



UNIVERSIDAD DE CÓRDOBA



Instituto de Agricultura Sostenible

UNIVERSIDAD DE CÓRDOBA

**Instituto de Agricultura Sostenible, Consejo Superior de
Investigaciones Científicas**

Programa de doctorado

**Ingeniería Agraria, Alimentaria, Forestal y de Desarrollo Rural
Sostenible**

TESIS DOCTORAL

**MECANISMOS GENÉTICOS Y MOLECULARES IMPLICADOS EN LA
COMPENSACIÓN PROTEICA EN EL GRANO DE TRIGO:
IMPLICACIONES PARA LA OBTENCIÓN DE VARIEDADES CON BAJO
CONTENIDO EN GLUTEN INMUNOGÉNICO**

**GENETIC AND MOLECULAR MECHANISMS INVOLVED IN PROTEIN
COMPENSATION IN WHEAT GRAIN: IMPLICATIONS FOR OBTAINING
VARIETIES WITH LOW-IMMUNOGENIC GLUTEN**

Autora

MIRIAM MARÍN SANZ

Director

Dr. Francisco Barro Losada

Departamento de Mejora Genética Vegetal, Instituto de Agricultura Sostenible,
Consejo Superior de Investigaciones Científicas (CSIC), España.

Córdoba, Marzo de 2023

TITULO: *Genetic and molecular mechanisms involved in protein compensation in wheat grain: implications for obtaining varieties with low-immunogenic gluten*

AUTOR: *Miriam Marín Sanz*

© Edita: UCOPress. 2023
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

<https://www.uco.es/ucopress/index.php/es/ucopress@uco.es>

“Sueña el rico en su riqueza,
que más cuidados le ofrece;
sueña el pobre que padece
su miseria y su pobreza;
sueña el que a medrar empieza,
sueña el que afana y pretende,
sueña el que agravia y ofende,
y en el mundo, en conclusión,
todos sueñan lo que son,
aunque ninguno lo entiende.”

Calderón de la Barca, La vida es Sueño.



**TÍTULO DE LA TESIS:**

Mecanismos genéticos y moleculares implicados en la compensación proteica en el grano de trigo: implicaciones para la obtención de variedades con bajo contenido en gluten inmunogénico

DOCTORANDO/A:

Miriam Marín Sanz

INFORME RAZONADO DEL/DE LOS DIRECTOR/ES DE LA TESIS

(se hará mención a la evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma).

Francisco Barro Losada, Profesor de Investigación del Instituto de Agricultura Sostenible (IAS), perteneciente al CSIC, como director de la tesis doctoral con título "Mecanismos genéticos y moleculares implicados en la compensación proteica en el grano de trigo: implicaciones para la obtención de variedades con bajo contenido en gluten inmunogénico" realizada por Miriam Marín Sanz,

INFORMA

Que dicha tesis ha sido realizada bajo mi dirección. El objetivo principal de la tesis ha sido la identificación de los mecanismos genéticos y moleculares que actúan en el grano de trigo como consecuencia del silenciamiento de las gliadinas mediante técnicas ARNi y/o CRISPR/Cas.

Los resultados y conclusiones obtenidas son de gran relevancia para comprender los mecanismos compensatorios que actúan en el grano de trigo tras el silenciamiento de las gliadinas. Estos resultados son de gran relevancia ya que permitirán controlar dichos procesos regulatorios utilizando técnicas ARNi y/o CRISPR/Cas, y permitirán el desarrollo de nuevas variedades con propiedades nutricionales específicas, incluyendo variedades no inmunogénicas para colectivos con intolerancia o sensibilidad al gluten.

En su etapa pre-doctoral, Miriam Marín Sanz ha realizado una estancia de 3 meses, desde el 14 de Julio hasta el 15 de Octubre de 2022, en el departamento de Crop Genetics del John Innes Centre (Norwich, Reino Unido), bajo la dirección del Profesor Cristobal Uauy. El objetivo principal de la estancia fue adquirir nuevas habilidades y conocimientos para el análisis de expresión génica en líneas de trigo con las alpha-gliadinas mutadas mediante CRISPR/Cas y ARNi. Los resultados obtenidos nos han permitido comparar ambas tecnologías y cuantificar su efecto sobre el transcriptoma del grano de trigo.

La doctoranda, además, ha asistido a congresos y colaborado en varias líneas de investigación de las que derivan las publicaciones/aportaciones siguientes:

Publicaciones:

Marín-Sanz, M., Giménez, M. J., Barro, F., & Savin, R. (2020) Prolamin Content and Grain Weight in RNAi Silenced Wheat Lines Under Different Conditions of Temperature and Nitrogen Availability. *Front. Plant Sci.* 11:314.

Guzmán-López, M. H., **Marín-Sanz, M.** (co-first author), Sánchez-León, S., & Barro, F. (2021). A Bioinformatic Workflow for InDel Analysis in the Wheat Multi-Copy α -Gliadin Gene Family Engineered with CRISPR/Cas9. *International Journal of Molecular Sciences*, 22(23), 13076.

Guzmán-López, M. H., Sánchez-León, S., **Marín-Sanz, M.**, Comino, I., Segura, V., Vaquero, L., Rivero-Lezcano, O. M., Pastor, J., Sousa, C., Vivas, S., & Barro, F. (2021). Oral Consumption of Bread from an RNAi Wheat Line with Strongly Silenced Gliadins Elicits No Immunogenic Response in a Pilot Study with Celiac Disease Patients. *Nutrients* 2021, 13, 4548.

Marín-Sanz, M., Iehisa, J. C. M., & Barro, F. (2022). New transcriptomic insights in two RNAi wheat lines with the gliadins strongly down-regulated by two endosperm specific promoters. *The Crop Journal*, 10(1), 194-203.

Haro, C., Guzmán-López, M. H., **Marín-Sanz, M.**, Sánchez-León, S., Vaquero, L., Pastor, J., Comino, I., Sousa, C., Vivas, S., Landa, B. B., & Barro, F. (2022). Consumption of Triticum Bread Reduces Immunogenic Gluten Intake without Altering the Gut Microbiota. *Foods*, 11(10), 1439.

Marín-Sanz, M., Sánchez-León, S., León, E., & Barro, F. (2022). Comparative characterization of the gluten and fructan contents of breads from industrial and artisan bakeries: a study of food products in the Spanish market. *Food & Nutrition Research*, 66.

Marín-Sanz, M., & Barro, F. (2022). RNAi silencing of wheat gliadins alters the network of transcription factors that regulate the synthesis of seed storage proteins toward maintaining grain protein levels. *Frontiers in plant science*, 13.

Chaouachi, C., **Marín-Sanz, M.**, Kthiri, Z., Boukef, S., Harbaoui, K., Barro, F., & Karmous, C. (2023). Opportunity of using durum wheat landraces to tolerate 2 drought stress: screening morpho-physiological components. Under revision in *AoB PLANTS*.

Marín-Sanz, M., Barro, F., & Sánchez-León, S. (2023). Unraveling the celiac disease-related immunogenic complexes in a set of wheat and tritordeum genotypes: implications for low-gluten precision breeding in cereal crops. Under revision in *Frontiers in plant science*.

Aportaciones a congresos:

Marín-Sanz, M., Giménez, M. J., Barro, F., & Savin, R. Contenido de prolamina y peso del grano de líneas de trigo RNAi con bajo contenido en gluten bajo diferentes condiciones de temperatura y nitrógeno. VII Congress of the Spanish Celiac Disease Society, 2021.

Actividades de doctorado:

Asistencia al "VI Congreso Nacional de la Sociedad Española de Enfermedad Celíaca (SEEC)", celebrado en Córdoba en 2018.

Visita a la Estación Experimental del Zaidín organizado por la Universidad de Córdoba, el 29 de marzo de 2019 en Granada.

Organización del seminario de actualidad "LA TESIS DOCTORAL Y LAS OPORTUNIDADES EN MATERIA DE TRANSFERENCIA TECNOLÓGICA" el 16 de mayo de 2019 en la Universidad de Córdoba.

Asistencia al seminario de actualidad "IX ENCUENTRO DE ESTUDIANTES DE DOCTORADO" del IAS-CSIC el 14 de Junio de 2019.

Participación en el taller de mejora de la empleabilidad realizado el 16 de Octubre de 2019 en la Universidad de Córdoba.

Asistencia a los seminarios de actualidad "EL DOCTORADO EN LA UNIVERSIDAD DE CÓRDOBA, MARCO NORMATIVO, PROCESOS Y PROCEDIMIENTOS, 2019" el 13 de Diciembre en el Rectorado de la Universidad de Córdoba.

Asistencia al seminario de actualidad "X ENCUENTRO DE ESTUDIANTES DE DOCTORADO" en el IAS-CSIC el 29 de Octubre de 2020.

Realización de una estancia internacional de tres meses en el John Innes Centre, Norwich, Reino Unido, del 14 de Julio al 15 de Octubre de 2022.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 23 de Marzo de 2023

Firma del/de los director/es

Fdo.: Francisco Barro Losada

**DOCTORANDA/O**

Miriam Marín Sanz

TÍTULO DE LA TESIS:

Mecanismos genéticos y moleculares implicados en la compensación proteica en el grano de trigo: implicaciones para la obtención de variedades con bajo contenido en gluten inmunogénico

INFORME RAZONADO DE LA TUTORA/OR**(Ratificando el informe favorable del director. Sólo cuando el director no pertenezca a la Universidad de Córdoba)**

El trabajo realizado por Miriam Marín Sanz, tanto en el aspecto formativo como investigador, ha sido excelente, cubriendo con creces los objetivos propuestos en el Programa. Quiero destacar el elevado número de publicaciones conseguidas en revistas Q1, (4 como primer autor firmante), la asistencia a Congresos y Reuniones Científicas y, sobre todo, su estancia en el John Innes Research Centre. El óptimo grado de formación está avalado por el informe de su Director, el Prf. Francisco Barro, y yo, como tutor, a través de las reuniones mantenidas con la Doctoranda, lo he podido constatar.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, a 23 de marzo de 2023

La/el Tutor/a

Fdo.: Jesús Valentín Jorrín Novo

Agradecimientos

La única certeza que considero es que no soy la misma persona que empezó la tesis doctoral. Durante estos años he crecido tanto profesional como personalmente, y esto es gracias a todos aquellos que han contribuido tanto directa como indirectamente en el presente trabajo.

