review & editing. SM: Formal Analysis. MS: Formal Analysis. ARS: Funding acquisition, Writing – review & editing. AG: Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

8 Funding

This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them. Additionally, this research was funded by Prima MED 2022-2025, section 2, AgroEcoMed and PRIN 2022 PNRR "Wheatsustain"P2022NJS4Y – CUP H53D23010650001.

9 Data Availability Statement

The datasets for this study will be available under request to the authors.

10 References

Akhunov, E., Nicolet, C., Dvorak, J. 2009. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theor Appl Genet.* 119(3):507–517

Akond, Z., Alam, M.J., Hasan, M.N., Uddin, M.S., Alam, M., Mollah, M.N.H. 2019. A Comparison on Some Interval Mapping Approaches for QTL Detection. *Bioinformation*. 28;15(2):90-94. doi: 10.6026/97320630015090.

Alzuwaid, N.T., Pleming, D., Fellows, C.M., Laddomada, B., Sissons, M. 2021. Influence of durum wheat bran particle size on phytochemical content and on leavened bread baking quality. *Foods.* 10, 489. doi: 10.3390/foods10030489.

Arriagada, O., Gadaleta, A., Marcotuli, I., Maccaferri, M., Campana, M., Reveco, S., Alfaro, C., Matus, I., Schwember, A.R. 2022. A comprehensive meta-QTL analysis for yield-related traits of durum wheat (*Triticum turgidum* L. var. *durum*) grown under different water regimes. *Front Plant Sci*. 13, 3329. doi: 10.3389/fpls.2022.984269

Avni, R., Nave, M., Eilam, T., Sela, H., Alekperov, C., Peleg, Z., Dvorak, J., Korol, A., Distelfeld, A. 2014. Ultra-dense genetic map of durum wheat × wild emmer wheat developed using the 90K iSelect SNP genotyping assay. *Mol. Breed.* 34,1549-1562. doi:10.1007/s11032-014-0176-2

Balyan, H. S., Gahlaut, V., Kumar, A., Jaiswal, V., Dhariwal, R., Tyagi, S., Agarwal, P., Kumari, S., Gupta, P.K. 2016. Nitrogen and phosphorus use efficiencies in wheat: Physiology, phenotyping, genetics, and breeding. *Plant breed rev*. vol. 40. (Hoboken, New Jersey: John Wiley & Sons, Ltd), 167–234. doi: 10.1002/9781119279723.ch4.

Biradar, S.S., Yashavanthakumar, K.J., Navathe, S., Reddy, U.G., Baviskar, V.S., Gopalareddy, K., Lamani, K., Desai, S.A. 2022. Dicoccum wheat: Current status and future perspectives. *New Horizons in Wheat and Barley Research*. Springer, Singapore, 531-563.doi: 10.1007/978-981-16-4449-8_21.

Blanco, A., Mangini, G., Giancaspro, A., Giove, S., Colasuonno, P., Simeone, R. 2012. Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars. *Mol Breed.* 30, 79–92. doi: 10.1007/s11032-011-9600-z.

Blanco, A., Pasqualone, A., Troccoli, A., Di Fonzo, N., Simeone, R. 2002. Detection of grain protein content QTLs across environments in tetraploid wheats. *Plant Mol Biol.* 48, 615-623. doi: 10.1023/A:1014864230933.

Bogard, M., Allard, V., Brancourt-Hulmel, M., Heumez, E., Machet, J.M., Jeuffroy, M.H., Gate, P., Martre, P., Le Gouis, J. 2010. Deviation from the grain protein concentration–grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. *J Exp Bot.* 61, 4303—4312. doi: 10.1093/jxb/erq238.

Brevis, J. C., and Dubcovsky, J. 2010. Effects of the chromosome region including the Gpc-B1 locus on wheat grain and protein yield. *Crop Sci.* 50, 93-104. doi: 10.2135/cropsci2009.02.0057.

Cabas-Lühmann, P., Arriagada, O., Matus, I., Marcotuli, I., Gadaleta, A., Schwember, A.R. 2023. Comparison of durum with ancient tetraploid wheats from an agronomical, chemical, nutritional, and genetic standpoints: a review. *Euphytica.* 219, 61. doi: 10.1007/s10681-023-03188-z.

Cabas-Lühmann, P., Schwember, A.R., Arriagada, O., Marcotuli, I., Matus, I., Alfaro, C., Gadaleta, A. 2024. Meta-QTL analysis and candidate genes for quality traits, mineral content, and abiotic-related traits in wild emmer. *Front Plant Sci.* 15, 1305196. doi: 10.3389/fpls.2024.1305196.

Campos, K.M., Royo, C., Schulthess, A., Villegas, D., Matus, I., Ammar, K., Schwember, A.R. 2016. Association of *phytoene synthase Psy1-A1* and *Psy1-B1* allelic variants with semolina yellowness in durum wheat (*Triticum turgidum* L. var. *durum*). *Euphytica.* 207, 109–117. doi: 10.1007/s10681-015-1541-x.

Castaings, L., Camargo, A., Pocholle, D., Gaudon, V., Texier, Y., Boutet-Mercey, S., Taconnat, L., Renou, J.P., Daniel-Vedele, F., Fernandez, E., Meyer, C., Krapp, A. 2009. The nodule inception-like

583 protein 7 modulates nitrate sensing and metabolism in *Arabidopsis*. *Plant J.* 57, 426-435. doi:
584 10.1111/j.1365-313X.2008.03695.x.

585 Cazzonelli, C.I., Nisar, N., Hussain, D., Carmody, M.E., Pogson, B.J. 2010. Biosynthesis and
586 regulation of carotenoids in plants—micronutrients, vitamins and health benefits. *Plant*
587 *developmental biology – biotechnological perspectives*. 117-137. doi: 10.1007/978-3-642-04670-
588 4_7.

589 Clarke B, Liang R, Morell M, Bird A, Jenkins C, Li Z. 2008. Gene expression in a starch synthase
590 IIa mutant of barley: changes in the level of gene transcription and grain composition. *Funct Integr*
591 *Genomics.* 8(3):211–21.

592 Colasuonno, P., Lozito, M. L., Marcotuli, I., Nigro, D., Giancaspro, A., Mangini, G., De Vita, P.,
593 Mastrangelo, A.M., Pecchioni, N., Houston, K., Simeone, R., Gadaleta, A., Blanco, A. 2017. The
594 carotenoid biosynthetic and catabolic genes in wheat and their association with yellow pigments.
595 *BMC Genomics.* 18, 122. doi: 10.1186/s12864-016-3395-6.

596 Colasuonno, P., Marcotuli, I., Blanco, A., Maccaferri, M., Condorelli, G.E., Tuberosa, R., Parada, R.,
597 Costa, A., Schwember, A.R., Gadaleta, A. 2019. Carotenoid pigment content in durum wheat
598 (*Triticum turgidum* L. var *durum*): An overview of quantitative trait loci and candidate genes. *Front*
599 *Plant Sci.* 10, 1347. doi:10.3389/fpls.2019.01347

600 Cong, L., Wang, C., Li, Z., Chen, L., Yang, G., Wang, Y., He, G. (2010). cDNA cloning and
601 expression analysis of wheat (*Triticum aestivum* L.) phytoene and ζ-carotene desaturase genes. *Mol.*
602 *Biol. Rep*. 37, 3351-3361. doi: 10.1007/s11033-009-9922-7.