En primer lugar, quiero agradecer a Francisco Barro su labor como director y consejero. He aprendido mucho de él durante estos años. La línea de investigación en la que tanto trabajo ha invertido no es fácil de abordar y, tristemente, su potencial se ve opacado bajo regulaciones inflexibles y tajantes, y mentes más inflexibles aún. Sin embargo, esto no es suficiente para rebajar el entusiasmo con el que afronta el trabajo. Este entusiasmo es, desde luego, contagioso. Gracias por no desistir en el empeño del trigo sin gluten que, estoy segura, algún día veremos en nuestra despensa.

A mis compañeros de laboratorio, Ana Adela, Ana Martínez, Marta, Cande, Luca y José, quiero agradecerles hacerme sentir en familia en el trabajo. A Susana, por su apoyo y su ayuda en el laboratorio, y por estar siempre cuidando de todos. A Helena, con quien entré en el IAS por primera vez, y que ha sido mi mayor consuelo cuando las cosas se ponían difíciles. Quiero agradecer a Latifa el que decidiese venir hasta Córdoba desde tan lejos, y que, cuando fui yo la que estuvo lejos, me hiciese sentir como en casa. A María José quiero agradecerle tantas cosas que me limitaré a decir lo excepcional que me pareció cuando la conocí y lo afortunada que me siento de haber estado ahí para verlo. Al personal del IAS, a sus maravillosos técnicos, conserjes y doctorandos. Para no escribir un desfile de nombres, me limitaré a mencionar a Manu, Luis y Martina, los cuatro comenzamos a la vez esta etapa que ahora se cierra para dar paso a nuevas experiencias. A Wohor, la humildad hecha persona. Y a Álvaro, con quien hablar resulta siempre fácil.

Quiero agradecer a Jesús Jorrín su labor como tutor en la presente tesis y el estar disponible siempre que lo he necesitado.

Quiero agradecer a Cristóbal el acogerme en su grupo durante la estancia y poder trabajar con su asombroso equipo. En especial a Matilde, Ruth y Lucía, por convertir la ciudad en un lugar mucho más divertido. Quiero agradecer también el haber descubierto un oasis hecho ciudad, un paraíso construido de historias viejas que tantas lecciones tienen para los novatos como yo.

A mis amigos de la universidad, del colegio, a los que conocí por casualidad y a los que hace mucho que no veo, quiero agradecerles su cariño y su apoyo en esta y cualquier otra etapa de mi vida. A Antonio T. y Marina, por su cariño y por su humor, un soplo de aire fresco en un mundo hecho de cristal. Quiero agradecer por su paciencia infinita a los 'informáticos', incomprensiblemente incomprensidos y que, aun así, nunca lo dan por perdido. Quiero agradecer a mi padre el que tuviese un piano en casa. A los músicos el que me enseñasen a entenderles. Y a la música misma, que soñase con ella.



Y ahora, me permitiré hacer alarde de mi familia, la mejor familia en la que me podría haber tocado vivir. A mi madre, la persona más dulce del planeta y que deja huella en cada corazón que la conoce. A mi padre, seguramente la persona más valiente e inteligente que encontraré nunca, si él lo permitiera, gobernaría el mundo. A mi hermano Pipón, de naturaleza autodidacta y capaz de sobresalir en todo lo que se propone; y Tomás, el hermano pequeño con el corazón más grande que haya visto. A toda mi familia quiero agradecerles, sinceramente, los innumerables buenos momentos que hemos vivido juntos. En especial a mi tía May, nuestra doctora, cuyo consejo ha sido y es muy valioso para mí. Quiero agradecer de todo corazón el cariño incondicional con el que mi abuela Mercedes nos trata a todos en la familia, y a mis abuelos José Luis, Antonio y Lourdes, el que nos hayan querido hasta el último de sus días.

Quiero agradecer a la buena fortuna el que pusiera a Rafa en mi camino. No tengo ni tendré lo suficiente para saldar la deuda con ella. Gracias Rafa por ser lo que mi corazón espera. Gracias por crecer juntos. Gracias también por tu familia, de la que me siento parte.

Por último, estoy agradecida con la razón de empezar la tesis y de mis continuos quebraderos de cabeza. Lejos de ser una debilidad, me has convertido en lo que soy ahora.



GENERAL INDEX

SUMMARY	13
RESUMEN	15
CHAPTER I	19
<i>General introduction and objectives</i>	
1. Wheat.....	20
2. Wheat grain composition.....	21
3. Wheat-related pathologies.....	24
3.1. CD epitopes and immunogenic peptides.....	26
4. Biotechnological approaches for obtaining low-gluten wheat.....	27
4.1. RNA of interference.....	28
4.2. CRISPR/Cas9 gene editing.....	28
5. Grain protein compensation after gliadin silencing.....	29
6. Objectives.....	32
7. References.....	33
CHAPTER II	39
<i>Prolamin Content and Grain Weight in RNAi Silenced Wheat Lines Under Different Conditions of Temperature and Nitrogen Availability</i>	
Abstract.....	40
Resumen.....	41
1. Introduction.....	42
2. Materials and Methods.....	45
2.1. Plant Material, Chamber Experiment and Treatments.....	45
2.2. Grain Weight and Total Protein Determination.....	45
2.3. Prolamins Quantification by RP-HPLC.....	45
2.4. Non-gluten Proteins (NGPs) Determination.....	46
2.5. Data Processing.....	46
2.6. Statistical Analysis.....	46
3. Results.....	47
3.1. Heat Stress and Nitrogen Treatment Effects on Grain Weight and Total Protein.....	47
3.2. Heat Stress and Nitrogen Treatment Effects on Gliadins and Glutenins.....	48
4. Discussion.....	51



5. Conclusion.....	53
6. References.....	54
7. Supplementary Materials	59
CHAPTER III	64
<i>New transcriptomic insights in two RNAi wheat lines with the gliadins strongly down-regulated by two endosperm specific promoters</i>	
Abstract.....	65
Resumen	66
1. Introduction	67
2. Material and Methods	68
2.1. Plant material.....	68
2.2. Prolamins determination	68
2.3. Total protein, NGPs, gluten content by ELISA, fructans and starch determination	68
2.4. RNA extraction.....	69
2.5. cDNA synthesis and qPCR experiments	69
2.6. RNA-seq data and bioinformatic analysis.....	69
2.7. Functional annotation.....	70
3. Results.....	70
3.1. Grain composition, quality and agronomic features of RNAi lines	70
3.2. Transcriptomic analysis of grains from RNAi lines.....	72
3.3. Expression of prolamins and other grain protein genes	74
3.4. Expression of genes related to stress response in RNAi lines	75
3.5. Expression of genes related to carbohydrate metabolism and other metabolic process in RNAi lines	77
4. Discussion.....	78
5. Conclusions.....	81
6. References.....	82
7. Supplementary Materials	87
CHAPTER IV	101
<i>RNAi silencing of wheat gliadins alters the network of transcription factors that regulate the synthesis of seed storage proteins toward maintaining grain protein levels</i>	
Abstract.....	102
Resumen	103



1. Introduction	104
2. Materials and Methods.....	106
2.1. Plant material and experimental design.....	106
2.2. RNA extraction and sequencing.....	106
2.3. RNA-seq data analysis	106
2.3.1. Trimminig.....	106
2.3.2. Mapping.....	107
2.3.3. Differential gene expression (DGE) analysis	107
2.3.4. Gene Ontology (GO) and pathway enrichment analyzes.....	107
2.4. Seed-borne allergens and immune-responsive proteins	107
2.5. Gene Network Inference with Ensemble of trees (GENIE3) and Weighted Gene Co-expression Network Analysis (WGCNA) networks in bread wheat	107
2.6. Gene expression by quantitative real-time PCR (qPCR).....	108
2.7. Prolamin quantification by Reverse Phase-High Performance Liquid Chromatography (RP-HPLC).....	108
2.8. Total protein, thousand kernel weight, and starch content	108
2.9. Proteomic data analysis	108
3. Results.....	109
3.1. Grain proteins are strongly affected by the RNAi	110
3.2. Starch synthesis-related genes are also affected in the RNAi line	111
3.3. Candidate TFs for the grain protein and starch synthesis regulation in the E82 RNAi line	112
3.4. The source-sink communication during grain filling is greatly affected by RNAi silencing	115
3.4.1. Proteases and protease inhibitors in the leaf of E82	116
3.4.2. Sugar and amino acid transporters in the leaf and grain of E82.....	117
4. Discussion.....	118
5. References.....	124
6. Supplementary Materials	130
CHAPTER V	160
<i>Unraveling the celiac disease-related immunogenic complexes in a set of wheat and tritordeum genotypes: implications for low-gluten precision breeding in cereal crops</i>	
Abstract.....	161



Resumen	162
1. Introduction	163
2. Materials and Methods.....	165
2.1. Plant material	165
2.2. DNA extraction, PCR and NGS sequencing.....	165
2.3. Amplicon analysis	166
2.3.1. Epitope search and amplicon types	167
2.4. Ranking of cereal lines' immunogenic potential by Analytic Hierarchy Process (AHP).....	168
3. Results.....	168
3.1. Epitope abundance in the set of cereal genotypes	169
3.2. Gliadin amplicons are grouped into different types depending on the number of CD epitopes.....	170
3.3. Clustering of wheat and tritordeum lines based on CD epitopes profile.....	173
3.4. Identification of CD alpha-gliadin epitopes with one mismatch.....	177
4. Discussion.....	177
5. Acknowledgements	182
6. References.....	183
7. Supplementary Materials.....	187
CHAPTER VI	226
<i>Conclusions</i>	



Summary

Prolamins, or gluten proteins, play an important role in the bread-making quality of wheat. These proteins are divided into two main fractions: gliadins and glutenins. In addition to their viscoelastic properties, gluten proteins are responsible for triggering the autoimmune-response in patients suffering from celiac disease (CD), a chronic enteropathy given in genetically predisposed individuals that affects 0.7% of the global population. The only available treatment is to follow a life-long gluten-free diet, which is difficult due to the large number of products that are based on wheat or present it as an additive. Biotechnological tools have been applied to obtain wheat varieties with low-gluten content. RNA of interference (RNAi) has been one of those, providing the silencing of the three groups of gliadins: ω -, α - and γ -gliadins.

The synthesis of gluten proteins is influenced by the environment: among the factors that affect the grain composition, temperature and nitrogen (N) stand out, especially during the grain-filling stage. In the **second chapter** of this thesis, gliadins and glutenins from four RNAi lines with low-gluten content have been studied under heat stress and two nitrogen (N) treatments. Quantification was performed by reverse-phase high-performance liquid chromatography (RP-HPLC). In addition to the effect of heat stress decreasing grain weight in all the lines tested, an increase in the prolamin content was observed in the control lines (without the silencing fragment) due to temperature. Except for ω -gliadins, the rest of the prolamins of the RNAi lines did not present any effect on their content in the environmental conditions tested, while the proteins of the control lines were affected. This indicates great stability of the silencing character in the RNAi lines, which gives them value in the context of increased temperature and heat stress events, and environments with different N availability. In this work, as well as in previous ones, the RNAi lines showed a total grain N content similar to that of the control lines, despite the great decrease in the gliadin content. This phenomenon of compensation or readjustment of the grain protein composition gave rise to the next two chapters.

The **third chapter** presents the results of gene expression by RT-qPCR and RNA-seq from two RNAi lines with two different endosperm promoters driving the silencing fragment. In both lines, the compensation was produced by the increase in the non-gluten proteins (NGPs) and the high molecular weight (HMW) subunits of glutenins. The increase in the expression of the genes that encode these subunits is already notable in the RNAi lines at 18 days after anthesis (DAA), as well as the transcription factors (TFs) related to the regulation of prolamins synthesis. In addition, the expression of genes encoding NGPs (globulins, serpins, ns-LTPs, and triticins) was also higher in the RNAi lines. Moreover, other changes at the transcriptome level were observed in the RNAi lines, particularly those related to enzyme regulation activity, carbohydrate metabolism, and stress response. Regarding the latter, many of the genes differentially expressed in the RNAi lines involved in this process are regulated by the abscisic acid (ABA), which could suggest the involvement of this phytohormone in the stress response observed at the expression level in the RNAi lines.



The transcriptome of another RNAi line with all three gliadin fractions strongly silenced was analyzed in the **fourth chapter**. The RNAi E82 line was chosen for the RNA-Seq analysis, and samples were taken from two tissues: the grain and the leaf in an intermediate stage of grain filling. The E82 line is one of the RNAi lines with the highest gluten reduction (ppm) and was used for two clinical trials with CD and non-celiac wheat sensitivity (NCWS) patients. In this work, the focus was mainly on the regulatory elements at the transcriptional level, thus obtaining a network of candidate TFs for regulating the synthesis of seed storage proteins, including prolamins, and starch in the RNAi lines. Both the genes encoding TFs of the network, as well as those of the prolamins and other storage proteins, were down-regulated in the E82 line. The results obtained in the leaf of the RNAi line showed a large number of down-regulated genes related to protease activity, carbohydrate and amino-acid metabolisms and their transport. Regarding the last one, it should be noted that specific proline transporters were down-regulated in the E82 line, while those of lysine-histidine were up-regulated. In addition to the lower gluten content, this E82 line differs from those analyzed in the previous chapter in that grain protein compensation is mainly by the NGPs, specifically the lysine-rich globulins but not the HMWs. The genes encoding globulins were not down-regulated in the RNAi line nor did they show any relationship with the TFs network, suggesting that they are regulated independently from the rest of the seed storage proteins. Both the TFs network described, and the information related to globulins and leaf processes, contribute to increasing the knowledge of the regulation of prolamins and protein compensation mechanisms in this RNAi line.

In addition to RNAi, CRISPR/Cas technology has been used to obtain lines with low-gluten content. Due to the genomic complexity of the prolamins coding regions and the fact that not all gluten proteins are equally immunogenic, the characterization of these genomic regions is important for designing precision breeding programs, both based on traditional and new plant breeding techniques like CRISPR/Cas. With this objective the **fifth chapter** was carried out, in which the amplicons of two CD immunogenic gliadin complexes were sequenced in a set of lines of bread wheat, durum wheat and tritordeum, including wheat lines with the 1BL/1RS rye translocation. Variability for CD immunogenic epitopes and peptides was studied, and their abundances were quantified in amplicons of α - and γ -gliadins, including also 40k- γ -secalins. The results from this chapter allowed us to identify and classify the lines with the highest and lowest immunogenic potential, as well as to quantify the abundance of CD-related epitopes, including 33-mer, one of the most immunodominant peptides in CD. These results are of great importance to understanding the structural complexity of the immunogenic complexes in wheat, and to select wheat varieties for subsequent efficient transformation by gene editing with CRISPR/Cas.



Resumen

Las prolaminas, o proteínas del gluten, tienen un papel fundamental en la calidad harino-panadera del trigo. Estas proteínas se dividen en dos grandes fracciones: gliadinas y gluteninas. Además de sus propiedades viscoelásticas, las proteínas del gluten son las responsables de desencadenar la respuesta autoinmune en pacientes con enfermedad celíaca (EC), una enteropatía crónica en individuos genéticamente predispuestos que afecta al 0.7% de la población mundial. El único tratamiento posible es seguir una dieta libre de gluten de por vida, lo que resulta difícil debido a la gran cantidad de productos basados en trigo o que lo presentan como aditivo. La aplicación de herramientas biotecnológicas ha permitido obtener variedades de trigo con bajo contenido en gluten. El ARN de interferencia (ARNi) ha sido una de las utilizadas, teniendo como objetivo las tres fracciones de gliadinas: ω -, α - y γ .

La síntesis de estas proteínas está influenciada por el ambiente: entre los factores que afectan a la composición del grano, destacan la temperatura y la disponibilidad de nitrógeno (N), especialmente durante el estadio de llenado del grano. En el **segundo capítulo** de este trabajo, gliadinas y gluteninas de cuatro líneas ARNi con bajo contenido en gluten han sido cuantificadas bajo estrés térmico y dos tratamientos de N. La cuantificación se realizó mediante cromatografía líquida de alta resolución en fase reversa (RP-HPLC, de sus siglas en inglés). Además del efecto del estrés térmico disminuyendo el peso del grano en todas las líneas ensayadas, se observó un incremento de las prolaminas en las líneas control (sin el fragmento de silenciamiento) debido a la temperatura. A excepción de las ω -gliadinas, el resto de las prolaminas de las líneas ARNi no se vieron afectadas por los cambios ambientales ensayados, mientras que estos sí influyeron en las proteínas de las líneas control. Esto indica una gran estabilidad del carácter de silenciamiento en las líneas ARNi, lo que las pone en valor en un contexto de cambio climático con incremento de temperatura y eventos de estrés térmico, y ambientes con distinta disponibilidad de N. En este trabajo, así como en otros anteriores, las líneas ARNi presentaron un contenido de N total en el grano similar al de las líneas control, pese a la gran disminución en el contenido de gliadinas. Este fenómeno de compensación o reajuste de la composición proteica del grano dio lugar a los dos siguientes capítulos.

En el **tercer capítulo** se exponen los resultados del análisis de la expresión génica mediante RT-qPCR y RNA-seq de dos líneas ARNi con silenciamiento de gliadinas donde el fragmento de silenciamiento estaba bajo el control de dos promotores de endospermo distintos. En ambas líneas, la compensación se produjo por el incremento de las proteínas no pertenecientes al gluten (NGPs, de sus siglas en inglés) y de las subunidades de gluteninas de alto peso molecular (HMW, de sus siglas en inglés). El incremento de la expresión de los genes que codifican estas subunidades ya es notable en las líneas ARNi a los 18 días después de la antesis (DAA, de sus siglas en inglés), al igual que factores de transcripción (FTs) relacionados con la regulación de la síntesis de las prolaminas. Además, la expresión de los genes que codifican para las NGPs (globulinas, serpinas, ns-LTPs y triticinas) también fue mayor en las líneas con



silenciamiento de gliadinas. También se observaron otros cambios a nivel de transcriptoma en las líneas ARNi, particularmente los relacionados con la regulación de la actividad enzimática, el metabolismo de los carbohidratos y la respuesta a estrés. En relación con este último, muchos de los genes implicados en este proceso, y diferencialmente expresados en las líneas ARNi, están regulados por el ácido abscísico (ABA), lo que podría sugerir la implicación de esta fitohormona en la respuesta a estrés observada a nivel de expresión en las líneas ARNi.