603 Delcour, J. and Hoseney, R. C. 2010. Principles of cereal science and technology (3rd ed.). *USA:*
604 *AACC International*. doi: 10.1094/9781891127632.

605 Dibari, B., Murat, F., Chosson, A., Gautier, V., Poncet, C., Lecomte, P., Mercier, I., Bergès, H., Pont,
606 C., Blanco, A. Salse, J. 2012. Deciphering the genomic structure, function and evolution of
607 carotenogenesis related phytoene synthases in grasses. *BMC Genomics*. 13, 1-14. doi: 10.1186/1471-
608 2164-13-221.

609 Digesù, A. M., Platani, C., Cattivelli, L., Mangini, G., Blanco, A. 2009. Genetic variability in yellow
610 pigment components in cultivated and wild tetraploid wheats. *J. Cereal Sci.* 50, 210–218.
611 doi:10.1016/j.jcs.2009.05.002.

612 Elouafi, I., Nachit, M.M., Martin, L.M. 2001. Identification of a microsatellite on chromosome 7B
613 showing a strong linkage with yellow pigment in durum wheat (*Triticum turgidum* L. var. *durum*).
614 *Hereditas*. 135, 255–261. doi: 10.1111/j.1601-5223.2001.t01-1-00255.x

615 Evanno, G., Regnaut, S., Goudet, J. 2005. Detecting the number of clusters of individuals using the
616 software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611-2620. doi: 10.1111/j.1365-
617 294X.2005.02553.x.

618 Fatiukha, A., Filler, N., Lupo, I., Lidzbarsky, G., Klymiuk, V., Korol, A.B., Pozniak, C., Fahima, T.,
619 Krugman, T. 2020. Grain protein content and thousand kernel weight QTLs identified in a
620 durum × wild emmer wheat mapping population tested in five environments. *Theor. Appl. Genet.* 133,
621 119-131. doi:10.1007/s00122-019-03444-8

622 Ficco, D.B., Mastrangelo, A.M., Trono, D., Borrelli, G.M., De Vita, P., Fares, C., Beleggia, R.,
623 Platani, C., Papa, R. 2014. The colours of durum wheat: a review. *Crop Pasture Sci*. 65, 1–15.
624 doi:10.1071/CP13293.

625 Fiedor, J., and Burda, K. 2014. Potential role of carotenoids as antioxidants in human health and
626 disease. *Nutrients*, 6, 466-488. doi: 10.3390/nu6020466.

627 Fontaine, J.X., Ravel, C., Pageau, K., Heumez, E., Dubois, F., Hirel, B., Le Gouis, J. 2009. A
628 quantitative genetic study for elucidating the contribution of glutamine synthetase, glutamate
629 dehydrogenase and other nitrogen-related physiological traits to the agronomic performance of
630 common wheat. *Theor. Appl. Genet*. 119, 645–662. doi: 10.1007/s00122-009-1076-4.

631 Fortunato, S., Nigro, D., Lasorella, C., Marcotuli, I., Gadaleta, A., de Pinto, M. C. 2023. The role of
632 *glutamine synthetase* (*GS*) and *glutamate synthase* (*GOGAT*) in the Improvement of nitrogen use
633 efficiency in cereals. *Biomolecules*. 13, 1771. doi: 10.3390/biom13121771.

634 Foulkes, M.J., Hawkesford, M.J., Barraclough, P.B., Holdsworth, M.J., Kerr, S., Kightley, S.,
635 Shewry, P.R. 2009. Identifying traits to improve the nitrogen economy of wheat: Recent advances
636 and future prospects. *Field Crops Res*. 114, 329–342. doi: 10.1016/j.fcr.2009.09.005

637 Fu, B. X., Schlichting, L., Pozniak, C. J., Singh, A. K. 2011. A fast, simple, and reliable method to
638 predict pasta yellowness. *Cereal Chem*. 88, 264-270. doi: 10.1094/CCHEM-12-10-0173.

639 Gadaleta, A., Nigro, D., Giancaspro, A., Blanco, A. 2011. The *glutamine synthetase* (*GS2*) genes in
640 relation to grain protein content of durum wheat. *Funct. Integra. Genomics*. 11, 665-670. doi:
641 10.1007/s10142-011-0235-2.

642 Gadaleta, A., Nigro, D., Marcotuli, I., Giancaspro, A., Blanco, A. 2014. Molecular characterization
643 of candidate genes involved in nitrogen metabolism and relationship with the grain protein content
644 of wheat. *OPTIONS MÉDITERRANÉENNES*. SÉRIE A: SÉMINAIRES MÉDITERRANÉENS.

645 Gallagher, C.E., Matthews, P.D., Li, F., Wurtzel, E.T. 2004. Gene duplication in the carotenoid
646 biosynthetic pathway preceded evolution of the grasses. *Plant Physiol*. 135, 1776–1783. doi:
647 10.1104/pp.104.039818.

648 Gao, R., Curtis, T. Y., Powers, S. J., Xu, H., Huang, J., Halford, N. G. 2016. Food safety: structure
649 and expression of the asparagine synthetase gene family of wheat. *J. Cereal Sci*. 68, 122-131. doi:
650 10.1016/j.jcs.2016.01.010.

651 Garnett, T., Conn, V., Kaiser, B.N. 2009. Root based approaches to improving nitrogen use efficiency
652 in plants. *Plant Cell Environ.* 32, 1272-1283. doi: 10.1111/j.1365-3040.2009.02011.x.

653 Habash, D.Z., Bernard, S., Schondelmaier, J., Weyen, J., Quarrie, S.A. 2007. The genetics of nitrogen
654 use in hexaploid wheat: N utilisation, development and yield. *Theor. Appl. Genet*. 114, 403–419. doi:
655 10.1007/s00122-006-0429-5.

656 He, X. Y., Zhang, Y. L., He, Z. H., Wu, Y. P., Xiao, Y. G., Ma, C. X., et al. 2008. Characterization
657 of phytoene synthase 1 gene (Psy1) located on common wheat chromosome 7A and development of
658 a functional marker. *Theor. Appl. Genet*. 116, 213–221. doi: 10.1007/s00122-007-0660-8

He, X.Y., He, Z.H., Ma, W., Appels, R., Xia, X.C. 2009. Allelic variants of phytoene synthase 1 (Psy1) genes in Chinese and CIMMYT wheat cultivars and development of functional markers for flour colour. Mol. Breeding, 23, 553-563. doi: 10.1007/s11032-009-9255-1.

He, X.Y., Zhang, Y.L., He, Z.H., Wu, Y.P., Xiao, Y.G., Ma, C.X., Xia, X.C. 2008. Characterization of *phytoene synthase 1* gene (*Psy1*) located on common wheat chromosome 7A and development of a functional marker. *Theor. Appl. Genet*. 116, 213-221. doi: 10.1007/s00122-007-0660-8.

Hessler, T. G., Thomson, M. J., Benscher, D., Nachit, M., and Sorrells, M. 2002. Association of a lipoxygenase locus, Lpx-B1, with variation in lipoxygenase activity in durum wheat seeds. *Crop Sci*. 42, 1695–1700. doi: 10.2135/ cropsci2002.1695.