El transcriptoma de otra línea ARNi con fuerte silenciamiento de todas las fracciones de gliadinas fue analizada en el **cuarto capítulo**. En esta ocasión se escogió la línea E82 para el análisis de expresión diferencial y se tomaron muestras de dos tejidos: el grano y la hoja en un estadio intermedio del llenado del grano. La línea E82 es una de las líneas ARNi con mayor reducción de gluten (ppm) y que, además, ha sido utilizada en ensayos clínicos con pacientes celíacos y sensibles al trigo. En este trabajo, el foco se puso principalmente en los elementos reguladores a nivel transcripcional, obteniéndose así una red de FTs candidatos a regular la síntesis de las proteínas de reserva de la semilla, incluyendo las prolaminas, y el almidón en el grano en las líneas ARNi. Tanto los genes codificantes de los FTs de la red, como los de las prolaminas y otras proteínas de reserva del grano, estuvieron subexpresados en la línea E82. Los resultados obtenidos en la hoja de la línea ARNi mostraron una gran cantidad de genes subexpresados relacionados con la actividad de las proteasas, el metabolismo de los amino-ácidos y de los azúcares, y su transporte. En relación a este último, cabe destacar que los transportadores específicos de prolina estuvieron subexpresados en la línea E82, mientras que los de lisina-histidina estuvieron sobreexpresados. Además del menor contenido en gluten, esta línea difiere de las analizadas en el capítulo anterior en que la compensación proteica en el grano se produce principalmente por las NGPs, concretamente las globulinas ricas en lisina, y no por las HMW. Los genes que codifican para las globulinas no estuvieron subexpresados en la línea ARNi ni presentaron relación alguna con la red de FTs, sugiriendo que son reguladas independientemente del resto de proteínas de reserva en el grano. Tanto la red de FTs descrita como la información relativa a las globulinas y los procesos en la hoja, contribuyen a incrementar el conocimiento relativo a la regulación de las prolaminas y la compensación proteica en las líneas con silenciamiento de las gliadinas.

Además del ARNi, se ha utilizado la tecnología CRISPR/Cas para la obtención de líneas con bajo contenido en gluten. Debido a la complejidad de las regiones genómicas que codifican los genes de prolaminas y a que no todo el gluten es igual de inmunogénico, la caracterización de estas regiones es importante para diseñar programas de mejora genética de precisión, tanto tradicionales como basados en las nuevas herramientas de mejora como CRISPR/Cas. Con este objetivo se realizó el **quinto capítulo**, donde los complejos inmunogénicos de las α - y γ -gliadinas fueron caracterizados mediante secuenciación masiva en un conjunto de líneas de trigo harinero, trigo duro y tritordeo, incluyendo líneas de trigo con la translocación de centeno 1BL/1RS. Se identificaron y cuantificaron los epítomos y péptidos



inmunogénicos relacionados con la EC en los amplicones de las α - y γ -gliadinas, incluyendo también las 40k- γ -secalinas. Los resultados de este capítulo permitieron conocer la complejidad estructural de estos complejos inmunogénicos en trigo, y clasificar las líneas con mayor y menor potencial inmunogénicos, así como cuantificar la abundancia de los epítomos relacionados con la EC, incluyendo el 33-mer, uno de los péptidos más inmunodominantes que se conocen. Estos resultados permitirán seleccionar variedades y su incorporación en programas de mejora genética clásica o su uso directo mediante edición genética con CRISPR/Cas.





CHAPTER I

General introduction and objectives



1. Wheat

Wheat was one of the first crops domesticated during the Agricultural Revolution 10,000 years ago. Since then, wheat became an essential part of human civilization as staple food in their diet (Feldman and Levy, 2015). The genus *Triticum* comprises diploid, tetraploid, and hexaploid species, derived by inter-specific or inter-generic hybridization followed by chromosome doubling. The allotetraploid wheat *Triticum turgidum ssp. durum* (AABB, $2n=4x=28$) contains the subgenomes A and B. The A subgenome comes from the diploid species *T. urartu* (A^uA^u), while the origin of the B subgenome remains unresolved, but genetic, biochemical, and morphological evidences suggest that *Aegilops speltoides* (BB) was the donor of this subgenome (de Sousa et al., 2021). The hybridization of *T. urartu* and *A. speltoides* gave rise to the wild emmer *T. turgidum ssp. dicoccoides*, the ancestor of domesticated emmer, from which some subspecies developed free-threshing grains, known as durum wheat (*T. turgidum ssp. durum*) (Figure 1.1 A). The allohexaploid wheat *T. aestivum* (AABBDD, $2n=6x=42$) contains the A and B subgenomes from the tetraploid wheat, and the D subgenome from *A. tauschii*. Both belong to the Triticeae tribe of the grass family Poaceae (Levy and Feldman, 2022; Pont et al., 2019) (Figure 1.1 A).

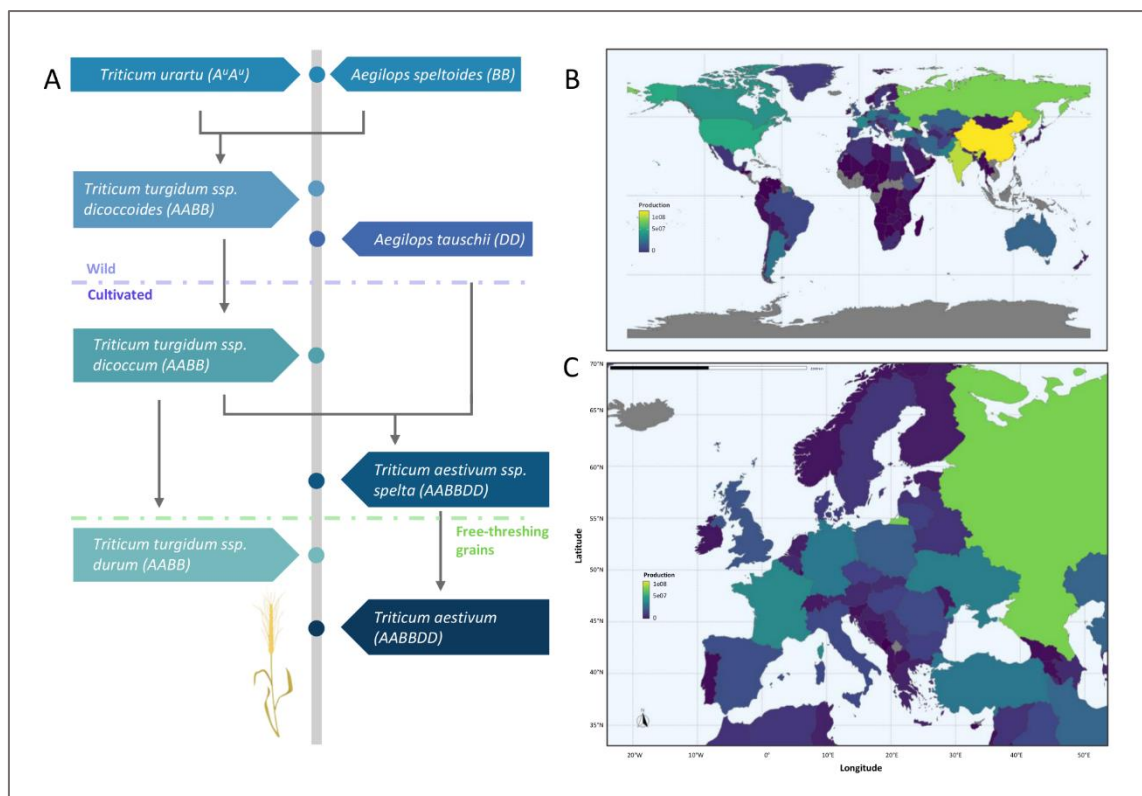


Figure 1.1. A) Phylogeny of wheat. B) World-wide and C) Europe production of wheat. The color code indicates the value of production (tons). The data showed in B and C were obtained by (FAO, 2022), and it reflects the production values for the 2020 year.



The wheat crop is cultivated all over the world, from areas as far north as Norway, Finland and Russia, such as southern Argentina (Figure 1.1 B,C). It is mainly cultivated in temperate regions, although it is also found in subtropical and tropical regions (Levy and Feldman, 2022). The 95% of the wheat grown globally is bread wheat (*T. aestivum*), and the remaining 5% is mainly durum wheat (*T. turgidum ssp. durum*). The production of wheat 'ancient' species like *T. monococcum* (einkorn, diploid), *T. turgidum ssp. dicoccum* (emmer, tetraploid), and *T. turgidum ssp. spelta* (spelt, hexaploid) comprises a small part of the remaining 5%. However, in the last decades these crops increased their popularity as they have been associated with health benefits, disease tolerance or adaptation to wide climate conditions (Longin et al., 2016). This is notable for spelt crops, as they present higher protein, lipid, nutrient, carbohydrates and vitamins contents than bread wheat (Kohajdová and Karovicova, 2007).

Durum wheat grows in semiarid environments (e.g. Mediterranean countries), associated more with social traditions than climate limitations. Canada is the main cultivator, however, most durum wheat consumers and production processes take place on the Mediterranean region (Xynias et al., 2020). This crop is used mainly for pasta production, as other traditional dishes like couscous (Elias, 1995). Bread wheat consumption has increased globally, even in countries with unfavourable climates for wheat production. It is in the top three of the most important crops in the world, reaching production values of 760 million tons of grain per year, only below maize (FAO, 2022). In Spain, in terms of production, it is ranked first, with values reaching 8.1 million tons per year (FAO, 2022) (Figure 1.1 B,C). While the main wheat export countries were Russia, the United States and Canada in 2020, the main wheat import countries were located in Asia and Eastern Europe (FAO, 2022). The unique properties of wheat doughs allow the production of products of daily consumption in the western lifestyle, such as bread, baked goods, noodles and pasta (Shewry and Hey, 2015). Although wheat is considered a primary source of energy (carbohydrates), it is also an important source of protein, in addition to dietary fiber, B vitamins, minerals, and other phytochemicals in the human diet (Shewry and Hey, 2015). All the essential amino acids, except lysine, are present in adequate amounts for an adult in the wheat grain. White flours have a lower proportion of essential amino acids compared to whole meal flours, which is related to the high content of seed storage proteins (SSPs) in the endosperm, rich in proline and glutamine and poor in lysine, contrasting with the proteins from other grain tissues rich in lysine (Shewry and Hey, 2015; Shewry, 2009).

2. Wheat grain composition

The wheat grain is composed of 2-3% germ, 13-17% bran and 80-85% endosperm. The germ, located in one end of the grain, is rich in proteins, lipids, minerals, and vitamin E. The bran is comprised of the outer layers, and it is rich in B vitamins and minerals. It contains a high proportion of fibre (53%). The aleurone layer, rich in proteins and enzymes, comprises the outer endosperm, while the inner part is referred to as the



starchy endosperm, the principal source of carbohydrates and a high source of proteins (Belderok et al., 2000; Cornell and Cauvain, 2003).

At harvesting, the wheat grain consists of 85% of carbohydrates, where 80% is starch present only in the starchy endosperm. Among carbohydrates, the wheat grain also contains fructans, composed of several fructose units bonded to a single glucose unit (Shewry and Hey, 2015). The importance of fructans has increased in recent years as they are one of the major components of the group of small fermentable carbohydrates named as FODMAPs (Fermentable, Oligo-, Di-, Mono-saccharides, and Polyols), which has been related to several human pathologies, including irritable bowel syndrome (IBS) and inflammatory bowel disease (Crohn's disease and ulcerative colitis) (Gibson and Shepherd, 2010).

The protein content of the grain is the major determinant of the bread-making quality of wheat (Shewry, 2009). Proteins represent 10-15% of the dry weight of the grain and they can be classified by their functions into storage proteins, metabolic proteins, protective proteins, and miscellaneous proteins with further specific functions (Shewry and Hey, 2015). However, wheat proteins have traditionally been classified into four fractions according to their solubility in albumins (soluble in water), globulins (soluble in dilute saline), prolamins (soluble in alcohol), and glutelins (insoluble in the other solvent but can be extracted in alkali). In the albumin and globulin fractions, regulatory proteins with metabolic function and protective proteins are included. These represent 20% of the total protein of the wheat grain (Wieser et al., 2023). The prolamins are named gliadins in wheat. Despite glutelins comprising proteins which are insoluble in the three first solvents, its major components are in fact prolamins subunits not extractable in alcohol because of their polymer nature, as they present inter-chain disulphide bonds. These proteins are called glutenins in wheat and now are classified as prolamins together with gliadins (Shewry, 2019), and both represent about 80% of the grain protein. Prolamins are also called gluten in wheat, and they are located in the grain endosperm (Shewry, 2009). In addition, the monomeric gliadins can be divided into three structural groups by their mobility in acid polyacrylamide gel electrophoresis (A-PAGE): ω -, α/β - and γ -gliadins. The α - and β -gliadins present highly similar amino acid sequences and are often grouped into a single group called α -gliadins. The glutenins are non-alcohol-soluble polymeric proteins linked by inter-chain disulphide bonds. They can be divided into High Molecular Weight (HMW) and Low Molecular Weight (LMW) glutenin subunits by their mobility in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 1.2). Prolamins are synthesized and deposited into protein bodies in the starchy endosperm cells during the grain-filling stage. Their synthesis begins approximately 10 days after anthesis (DAA) and ends 20 days later (Wieser et al., 2023). They constitute, together with starch, an important nutrient reservoir during germination (Shewry and Halford, 2002; Shewry, 2019).



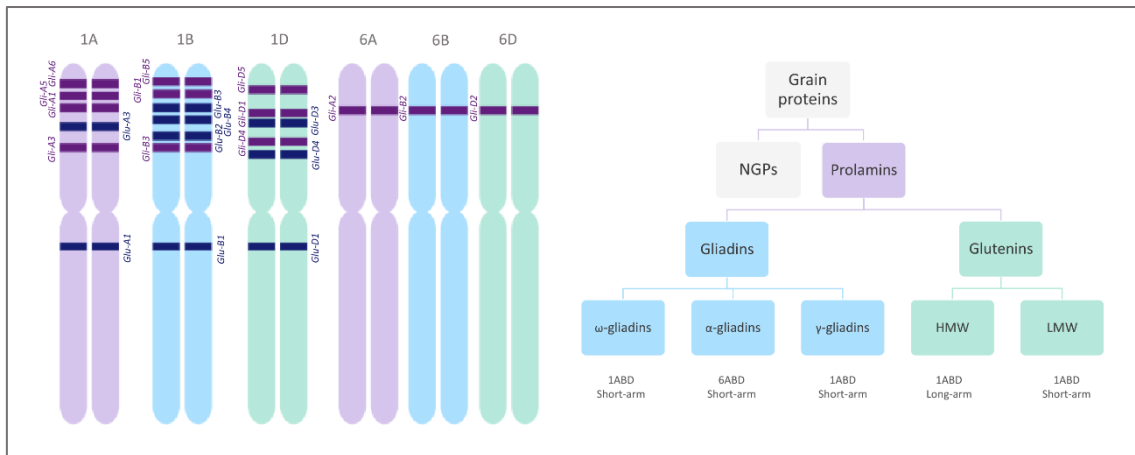


Figure 1.2. Gluten protein classification and their chromosome loci in bread wheat. At the left, the purple bands indicate loci of gliadin genes, and blue bands indicate the glutenin loci.

In wheat, the profile of protein distribution and their content is determined by genetic factors, but also by environmental ones, mainly the availability of nitrogen (N) fertilization and the timing of its application. Generally, the increase of N fertilization increases the protein content of the grain by regulating the expression levels of the genes involved in the synthesis of these storage proteins (Yu et al., 2017). The N application also affects the distribution of starch granules and the properties of starch (Xiong et al., 2014). The effect of high temperature and heat stress treatments during the grain filling phase reduces the starch content and affects the activity of the enzymes that catalyze its synthesis. Despite starch synthesis is more susceptible to heat stress than protein synthesis, this environmental factor also affects the accumulation of storage protein in wheat, increasing gliadins and decreasing glutenins (Triboï et al., 2003; DuPont and Altenbach, 2003).

Gliadins and glutenins are encoded by genes located in the groups 1 and 6 chromosomes from the A, B and D bread wheat subgenomes. Gliadin and LMW genes are located in the short arm of chromosomes 1A, 1B, 1D, 6A, 6B, and 6D, while the genes encoding for the HMW are located on the long arm of group 1 chromosomes (*Glu-1* loci) (Figure 1.2). The HMW subunits of glutenins are encoded by single genes named according to the subgenome, the subunit's size (x-type: higher molecular weight, or y-type: lower molecular weight), and its mobility on SDS-PAGE gels (numbered 1-12) in many cases (e.g. HMW 1Dx5) (Wieser et al., 2023). There are two closely linked genes encoding an x- and y-type subunit in each locus so, theoretically, there are six expected subunits in bread wheat, but not all the HMW genes are expressed in all genotypes, varying from three to five subunits some genes are silenced by transposon-like element insertion or nucleotide mutation in the coding/promoter region (Payne, 1987; Galili, 1997; Zu-Jun et al., 2006). In contrast, the LMW and the gliadins are encoded by gene families with multiple copies in the genome and arranged in tandem as a consequence of gene duplication, making difficult of obtaining high-quality gene sequences and analyzing them. Therefore, there is a high polymorphism in the gluten proteins among



wheat varieties, not only due to the variability of the protein sequences but also to the variability of the gene expression (Shewry, 2019). Unlike other species, bread wheat has a large and complex genome that contains more than 85% of repetitive DNA. In addition to its hexaploid nature, these traits make a challenge the task of assembly wheat genome. Few years ago, the high quality annotated genome of bread wheat was released (Appels et al., 2018). The availability of the wheat reference genome allowed also the genome-wide identification of multiple gene families, including genes coding for the seed-borne allergens as gluten proteins and other genes related to adverse reactions to wheat (Juhász et al., 2018). The identification of the wheat genes encoding prolamins and their chromosomal locations were achieved combining wheat proteins/peptides public databases and the reference genome sequence. In complex genomic regions, the short sequence reads are not enough to assembly them. Consequently, another approaches has been used to solve the problem of sequencing these genes: aligning PacBio-based sequence contigs with BioNano genome maps (Huo et al., 2018b; Huo et al., 2018a). With this approach, 47 α -gliadin genes (26 encoding intact full-length proteins), 14 γ -gliadin, 19 ω -gliadin and 17 LMW genes were reported in the model Chinese spring bread wheat cultivar.

Prolamin-like proteins, particularly gliadins, are also present in other cereal crops, being termed as hordeins in barley (*Hordeum vulgare*), secalins in rye (*Secale cereale*) and avenins in oat (*Avena sativa*). In addition, sequences encoding proteins with high homology to oat avenins, and named avenin-like proteins, have been identified in bread and durum wheat, and barley (Daly et al., 2020).

Wheat prolamins play a major role in the bread-making quality of wheat, being responsible for the viscoelastic properties of wheat flour. When water is added to flour during the bread-making process, gliadins and glutenins will form a cohesive viscoelastic protein network that will retain the gas released during the fermentation of the dough and will lead to the formation of a leavened, fluffy bread crumb (Delcour et al., 2012). In fact, the HMW has the greatest impact on dough quality (Payne et al., 1987). The bread-making properties of wheat gluten are not shared in other cereals crops, making wheat unique among cereals.

3. Wheat-related pathologies

Despite the importance of wheat gluten proteins, the consumption of wheat, and other gluten-containing cereals are associated with several adverse reactions in humans. The pathologies can be differentiated by their mechanism of activation in (i) HLA-associated Celiac Disease (CD), and (ii) IgE-mediated diseases such as wheat allergy, wheat-dependent exercise-induced anaphylaxis (WDEIA) or baker's asthma (Ludvigsson et al., 2013; Larré et al., 2011). There is a third pathology denoted as non-celiac wheat sensitivity (NCWS), whose causes and mechanisms are still unclear. In the first place, gluten was initially assigned as the causative agent of NCWS, however, the FODMAPs, particularly fructans, and α -amylase/trypsin inhibitors (ATIs) were later



proposed as possible triggering compounds in subsequent studies (Gibson et al., 2017; Catassi et al., 2013). The prevalence of NCWS ranges from 0.6% to 13%, due to cases of self-diagnoses and the absence of standardized diagnostic criteria or biomarkers (Aziz et al., 2016).