Hessler, T. G., Thomson, M. J., Benscher, D., Nachit, M., Sorrells, M., 2002. Association of a lipoxygenase locus, Lpx-B1, with variation in lipoxygenase activity in durum wheat seeds. *Crop Sci*. 42, 1695–1700. doi: 10.2135/cropsci2002.1695.

Hirschberg, J. 2001. Carotenoid biosynthesis in flowering plants. *Curr. Opin. Plant Biol*. 4, 210-218. doi: 10.1016/S1369-5266(00)00163-1.

Howitt, C.A., Cavanagh, C.R., Bowerman, A.F., Cazzonelli, C., Rampling, L., Mimica, J.L., Pogson, B.J. 2009. Alternative splicing, activation of cryptic exons and amino acid substitutions in carotenoid biosynthetic genes are associated with lutein accumulation in wheat endosperm. *Funct. Integr. Genomics.* 9, 363–376. doi: 10.1007/s10142-009-0121-3.

Huang, M., Liu, X., Zhou, Y., Summers, R.M., Zhang, Z. 2019. BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience.* giy154. doi:10.1093/gigascience/giy154

Kuhn, J.C., Stubbs, T.L. and Carter, A.H. 2016. Effect of the *Gpc-B1* Allele in Hard Red Winter Wheat in the US Pacific Northwest. Crop Science, 56: 1009-1017. https://doi.org/10.2135/cropsci2015.08.0470

Joppa, L.R., and Cantrell, R.G. (1990). Chromosomal location of genes for grain protein content of wild tetraploid wheat. *Crop Sci.* 30, 1059-1064. Doi:10.2135/cropsci1990.0011183X003000050021x

Joppa, L.R., Du, C., Hart, G., Hareland, G. 1997. Mapping gene(s) for grain protein in tetraploid wheat (*Triticum turgidum* l.) using a population of recombinant inbred chromosome lines. *Crop Sci*. 37, 1586-1589. doi:10.2135/cropsci1997.0011183X003700050030x

Joppa, L.R., Hareland, G.A., Cantrell, R.G. 1991. Quality characteristics of the Langdon durum-dicoccoides chromosome substitution lines. *Crop Sci.* 31, 1513-1517. doi:10.2135/cropsci1991.0011183X003100060024x

Kastori, R.R., and Petrović, N.M. 2004. Bioenergetic aspects of competition between nitrogen metabolism and carbohydrate synthesis in small grains. *Matica Srpska Proceedings for Natural Sciences*. (106), 21-27. doi: 10.2298/ZMSPN0416021K.

Kuchel H, Langridge P, Mosionek L, Williams K, Jefferies S P. 2006. The genetic control of milling yield, dough rheology and baking quality of wheat. *Theor. Appl. Genet.*, 112, 1487-1495.

Ishikawa, G.; Sakai, H.; Mizuno, N.; Solovieva, E.; Tanaka, T.; Matsubara, K. 2022. Developing core marker sets for effective genomic-assisted selection in wheat and barley breeding programs. *Breed Sci*. 72(3):257-266. doi: 10.1270/jsbbs.22004.

Lap, B., Rai, M., Tyagi, W. 2021. Playing with colours: genetics and regulatory mechanisms for anthocyanin pathway in cereals. *Biotechnol. Genet. Eng. Rev*. 37, 1-29. doi: 10.1080/02648725.2021.1928991.

Lea, P.J., Azevedo, R.A. 2007. Nitrogen use efficiency. 2. Amino acid metabolism. *Ann. Appl. Biol.* 151, 269–275. doi: 10.1111/j.1744-7348.2007.00200.x.

Leigh, F.J., Wright, T.I., Horsnell, R.A., Dyer, S., Bentley, A. R. 2022. Progenitor species hold untapped diversity for potential climate-responsive traits for use in wheat breeding and crop improvement. *Heredity*. 128, 291-303. doi: 10.1038/s41437-022-00527-z.

Linquist, B., Van Groenigen, K. J., Adviento-Borbe, M. A., Pittelkow, C., Van Kessel, C. 2012. An agronomic assessment of greenhouse gas emissions from major cereal crops. *Glob. Change Biol.* 18, 194-209. doi: 10.1111/j.1365-2486.2011.02502.x

Liu, Y., Fu, B., Zhang, Qiaofeng, Z., Jin. C., Wei. G., Wenling, Z., Jizhong, W. 2024. Genetic diversity and population structure of wheat landraces in Southern Winter Wheat Region of China. *BMC Genomics* **25**, 664. https://doi.org/10.1186/s12864-024-10564-z

Lu, J., Zhang, L., Lewis, R.S., Bovet, L., Goepfert, S., Jack, A.M., Crutchfield, J.D., Ji, H. and Dewey, R.E. 2016. Expression of a constitutively active nitrate reductase variant in tobacco reduces tobacco-specific nitrosamine accumulation in cured leaves and cigarette smoke. *Plant biotechnology journal*, *14*(7), pp.1500-1510.

Ma W, Daggard G, Sutherland M, Brennan P. 1999. Molecular markers for quality attributes in wheat. In: Williamson P, Banks P, Haak I, Thompson J, Campbell A, eds., *Proceedings of the 9th Assembly of the Wheat Breeding Society of Australia*. Toowoomba, Australia. pp. 115-117.

Maccaferri, M., Cane', M. A., Sanguineti, M.C., Salvi, S., Colalongo, M.C., Massi, A., Clarke, F., Knox, R., Pozniak, C.J., Clarke, J.M., Fahima, T., Dubcovsky, J., Xu, S., Ammar, K., Karsai, I., Vida, G., Tuberosa, R. 2014. A consensus framework map of durum wheat (*Triticum durum* Desf.) suitable for linkage disequilibrium analysis and genome-wide association mapping. *BMC Genomics*. 15, 1-21. doi: 10.1186/1471-2164-15-873.

Marcotuli, I., Houston, K., Schwerdt, J.G., Waugh, R., Fincher, G.B., Burton, R.A., Blanco, A., Gadaleta, A. 2016. Genetic diversity and genome wide association study of β-glucan content in tetraploid wheat grains. *PLoS One*. 11, e0152590. doi: 10.1371/journal.pone.0152590.b

Marcotuli, I., Soriano, J. M., Gadaleta, A. 2022. A consensus map for quality traits in durum wheat based on genome-wide association studies and detection of ortho-meta QTL across cereal species. *Front. Genet.* 13. doi: 10.3389/fgene.2022.982418

Mares, D. J. and Campbell, A.W. 2001. Mapping components of flour and noodle colour in Australian wheat. *Aust. J. Agric. Res*. 52, 1297–1309. doi: 10.1071/AR01048.

Moise, A.R., Al-Babili, S., Wurtzel, E.T. 2014. Mechanistic aspects of carotenoid biosynthesis. *Chem. Rev*. 114, 164–193. doi: 10.1021/cr400106y.

N'Diaye, A., Haile, J.K., Nilsen, K.T., Walkowiak, S., Ruan, Y., Singh, A.K., Clarke, F.R., Clarke, J.M., Pozniak, C.J. 2018. Haplotype loci under selection in Canadian durum wheat germplasm over 60 years of breeding: association with grain yield, quality traits, protein loss, and plant height. *Front. Plant Sci*. 9, 1589. doi: 10.3389/fpls.2018.01589.