Among the adverse reactions to gluten proteins, CD is the most known pathology. CD is a chronic enteropathy in genetically predisposed individuals after the intake of gluten, which triggers both an innate and adaptive immune response. The innate immune response takes place in the epithelial component of the intestinal mucosa and, in presence of gluten, the production of cytokines interleukin-15 (IL-15) is increased. The intraepithelial lymphocytes will be differentiated into cytotoxic CD8⁺ T cells and, in the end, it will produce damage in the intestinal mucosa such as villous atrophy and crypt hyperplasia (Green et al., 2015). In contrast, the adaptive immune response requires antigen-presenting cells (APCs), such as dendritic cells and B cells, that will interact with CD4⁺ T cells. This will produce inflammatory cytokines such as interferon- γ (IFN- γ) and IL-21. For gluten to be recognized by the HLA-DQ2 and HLA-DQ8 molecules from the APCs, the glutamine residues present in gluten peptides are deamidated to glutamate by tissue transglutaminase 2 (tTG2) in the lamina propria of the small intestine. Then, the HLA-DQ2 and HLA-DQ8 molecules are capable to bind to the deamidated gluten peptides, which are presented to the CD4⁺ T cells, inducing the activation and proliferation of these lymphocytes, followed by the production of proinflammatory cytokines, leading to histological changes in the intestinal mucosa like villous atrophy and crypt hyperplasia. In addition, the B cells are also activated and they can be differentiated into plasma cells, secreting IgA and IgG antibodies against tTG2 and deamidated gluten peptides (Tye-Din et al., 2018). There are a wide range of clinical manifestations as diarrhea and abdominal pain, among others (Caio et al., 2019; Green et al., 2015) (Figure 1.3).

In a recent meta-analysis, the global prevalence of CD (biopsy-confirmed) was 0.7%. However, the prevalence of CD varies with sex, age and geographic location (Singh et al., 2018). Europe (including data from Russia), together with Oceania, present the highest prevalence, and the highest estimates of CD were in Finland and Sweden inside Europe with values from 2% to 3% (Catassi et al., 2014). There is evidence that the prevalence of CD has increased over time (Singh et al., 2018). Although the origin of this increment is still unknown, several hypotheses were proposed: (i) the development of high-yield but also high-immunogenic varieties; (ii) the transformation of bread-making processes introduced by the expansion of industrial baking techniques with short-time fermentation which minimize the proteolytic of proteins and lead to a high immunogenic-undigested peptides (Gobbetti et al., 2007); (iii) the diagnostic tests, which are becoming increasingly accurate and widely used for CD screening (Catassi et al., 2015); and (iv) diet factors like the increasing of CD in countries where wheat was not the main staple food but are adopting the western lifestyle, including a gluten-rich diet (Catassi et al., 2015).



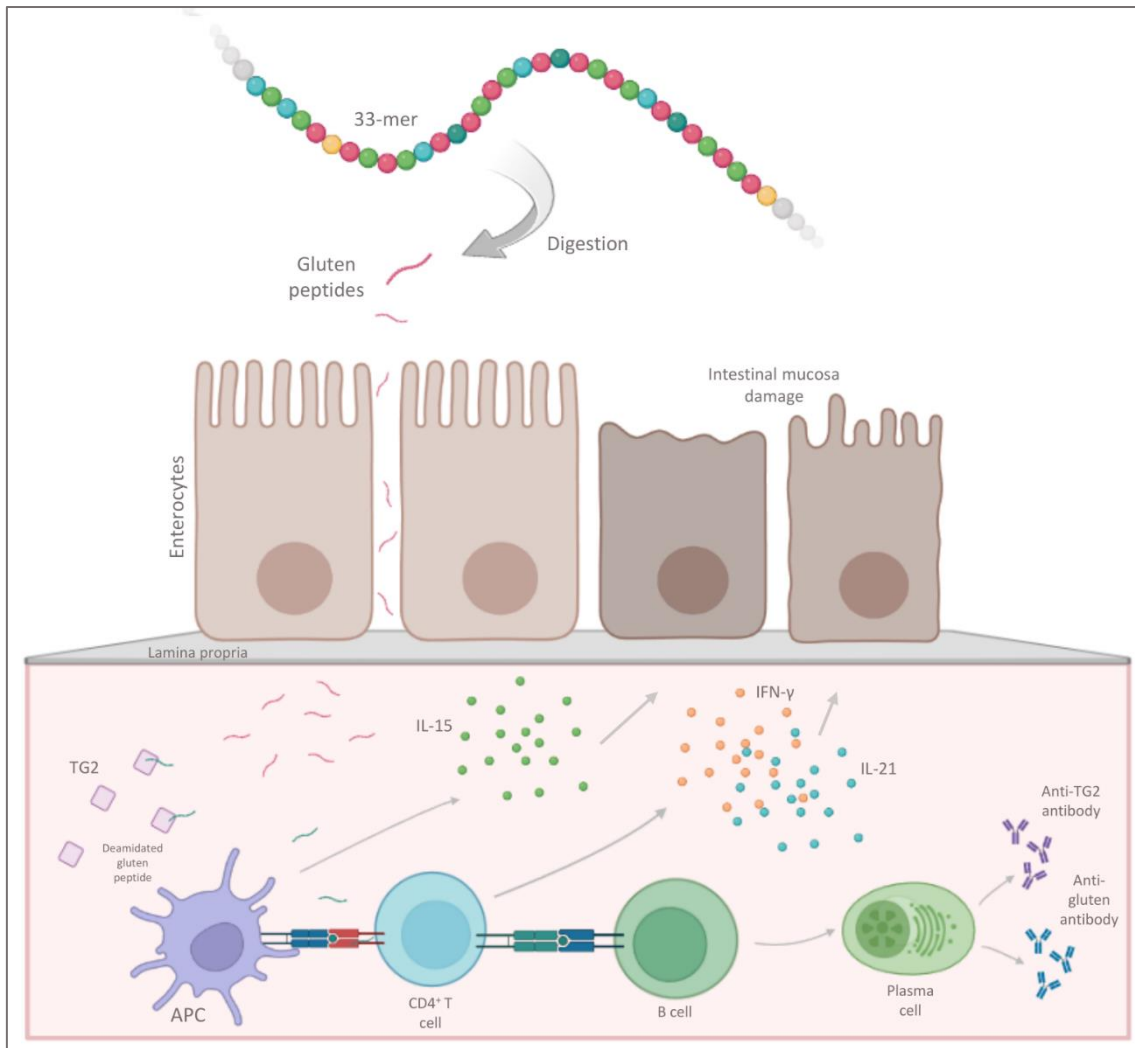


Figure 1.3. Celiac Disease pathogenesis. APC: antigen presenting cells, IL: interleukin, IFN- γ : interferon- γ , TG2: transglutaminase 2. The gluten peptide represented as an example is the region which comprises the 33-mer present in the α -gliadins.

3.1. CD epitopes and immunogenic peptides

The gluten epitopes are 9-mer amino acid core peptides that form stable complexes with HLA-DQ molecules. These epitopes are often localized in gluten peptides resistant to proteolytic degradation in the gastrointestinal lumen. They are rich in proline and glutamine residues and therefore the target of tTG2, increasing their binding properties to the CD-associated HLA-DQ molecules (Sollid et al., 2012; 2020).

Both the three gliadin families and the two glutenin families contain CD epitopes. Particularly, the α -gliadins present an immunogenic peptide called 33-mer which is resistant to digestion by intestinal brush-border enzymes (Shan et al., 2002), and comprises 6 overlapping copies of three highly immunogenic epitopes named DQ2.5_glia_ α 1a, DQ2.5_glia_ α 1b and DQ2.5_glia_ α 2 (Shan et al., 2002). This peptide can be found on the D-subgenome of bread wheat at low frequency, which means that



there are only a few copies of this peptide in the genome. It was proposed to evolve from one α -gliadin subtype found in *A. tauschii* (Ozuna et al., 2015). However, shorter versions of the 33-mer peptide, harboring 1-4 CD epitopes are widely distributed among durum and bread wheat cultivars. Another immunogenic peptide containing a great number of overlapping epitopes is the 26-mer peptide found in the γ -gliadins. In addition to being multivalent, this peptide is also resistant to proteolysis in the intestinal brush border membrane (Shan et al., 2005). As the 33-mer, the full 26-mer peptide is exclusive of the D-subgenome of bread wheat and it is present at low-frequency (Salentijn et al., 2012).

In addition to the CD epitopes associated with the adaptive immune response, the p31-43 peptide, related to the non-T cell-dependent innate immune response in CD was reported (Maiuri et al., 2003). This peptide is located in the N-terminal repetitive region of the α -gliadins. Altogether, this gliadin fraction is one of the most immunogenic and therefore is the most studied to determine the inter-specific variability and abundance of epitopes among cereals (Ozuna et al., 2015; Salentijn et al., 2013; Cho et al., 2018). To that, amplicon sequencing based on Next Generation Sequencing (NGS) is an excellent tool to analyze these immunogenic complexes with a high density of epitopes.

4. Biotechnological approaches for obtaining low-gluten wheat

So far, the treatment of CD is to follow a life-long gluten-free diet (GFD). However, following a GFD has many disadvantages: i) because there are many wheat-based products, ii) gluten is widely used as an additive in many foods that naturally do not contain gluten, iii) GFD was reported to have negative effects on gut microbiota in patients with CD, NCWS and healthy people (Caio et al., 2020), and iv) GFD products, particularly gluten-free bread, have a higher content of both saturated and hydrogenated fatty acids than their normal counterpart, resulting in a high glycemic index that can increase the risk of obesity, insulin resistance and cardiovascular diseases (Scaramuzza et al., 2013). Therefore, the development of wheat varieties without the immunogenic components, but maintaining the organoleptic properties of standard wheat, is a very appealing objective. Although naturally there is a wide range of variability for immunogenic gliadins among hexaploid, tetraploid and diploid wheat species, the structural complexity of the gliadin immunogenic loci, the high copy number of gliadin genes and the polyploid character of wheat have prevented the effective application of classical cross-based breeding techniques to develop wheat varieties with low-gluten and a low-capacity to trigger the immunogenic response. In contrast, biotechnological approaches based on RNAi and new plant breeding techniques (NPBTs) such as CRISPR/Cas have made it possible to develop this kind of wheat lines. For both RNAi and CRISPR/Cas9 techniques, the selection of the best wheat varieties, and candidates to be transformed is crucial. Therefore, the characterization of the gliadin CD-related complexes of wheat varieties would allow the improvement of gene editing and RNAi silencing results.

4.1. RNA of interference



The low-gluten RNA of interference (RNAi) wheat lines were obtained by the application of RNAi silencing technology. This technology is based on the post-transcriptional RNAi mechanism existing in eukaryotes cells. It is based on a hairpin RNA (hpRNA) vector composed of an inversely repeated (IR) sequence of the target gene, with a spacer region in between. The transcribed RNA hybridizes with itself to form a hairpin structure, which is then processed by the ribonuclease III enzyme Dicer to produce small interfering RNAs (siRNAs). The siRNA will be bound to Argonaute family proteins (AGOs) to form the AGO/siRNA complexes, which will join the RNA-induced silencing complex (RISC). The complementary mRNA of the RISC siRNA will be degraded (Liu et al., 2021). This mechanism prevents the translation of this mRNA into a protein, and it is used for gene silencing (Figure 1.4).

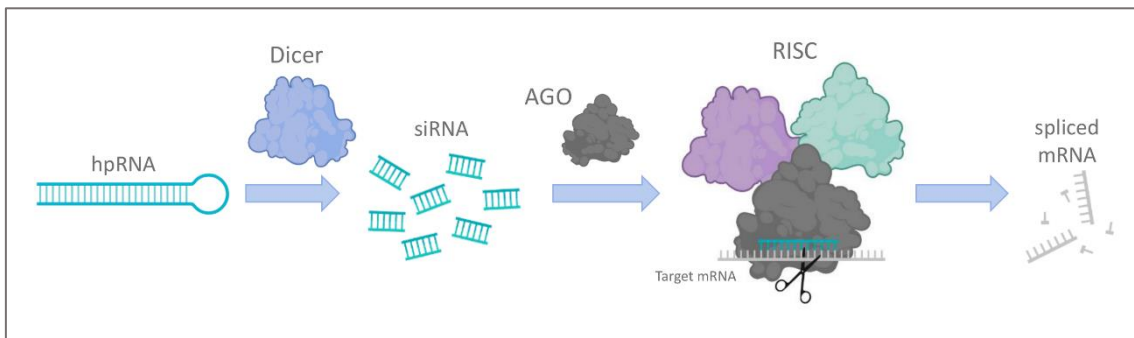


Figure 1.4. RNA of interference mechanism. hpRNA is formed by the inversely repeated (IR) sequence of the target gene and with a spacer region. siRNA: small interfering RNA, AGO: argonaute protein, RISC: RNA-induced silencing complex.

The RNAi technology has been used to silence the genes for each gliadin fraction separately, as shown for the γ - (Gil-Humanes et al., 2008), the ω -1,2 (Altenbach et al., 2019), the ω -5 (Altenbach and Allen, 2011; Altenbach et al., 2015), and α -gliadins (Altenbach et al., 2020). But, the wheat lines with the lowest gluten content were obtained by targeting all the gliadin fractions and transformed with one ω/α hpRNA or both ω/α and γ hpRNAs (Gil-Humanes et al., 2010). One of these RNAi lines, named E82, presents a reduction of 98% in the gluten content (Gil-Humanes et al., 2010). These lines present breadmaking and sensory properties similar to the wild-type lines (Gil-Humanes et al., 2014). Moreover, bread made with E82 flour was tested in clinical trials with NCWS and CD patients. Notably, the E82 bread did not elicit an immune response in DQ2.5-positive patients with CD after a short-term oral challenge (Guzmán-López et al., 2021) while positive changes in the gut microbiota were reported for NCWS patients (Haro et al., 2018).

4.2. CRISPR/Cas9 gene editing

The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) proteins were discovered as part of the adaptive immune system in prokaryotes (Mojica et al., 2005), and adapted for genome engineering (Doudna and Charpentier, 2014). This is a sequence-specific nuclease (SSN) system



which targets specific DNA sequences and make double-stranded breaks (DSBs) in the target site by a bacterial-derived DNA endonuclease (Cas9). The Cas9 is guided to the target site by a 20 bp single guide RNA (sgRNA) sequence. The requirement for Cas9 binding specificity is determined first by the presence of a protospacer-adjacent motif (PAM) downstream of the target region, and second by the pairing of the sgRNA to the target region. The DSBs are repaired by two different pathways: non-homologous end joining (NHEJ) or homology directed repair (HDR) (Figure 1.5). The NHEJ is an error-prone pathway that can lead to insertions or deletions in the target site as the DNA is repaired imprecisely. In contrast, in the HDR pathways the DNA is repaired using a donor template similar to the target one (Samanta et al., 2016).

The α -gliadins genes have been also edited by CRISPR/Cas9 system (Sánchez-León et al., 2018). The immunoreactivity of the wheat lines edited by CRISPR for the α -gliadins was reduced by 85% measured by using the R5 and G12 monoclonal antibodies (Sánchez-León et al., 2018). This system was demonstrated very highly precise and efficient, as of the 45 genes identified in the wild type, up to 35 were mutated with no off-targets. Furthermore, transgene-free lines were obtained, being great candidates for classical breeding programs.

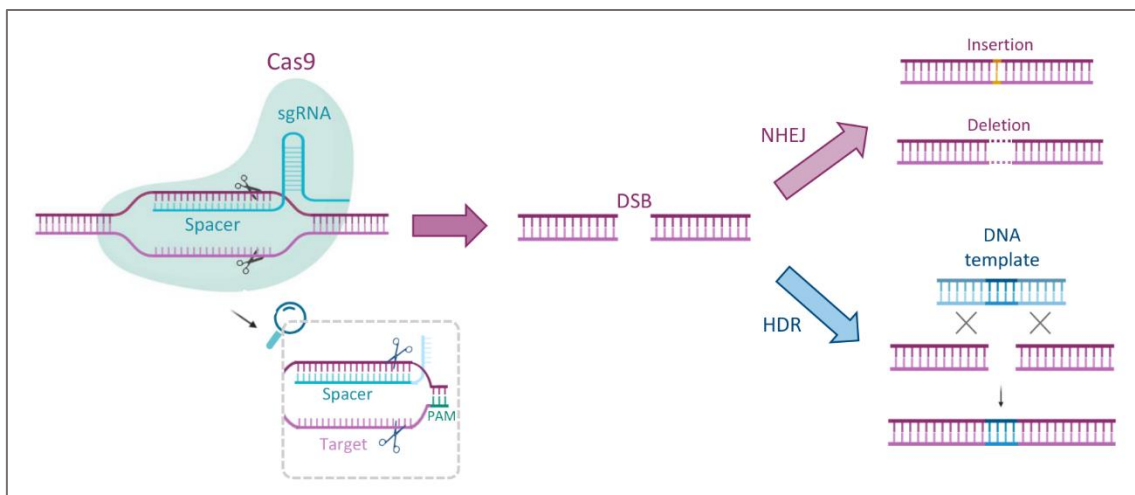


Figure 1.5. CRISPR gene editing scheme: Cas9 recognition of the target site, double strand break (DSB), and the NHEJ and HR DNA repair mechanisms.

5. Grain protein compensation after gliadin silencing

Despite the silencing of gliadins, the wheat RNAi lines present a similar N content to that of the wild-type line, indicating a protein compensation mechanism involving the major protein fractions of the grain to provide a total protein content that does not differ from that of the parental line. However, these readjustments vary depending on the target of the silencing. For example, for the RNAi lines with silencing for the γ -gliadins, the other two gliadin fractions (ω - and α -gliadins) are increased for equalizing the total protein content (Pistón et al., 2011). In contrast, when the wheat ω -5 and 1,2- ω -gliadins were silenced (Altenbach et al., 2019; Altenbach et al., 2014), the HMW and the non-gluten proteins (NGPs) increased their content with respect to the wild-type line. Both,



the LMW and the α -gliadins had opposite variations in both types of lines, increasing in the ω -5-gliadin line, and decreasing in the 1,2- ω -gliadin ones. These compensatory effects were also shown in CRISPR/Cas lines targeting the α -gliadins. These lines presented an increase in the HMW glutenin subunits, while some of them also present a LMW reduction (Sánchez-León et al., 2018).

This compensatory phenomenon seems to be a widespread mechanism as protein rearrangements have also been detected in maize, where RNAi lines for the 22-kDa α -zeins showed an increase in 19-kDa α -zeins, while the RNAi lines with silence for the 19-kDa α -zeins were compensated with the 22-kDa α -zeins (Huang et al., 2004). The compensatory effect of SSPs against glutelin decreasing has also been reported in rice (Iida et al., 1993).

In wheat, the prolamins are regulated mainly at the transcriptional level by the presence of trans-activating transcription factors (TFs) that interact with the cis-regulatory elements (CREs) present in prolamins' promoters (Kawakatsu and Takaiwa, 2010). There are some well-known TFs which recognized the upstream sequence of gliadins and glutenins genes. Among them, the Prolamin Binding Factor (*TaPBF*) stands out together with the Storage Protein Activator (*TaSPA*). These TFs bind to the bifactorial endosperm box (E box), which is comprised of the prolamin box (P box) and the GCN4-like motif (GLM box) (Kawakatsu and Takaiwa, 2010). The *TaPBF* recognizes the P box, while the *TaSPA* binds to the GCN4-like motif. In addition to the E box, there are other CREs widely studied in the SSPs promoters. For example, the RY motif (5'-CATGCA-3'), which is directly recognized by a B3-superfamily TF called *TaFUSCA3* (Sun et al., 2017); or the R2R3MYB recognition site (5'-AACAAAC-3'), to which *TaGAMyb* is bound (Guo et al., 2015).