Nigro, D., Blanco, A., Anderson, O.D., Gadaleta, A. 2014. Characterization of ferredoxin-dependent glutamine-oxoglutarate amidotransferase (*Fd-GOGAT*) genes and their relationship with grain protein content QTL in wheat. *PLoS One*. 9, e103869. doi: 10.1371/journal.pone.0103869.

Nigro, D., Fortunato, S., Giove, S. L., Mazzucotelli, E., Gadaleta, A. 2020. Functional validation of glutamine synthetase and glutamate synthase genes in durum wheat near isogenic lines with QTL for high GPC. *Int. J. Mol. Sci.* 21, 9253. doi: 10.3390/ijms21239253.

Nigro, D., Fortunato, S., Giove, S.L., Paradiso, A., Gu, Y.Q., Blanco, A., de Pinto, M.C., Gadaleta, A. 2016. Glutamine synthetase in durum wheat: genotypic variation and relationship with grain protein content. *Plant Sci.* 7, 971. doi: 10.3389/fpls.2016.00971

Nigro, D., Gadaleta, A., Mangini, G., Colasuonno, P., Marcotuli, I., Giancaspro, A., Giove, S.L., Simeone, R., Blanco, A. 2019. Candidate genes and genome-wide association study of grain protein content and protein deviation in durum wheat. *Planta*. 249, 1157-1175. doi: 10.1007/s00425-018-03075-1.

Nigro, D., Gu, Y.Q., Huo, N., Marcotuli, I., Blanco, A., Gadaleta, A., Anderson, O.D. 2013. Structural analysis of the wheat genes encoding NADH-dependent glutamine-2-oxoglutarate amidotransferases and correlation with grain protein content. *Plos One*. 8, e73751. doi: 10.1371/journal.pone.0073751.

Olmos, S., Diestelfeld, A., Chicaiza, O., Schatter, A.R., Fahima, T., Echenique, V., Dubcovsky, J. 2003. Precise mapping of a locus affecting grain protein content in durum wheat. *Theor. Appl. Genet.* 107, 1243–1251. doi:10.1007/s00122-003-1377-y

Ortiz-MonasterioR, J.I., Sayre, K.D., Rajaram, S., McMahon, M. 1997. Genetic progress in wheat yield and nitrogen use efficiency under four nitrogen rates. *Crop Sci*. 37, 898-904. doi: 10.2135/cropsci1997.0011183X003700030033x.

Palaisa, K. A., Morgante, M., Williams, M., Rafalski, A. 2003. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell*. 15, 1795-1806. doi: 10.1105/tpc.012526.

Parada, R., Royo, C., Gadaleta, A., Colasuonno, P., Marcotuli, I., Matus, I., Castillo, D., Costa de Camargo, A., Araya-Flores, J., Villegas, D., Schwember, A.R. 2020. *Phytoene synthase 1 (Psy-1)* and *lipoxygenase 1 (Lpx-1)* genes influence on semolina yellowness in wheat mediterranean germplasm. *Int. J. Mol. Sci.* 21, 4669. doi: 10.3390/ijms21134669.

Parker G D, Chalmers K J, Rathgen A J, Langridge P. 1998. Mapping loci associated with flour colour in wheat (Triticum aestivum L.). *Theor. Appl. Genet.*, 97, 238-245.

Patil, R., Oak, M., Deshpande, A., Tamhankar, S. 2018. Development of a robust marker for Psy-1 homoeologs and its application in improvement of yellow pigment content in durum wheat. *Mol. Breed.* 38, 136. doi: 10.1007/s11032-018-0895-x.

773 Pozniak, C. J., Knox, R. E., Clarke, F. R., & Clarke, J. 2007. Identification of QTL and association
774 of a phytoene synthase gene with endosperm colour in durum wheat. *Theor. Appl. Genet.* 114, 525–
775 537. doi: 10.1007/s00122-006-0453-5.

776 Quarrie, S.A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy, C., Steele, N., et al.
777 2005. A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese
778 Spring x SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor.*
779 *Appl. Genet*. 110, 865-880. doi: 10.1007/s00122-004-1902-7.

780 Ramachandran, A., Pozniak, C.J., Clarke, J.M., Singh, A.K. 2010. Carotenoid accumulation during
781 grain development in durum wheat. *J. Cereal Sci*. 52, 30–38. doi:10.1016/j. jcs.2010.02.014

782 Richaud, D., Stange, C., Gadaleta, A., Colasuonno, P., Parada, R., Schwember, A.R. 2018.
783 Identification of Lycopene epsilon cyclase (LCYE) gene mutants to potentially increase β-carotene
784 content in durum wheat (*Triticum turgidum* L. ssp. *durum*) through TILLING. *PloS One.* 13,
785 e0208948. doi: 10.1371/journal.pone.0208948.

786 Rufo, R., López, A., Lopes, M.S., Bellvert, J., Soriano, J.M. 2021. Identification of Quantitative Trait
787 Loci Hotspots Affecting Agronomic Traits and High-Throughput Vegetation Indices in Rainfed
788 Wheat. *Front. Plant Sci*. 12, 735192. doi: 10.3389/fpls.2021.735192.

789 Ruiz-Sola, M.Á., Rodríguez-Concepción, M. 2012. Carotenoid biosynthesis in *Arabidopsis*: a
790 colorful pathway. *Arabidopsis Book*, 10, e0158. doi: 10.1199/tab.0158.

791 Saini P, Sheikh I, Saini DK, Mir RR, Dhaliwal HS and Tyagi V. 2022. Consensus genomic regions
792 associated with grain protein content in hexaploid and tetraploid wheat. *Front. Genet.* 13:1021180.
793 doi: 10.3389/fgene.2022.1021180

794 Sandmann, G. 2001. Genetic manipulation of carotenoid biosynthesis: strategies, problems and
795 achievements. *Trends Plant Sci.* 6, 14–17. doi:10.1016/S1360-1385(00)01817-3

796 Schulthess, A., Matus, I., Schwember, A. 2013. Genotypic and environmental factors and their
797 interactions determine semolina color of elite genotypes of durum wheat (*Triticum turgidum* L. var.
798 *durum*) grown in different environments of Chile. *Field Crops Res*. 149, 234–244. doi:
799 10.1016/j.fcr.2013.05.001.

800 Seo, M., Koshiba, T. 2002. Complex regulation of ABA biosynthesis in plants. *Trends Plant Sci*. 7,
801 41-48. doi: 10.1016/S1360-1385(01)02187-2.

802 Shahi, D., Guo, J., Babar, M.A., Pradhan, S., Avci, M., Khan, N., McBreen, J., Rayamajhi, S., Liu,
803 Z., Bai, G., Amand, P. S., Bernardo, A., Reynolds, M., Molero, G., Sukumaran, S., Foulkes, J.,
804 & Khan, J. 2024. Deciphering the genetic basis of novel traits that discriminate useful and non-useful
805 biomass to enhance harvest index in wheat. *The Plant Genome*, 17,
806 e20512. https://doi.org/10.1002/tpg2.20512

807 Shewry, P.R. 2009. Wheat. *J. Exp. Bot*. 60, 1537-1553. doi: 10.1093/jxb/erp058.

808 Shrawat, A.K., Carroll, R.T., DePauw, M., Taylor, G.J., Good, A.G. 2008. Genetic engineering of
809 improved nitrogen use efficiency in rice by the tissue-specific expression of alanine aminotransferase.
810 *Plant Biotechnol. J.* 6, 337—345. doi: 10.1111/j.1467-7652.2008.00351.x.

Sissons, M., Smit, J., Kadkol, G., Sissons, S. 2020. A small-scale rapid screening protocol for quality traits of durum wheat. *Cereal Chem.* 97, 693-702. doi:10.1002/cche.10285

Subedi, M.; Ghimire, B.; Bagwell, J.W.; Buck, J.W.; Mergoum, M. 2023. Wheat end-use quality: State of art, genetics, genomics-assisted improvement, future challenges, and opportunities. *Front. Genet.* 13:1032601. doi: 10.3389/fgene.2022.1032601

Subedi, M., Bagwell, J. W., Lopez, B., Baik, B.-K., Babar, M. A., & Mergoum, M. 2024. A Genome-Wide Association Study Approach to Identify Novel Major-Effect Quantitative Trait Loci for End-Use Quality Traits in Soft Red Winter Wheat. *Genes*, *15*(9), 1177. https://doi.org/10.3390/genes15091177

Subira, J., Peña, R.J., Álvaro, F., Ammar, K., Ramdani, A., Royo, C. 2014. Breeding progress in the pasta–making quality of durum wheat cultivars released in Italy and Spain during the 20th Century. *Crop Pasture Sci*. 65, 16–26. doi: 10.1071/CP13238.

Tabbita, F., Pearce, S., & Barneix, A. J. 2017. Breeding for increased grain protein and micronutrient content in wheat: Ten years of the *GPC-B1* gene. *Journal of Cereal Science, 73*, 183-191. https://doi.org/10.1016/j.jcs.2017.01.003

Taylor, L., Nunes-Nesi, A., Parsley, K., Leiss, A., Leach, G., Coates, S., Wingler, A., Fernie, A.R., Hibberd, J.M. 2010. Cytosolic pyruvate, orthophosphate dikinase functions in nitrogen remobilization during leaf senescence and limits individual seed growth and nitrogen content. *Plant J.* 62, 641-652. doi: 10.1111/j.1365-313X.2010.04179.x.

Troccoli, A., Borrelli, G.M., De Vita, P., Fares, C., Di Fonzo, N. 2000. Durum wheat quality: a multidisciplinary concept. Mini review. *J. Cereal Sci*. 32, 99-113. doi: 10.1006/jcrs.2000.0322.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., Dubcovsky, J. 2006. A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314, 1298–1301. doi: 10.1126/science.1133649

Vishwakarma, M. K., Mishra, V. K., Gupta, P. K., Yadav, P. S., Kumar, H., & Joshi, A. K. 2014. Introgression of the high grain protein gene *Gpc-B1* in an elite wheat variety of Indo-Gangetic Plains through marker-assisted backcross breeding. *Current Plant Biology, 1*, 60-67. https://doi.org/10.1016/j.cpb.2014.09.003

Wang, J., Zhang, Z. 2021. GAPIT Version 3: Boosting Power and Accuracy for Genomic Association and Prediction. Genomics. *Proteomics Bioinformatics*. 19, 1–12. doi: 10.1016/j.gpb.2021.08.005.

Wang, M., Wang, Y., Wang, X., Wei, G., Yang, H., Yang, X., ... & Wu, Z. 2024. Integrated physiological, biochemical, and transcriptomics analyses reveal the underlying mechanisms of high nitrogen use efficiency of black sesame. *Plant Physiology and Biochemistry*. 206, 108205.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B.E., et al. 2014. Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183.

Xie, X., Yoneyama, K., Yoneyama, K. 2010. The strigolactone story. *Annu. Rev. Phytopathol.* 48, 93-117. doi: 10.1146/annurev-phyto-073009-114453.

Xu, G., Fan, X., Miller, A.J. 2012. Plant nitrogen assimilation and use efficiency. *Annu. Rev. Plant Biol*. 63, 153–182. doi: 10.1146/annurev-arplant-042811-105532.

Xu, Y., Wang, R., Tong, Y., Zhao, H., Xie, Q., Liu, D., Zhang, A., Li, B., Xu, H., An, D. 2014. Mapping QTLs for yield and nitrogen-related traits in wheat: influence of nitrogen and phosphorus fertilization on QTL expression. *Theor. Appl. Genet*. 127, 59–72. doi: 10.1007/s00122-013-2201-y.

Zhang L P, Yan J, Xia X C, He Z H, Sutherland M W. 2006. QTL mapping for kernel yellow pigment content in common wheat. *Acta Agronomica Sinica*, 32, 41-45.

Zhang, W., and Dubcovsky, J. 2008. Association between allelic variation at the Phytoene synthase 1 gene and yellow pigment content in the wheat grain. *Theor. Appl. Genet*. 116, 635–645. doi: 10.1007/s00122-007-0697-8.

Zhao Y, Islam S, Alhabbar Z, Zhang J, O'Hara G, Anwar M, Ma W. 2023. Current Progress and Future Prospect of Wheat Genetics Research towards an Enhanced Nitrogen Use Efficiency. *Plants*. 12(9):1753. doi: 10.3390/plants12091753. PMID: 37176811; PMCID: PMC10180859.

Tables

Table 1. Descriptive statistics for 144 tetraploid wheat genotypes for GPC and YI averaged per environment.

Trait	Statistics	Bari2019	Bari2022	Chile2021
GPC (%)	Mean	17.24	15.67	15.01
	SD (±)	1.60	1.87	1.62
	Median	17.10	15.47	14.78
	Min	13.29	11.6	10.59
	Max	27.35	21.7	22.68
	h^2	0.88	0.90	0.44
YI (CIE *b**)	Mean	9.92	13.06	13.05
	SD (±)	0.88	0.88	0.87
	Median	9.78	13.15	12.99
	Min	8.21	10.73	10.80
	Max	13.54	16.37	16.60
	h^2	0.88	0.66	0.40

***$P \leq 0.0001$; ni, the significance of the environment was not included because of the amount of missing data for the traits in Bari 22.

Table 2. SNP markers significantly associated (LOD≥3) with protein content and yellow index identified by GWAS in the whole tetraploid wheat collection evaluated in three environments (Chile2021, Bari2019 and Bari2022) and the associated genes. Green font indicates the QTL identified with the MLM model, red font the ones with the BLINK model, black font the QTL detected with both the methodology. The cM position is referred to the consensus map by Maccaferri et al. (2014), while the bp position is based on the reference Chinese Spring genome assembly (IWGSC RefSeq v2.1.).

Traits	QTL	Closest marker	SNP alleles	Chr	cM	Position (bp)	Environments									Mean			Candidate gene
							Chile2021			Bari2019			Bari2022						
							LOD	R^2	marker effect	LOD	R^2	marker effect	LOD	R^2	marker effect	LOD	R^2	marker effect	
GPC	*QGpc.bc.1B*	AX-158606326	C/T	1B	-	653,263,986	**3.26**	**0.10**	**-0.42**	-	-	-	2.98	0.09	-0.40	**3.53**	**0.11**	**-0.41**	-
	*QGpc.bc.2A**	BobWhite_c3833_815	G/T	2A	145.9	692,850,242	**3.04**	**0.08**	**1.12**	-	-	-	-	-	-	-	-	-	*glutamine synthetase (GS2)*

	QGpc.bc.3A.1	AX-158533209*	C/T	3A	-	26,002,005	2,99	0.08	3.98	-	-	-	**3.20**	**0.09**	**4.27**	-	-	-	-
	QGpc.bc.3A.2	AX-158532914	A/C	3A	-	697,783,401	-	-	-	-	-	-	**3.15**	**0.09**	**-4.39**	-	-	-	-
	QGpc.bc.3B	BS00029548_51	A/T	3B	75.2	207,223,999	-	-	-	**3.01**	**0.09**	**-0.47**	3.00	0.08	-0.39	-	-	-	*NADH-dependent glutamate synthase (NADH-GOGAT)*
	QGpc.bc.4B	RAC875_c35152_372	T/C	4B	62.9	165,019,431	-	-	-	**3.46**	**0.10**	**0.66**	-	-	-	-	-	-	-
	QGpc.bc.5A	BobWhite_rep_c63332_67	A/G	5A	184.4	677,845,586	**3.09**	**0.10**	**1.39**	-	-	-	2.99	0.08	0.59	**3.46**	**0.09**	**0.45**	-
	QGpc.bc.6B	RAC875_c45987_132	A/C	6B	190.4	712,243,648	-	-	-	**4.16**	**0.12**	**-0.72**	-	-	-	-	-	-	-
	QGpc.bc.7A-1	AX-86163414	G/T	7A	-	69,442,303	-	-	-	-	-	-	-	-	-	**3.11**	**0.08**	**3.26**	-
	QGpc.bc.7A-2	BS00022137_51	A/G	7A	203.2	729,397,521	-	-	-	**3.32**	**0.10**	**-0.49**	-	-	-	-	-	-	-
	QYI.bc.1B	AX-94522424	C/G	1B	-	679,342,319	2.99	0.09	0.54	3.02	0.08	0.96	-	-	-	**3.17**	**0.10**	**0.16**	-
	QYI.bc.2A	AX-109390022	C/T	2A	-	45,120,870	-	-	-	**3.21**	**0.10**	**0.40**	2.98	0.08	0.10	**2.98**	**0.09**	**0.20**	
	QYI.bc.2B	BS00083078_51	A/G	2B	71.0	131,130,446	**3.31**	**0.13**	**-0.29**	4.37	0.13	-0.58	-	-	-	-	-	-	-
	QYI.bc.3A	AX-158533209*	C/T	3A	-	26,002,005	-	-	-	2.97	0.09	1.86	**3.33**	**0.09**	**3.58**	**3.57**	**0.10**	**1.32**	-
	QYI.bc.4A	AX-158564262	T/C	4A	-	45,859,413	-	-	-	-	-	-	-	-	-	**3.73**	**0.11**	**0.18**	-
	QYI.bc.4B.1	Tdurum_contig67399_676	A/G	4B	13.9	13,945,623	-	-	-	3.37	0.09	0.91	-	-	-	-	-	-	-
YI	*QYI.bc.4B.2*	IAAV3421	C/T	4B	69.3	538,997,220	**4.10**	**0.12**	**0.85**	-	-	-	-	-	-	-	-	-	-
	QYI.bc.5A	AX-110949650	A/C	5A	-	540,369,204	-	-	-	-	-	-	**3.13**	**0.12**	**-0.32**	-	-	-	-
	QYI.bc.6A	AX-109333268	C/G	6A	-	580,190,968	3.30	0.09	0.90	-	-	-	-	-	-	-	-	-	-
	QYI.bc.6B.1	BS00022823_51	G/T	6B	52.4	94,000,112	-	-	-	3.87	0.11	0.93	-	-	-	-	-	-	-
	QYI.bc.7A.1	Kukri_c39894_178	A/G	7A	111.1	232,593,135	**5.25**	**0.15**	**-0.29**	-	-	-	-	-	-	-	-	-	*Phytoene synthase 1 (Psy-A1)*
	QYI.bc.7A.2	AX-89345811	A/G	7A	-	690,870,014	3.12	0.08	0.03	3.06	0.09	0.33	**3.82**	**0.14**	**0.44**	2.99	0.10	0.10	-

Figures

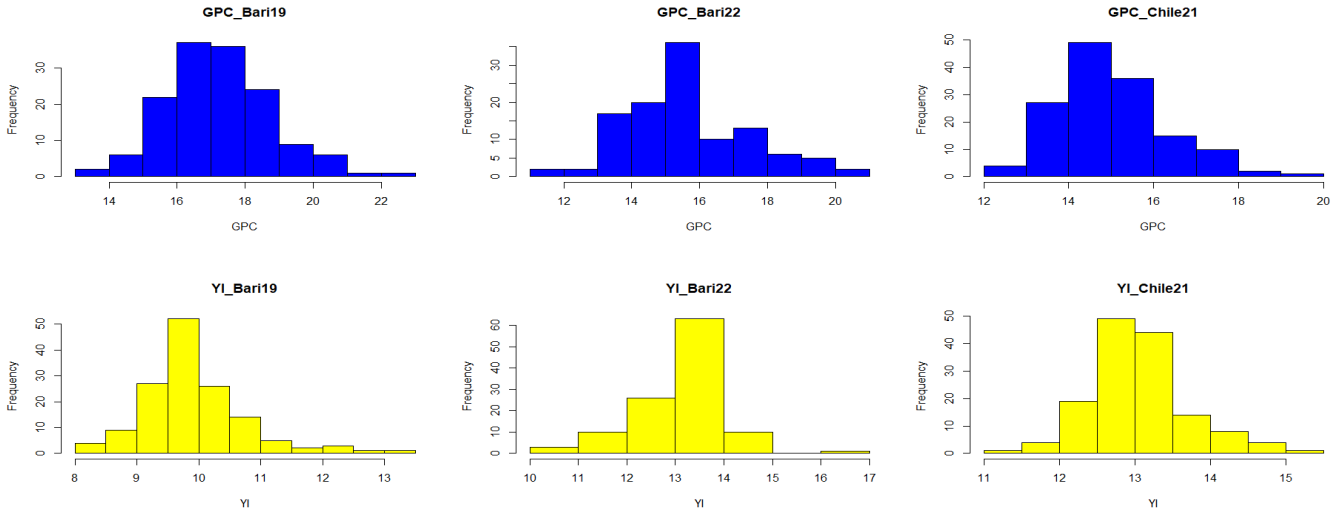


Figure 1. Frequency distribution plots of phenotypic values for grain protein content (GPC) and yellow index (YI) in 144 *Triticum turgidum* accessions grown in Bari 2019, Bari 2022 and Chile 2021.

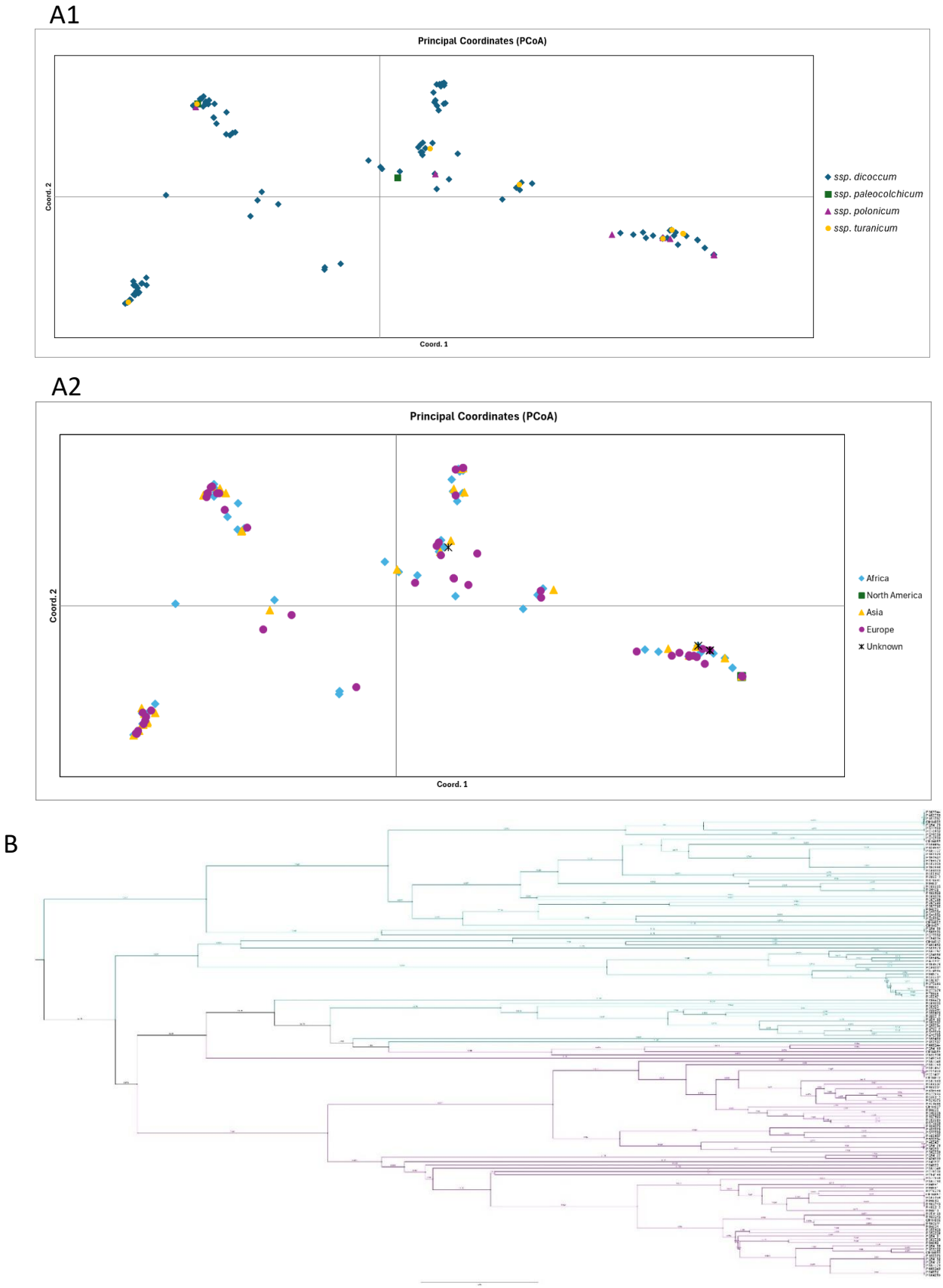
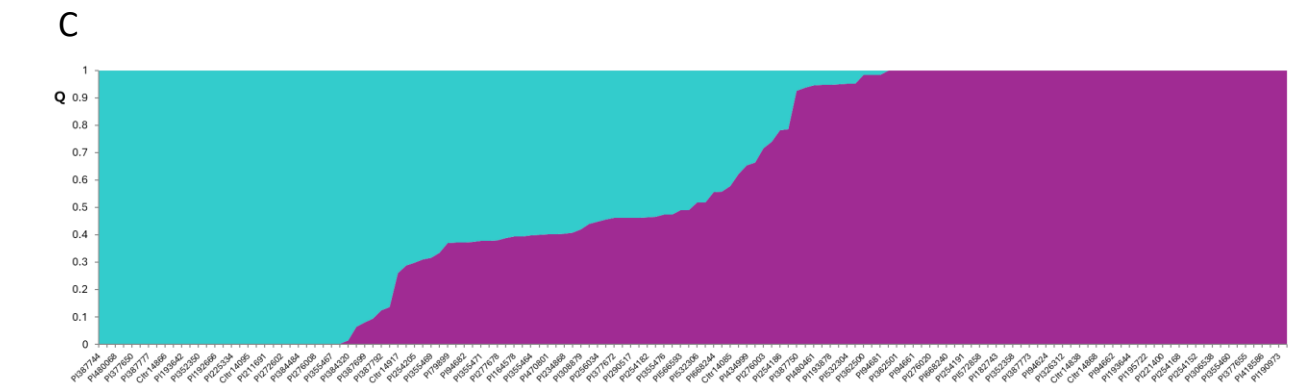


Figure 2. Genotypic data used to carry out a genome wide association study (GWAS). (A) Principal Coordinates Analysis (PCoA) plot of the first two components obtained from 7 K SNP iSelect for 144



tetraploid wheat accessions plotted using the subspecies (A1) and the geographical origin (A2). (B) Unrooted bayesian tree of 7 K SNP iSelect using the bootstrap method with a replication of 1000 times. (C) STRUCTURE bar plot for K = 2 based on 7 K SNP iSelect genotyping data. Q value represents proportion of ancestry to a given subpopulation.

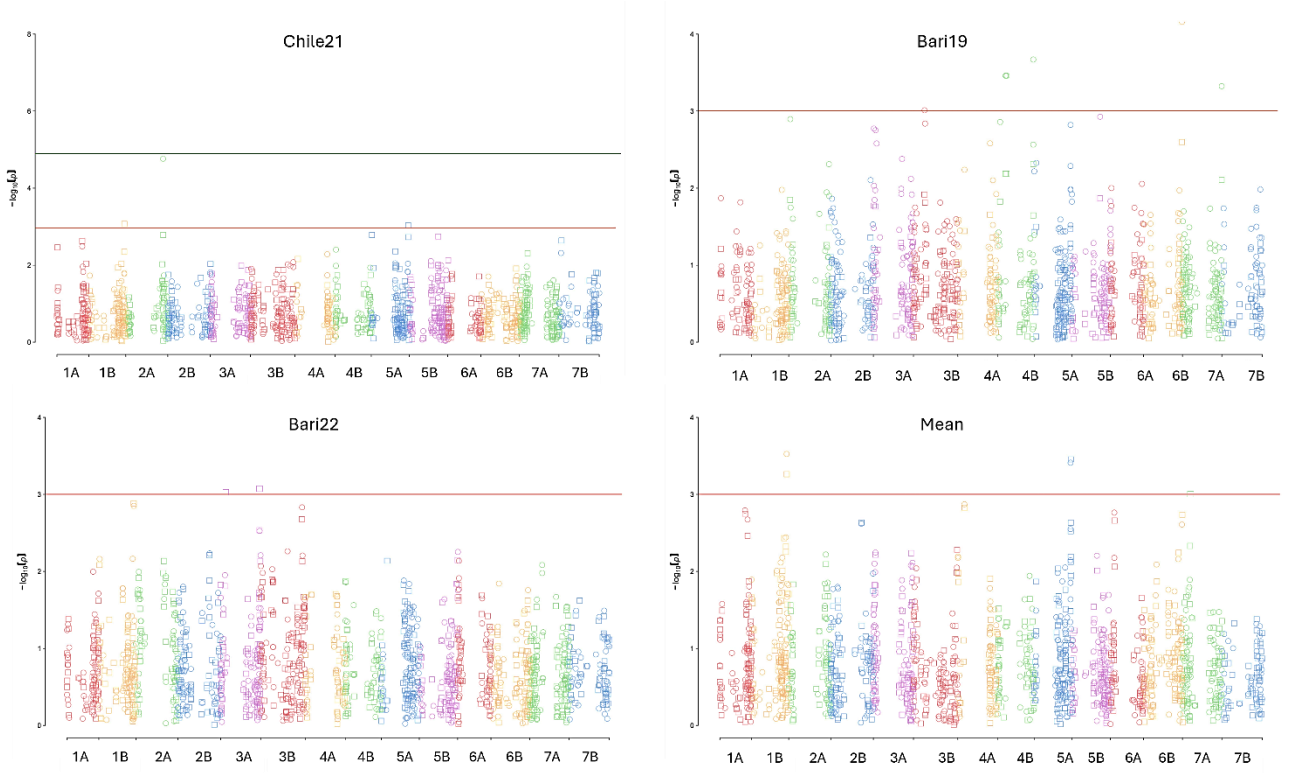


Figure 3. Manhattan plot of grain protein content from GWAS using the Mixed Linear Model (square symbol) and the BLINK model (circle symbol). The -log10 (p-values) from the GWAS are plotted according to the genetic position of the SNP markers on each of the 7 wheat chromosome pairs. In image with the doble threshold line the green one is for MLM model while the red one is for BLINK.

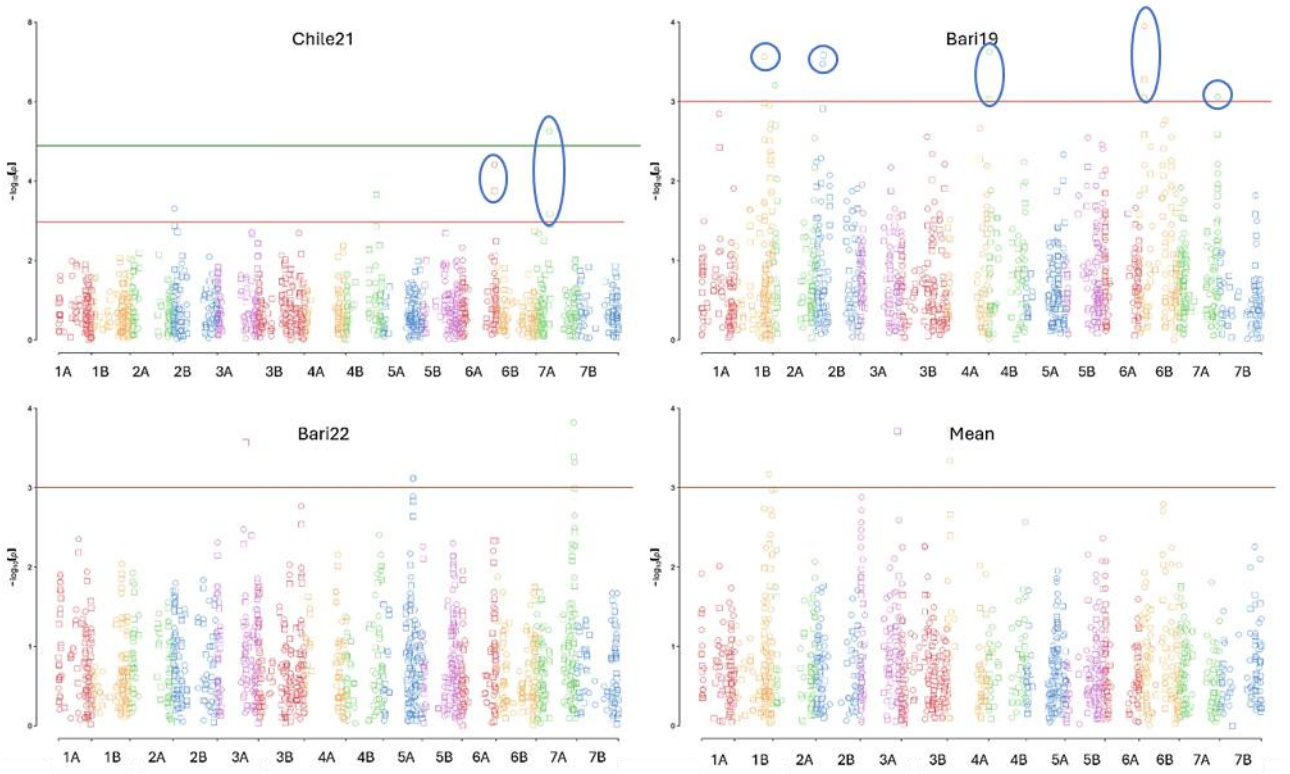


Figure 4. Manhattan plot of yellow index from GWAS using the Mixed Linear Model (square symbol) and the BLINK model (circle symbol). The -log10 (p-values) from the GWAS are plotted according to the genetic position of the SNP markers on each of the 7 wheat chromosome pairs. In image with the doble threshold line the green one is for MLM model while the red one is for BLINK. Blue circle highlighted the regions identified by both the methods.

CONCLUSION

The studies presented here highlight the growing importance of genetic improvement in wheat cultivation, with a particular focus on wild relatives and their potential impact on cereal nutritional quality. The *CslF6* gene from *Aegilops caudata*, as explored in the first study, reveals significant variations in amino acid sequences that modulate enzymatic activity and β-glucan biosynthesis, a crucial polysaccharide for human health. The observed variations in the DP3:DP4 ratio indicate how such genetic differences may influence the final β-glucan content in crops, paving the way for applications in crop improvement programs aimed at optimizing the nutritional value of cereals. The second study, focusing on β-glucan accumulation and *CslF6* gene expression, demonstrated that increased β-glucan content in wheat varieties through hybridization with wild relatives can significantly enhance the nutritional profile, suggesting that even small increases in β-glucan content can have a substantial impact on dietary quality. The use of advanced plant glycomics tools, such as HPAEC-PAD and GC-MS, emerges as a powerful approach to enhance the efficiency of selecting high β-glucan wheat relatives and hybrid lines. Finally, the third study underscores the importance of genomic tools in improving both grain protein content (GPC) and grain yield (GY) in wheat, opening new avenues for developing high-performance varieties that address both nutritional and market needs. The integration of genomic knowledge and the characterization of genes involved in biochemical pathways is critical to addressing the challenges of sustainable agriculture and wheat nutritional quality.

In summary, the combination of genomic approaches, comparative genomics, and traditional breeding holds great promise for revolutionizing crop improvement programs, offering innovative solutions for more sustainable and nutritionally enriched agricultural production.