In addition, some of the TFs actively involved in the transcriptional regulation of prolamins and other SSPs have also a regulatory role in the synthesis of starch. Particularly, they regulate the expression of the starch synthesis-related genes (SSRGs), encoding catalyzers such as starch synthase and starch branching enzyme (Shevkani et al., 2017). The *TaPBF* and *TaSPA* TFs, and their orthologs in maize, are also involved in the SSRGs regulation (Zhang et al., 2016; Orman-Ligeza et al., 2020; Guo et al., 2020).

In the case of the RNAi lines analyzed in this thesis, the silencing targets were the three groups of gliadins. All of them contain plasmids harboring IR sequences of the most conserved regions from the ω -, α - and γ -gliadin genes, driven by two different endosperm-specific promoters: γ -gliadin and D-hordein promoters. Interestingly, these promoters also contain the CREs described above, and therefore, the expression of the silencing fragment is also subjected to the same regulatory mechanisms than their target prolamin genes. In three of these lines – E82 (pGhp- ω/α + pghp8.1 plasmids), D793 (pGhp- ω/α), and D783 (pDhp- ω/α) –, an increase in the HMW was detected in two of them. Only one presented an increase in the LMW, while the other two lines showed a lower content of this protein fraction. The common factor for all three was the increase



in the NGPs, which mainly comprise globulins and albumins (Pistón et al., 2013; Barro et al., 2016).



6. Objectives

The objectives of the present thesis are the following:

- I. To evaluate the effectivity of gliadins silencing and the grain protein compensation in low-gluten RNAi wheat lines under different N availability regimes and temperature treatments.
- II. To understand and develop a model of protein compensation in the grain at transcriptional levels when gliadins have been silenced by RNAi, a model that can be used for the future redesign of wheat grain proteins.
- III. To study and elucidate at the transcriptomic level the role that leaves play through source-sink communications in the grain protein rearrangements in wheat RNAi lines.
- IV. Develop an amplicon sequencing pipeline to study the complexity of CD-related alpha and gamma gliadins in cereals for the identification and integration of varieties with low immunogenic capacity in precision breeding programs.



7. References

- Altenbach, S. B., and Allen, P. V (2011). Transformation of the US bread wheat 'Butte 86' and silencing of omega-5 gliadin genes. *GM Crops* 2, 66–73.
- Altenbach, S. B., Chang, H. C., Rowe, M. H., Yu, X. B., Simon-Buss, A., Seabourn, B. W., et al. (2020). Reducing the Immunogenic Potential of Wheat Flour: Silencing of Alpha Gliadin Genes in a U.S. Wheat Cultivar. *Front. Plant Sci.* 11, 20. doi:10.3389/fpls.2020.00020.
- Altenbach, S. B., Chang, H. C., Yu, X. B., Seabourn, B. W., Green, P. H., and Alaedini, A. (2019). Elimination of omega-1,2 gliadins from bread wheat (*Triticum aestivum*) flour: Effects on immunogenic potential and end-use quality. *Front. Plant Sci.* 10, 580. doi:10.3389/fpls.2019.00580.
- Altenbach, S. B., Tanaka, C. K., and Allen, P. V. (2014). Quantitative proteomic analysis of wheat grain proteins reveals differential effects of silencing of omega-5 gliadin genes in transgenic lines. *J. Cereal Sci.* 59, 118–125. doi:10.1016/j.jcs.2013.11.008.
- Altenbach, S. B., Tanaka, C. K., Pineau, F., Lupi, R., Drouet, M., Beaudouin, E., et al. (2015). Assessment of the allergenic potential of transgenic wheat (*Triticum aestivum*) with reduced levels of ω 5-gliadins, the major sensitizing allergen in wheat-dependent exercise-induced anaphylaxis. *J. Agric. Food Chem.* 63, 9323–9332.
- Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* (80-.). 361, eaar7191.
- Aziz, I., Dwivedi, K., and Sanders, D. S. (2016). From coeliac disease to noncoeliac gluten sensitivity; Should everyone be gluten free? *Curr. Opin. Gastroenterol.* 32, 120–127. doi:10.1097/MOG.0000000000000248.
- Barro, F., Iehisa, J. C. M., Giménez, M. J., García-Molina, M. D., Ozuna, C. V., Comino, I., et al. (2016). Targeting of prolamins by RNAi in bread wheat: Effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotechnol. J.* 14, 986–996. doi:10.1111/pbi.12455.
- Belderok, B., Mesdag, J., Mesdag, H., and Donner, D. A. (2000). *Bread-making quality of wheat: a century of breeding in Europe*. Springer Science & Business Media.
- Caio, G., Lungaro, L., Segata, N., Guarino, M., Zoli, G., Volta, U., et al. (2020). Effect of gluten-free diet on gut microbiota composition in patients with celiac disease and non-celiac gluten/wheat sensitivity. *Nutrients* 12, 1832.
- Caio, G., Volta, U., Sapone, A., Leffler, D. A., De Giorgio, R., Catassi, C., et al. (2019). Celiac disease: a comprehensive current review. *BMC Med.* 17, 1–20.
- Catassi, C., Bai, J. C., Bonaz, B., Bouma, G., Calabrò, A., Carroccio, A., et al. (2013). Non-celiac gluten sensitivity: The new frontier of gluten related disorders. *Nutrients* 5, 3839–3853. doi:10.3390/nu5103839.
- Catassi, C., Gatti, S., and Fasano, A. (2014). The new epidemiology of celiac disease. *J. Pediatr. Gastroenterol. Nutr.* 59, S7–S9. doi:10.1097/01.mpg.0000450393.23156.59.
- Catassi, C., Gatti, S., and Lionetti, E. (2015). World perspective and celiac disease epidemiology. *Dig. Dis.* 33, 141–146. doi:10.1159/000369518.
- Cho, K., Beom, H.-R., Jang, Y.-R., Altenbach, S. B., Vensel, W. H., Simon-Buss, A., et al. (2018). Proteomic profiling and epitope analysis of the complex α -, γ -, and ω -gliadin families in a commercial bread wheat. *Front. Plant Sci.* 9, 818.
- Cornell, H., and Cauvain, S. P. (2003). Bread making: improving quality. *Fish Shellfish Immunol. Dalmo, RA, Bogwald, J., Eds.*
- Daly, M., Bromilow, S. N., Nitride, C., Shewry, P. R., Gethings, L. A., and Mills, E. N. C. (2020). Mapping coeliac toxic motifs in the prolamin seed storage proteins of barley, rye, and oats using a curated



- sequence database. *Front. Nutr.* 7, 87.
- de Sousa, T., Ribeiro, M., Sabença, C., and Igrejas, G. (2021). The 10,000-year success story of wheat! *Foods* 10, 2124.
- Delcour, J. A., Joye, I. J., Pareyt, B., Wilderjans, E., Brijs, K., and Lagrain, B. (2012). Wheat gluten functionality as a quality determinant in cereal-based food products. *Annu. Rev. Food Sci. Technol.* 3, 469–492.
- Doudna, J. A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science* (80-.). 346, 1258096.
- DuPont, F. M., and Altenbach, S. B. (2003). Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. *J. Cereal Sci.* 38, 133–146.
- Elias, E. M. (1995). Durum wheat products. *Durum Wheat Improv. Mediterr. Reg. New Challenges, Ser. A Séminaires Méditerranéennes* 40, 23–31.
- FAO (2022). FAOSTAT. Available at: <https://www.fao.org/faostat/>.
- Feldman, M., and Levy, A. A. (2015). “Origin and evolution of wheat and related Triticeae species,” in *Alien introgression in wheat* (Springer), 21–76.
- Galili, G. (1997). The prolamin storage proteins of wheat and its relatives. *Cell. Mol. Biol. plant seed Dev.*, 221–256.
- Gibson, P. R., and Shepherd, S. J. (2010). Evidence-based dietary management of functional gastrointestinal symptoms: the FODMAP approach. *J. Gastroenterol. Hepatol.* 25, 252–258.
- Gibson, P. R., Skodje, G. I., and Lundin, K. E. A. (2017). Non-coeliac gluten sensitivity. *J. Gastroenterol. Hepatol.* 32, 86–89.
- Gil-Humanes, J., Pistón, F., Altamirano-Fortoul, R., Real, A., Comino, I., Sousa, C., et al. (2014). Reduced-gliadin wheat bread: An alternative to the gluten-free diet for consumers suffering gluten-related pathologies. *PLoS One* 9, e90898. doi:10.1371/journal.pone.0090898.
- Gil-Humanes, J., Pistón, F., Hernando, A., Alvarez, J. B., Shewry, P. R., and Barro, F. (2008). Silencing of γ -gliadins by RNA interference (RNAi) in bread wheat. *J. Cereal Sci.* 48, 565–568.
- Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L. M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proc. Natl. Acad. Sci. U. S. A.* 107, 17023–17028. doi:10.1073/pnas.1007773107.
- Gobbetti, M., Giuseppe Rizzello, C., Di Cagno, R., and De Angelis, M. (2007). Sourdough lactobacilli and celiac disease. *Food Microbiol.* 24, 187–196. doi:10.1016/j.fm.2006.07.014.
- Green, P. H. R., Lebowl, B., and Greywoode, R. (2015). Celiac disease. *J. Allergy Clin. Immunol.* 135, 1099–1106.
- Guo, D., Hou, Q., Zhang, R., Lou, H., Li, Y., Zhang, Y., et al. (2020). Over-expressing *taspa-b* reduces prolamin and starch accumulation in wheat (*Triticum aestivum* L.) grains. *Int. J. Mol. Sci.* 21, 3257. doi:10.3390/ijms21093257.
- Guo, W., Yang, H., Liu, Y., Gao, Y., Ni, Z., Peng, H., et al. (2015). The wheat transcription factor TaGAMyb recruits histone acetyltransferase and activates the expression of a high-molecular-weight glutenin subunit gene. *Plant J.* 84, 347–359. doi:10.1111/tpj.13003.
- Guzmán-López, M. H., Sánchez-León, S., Marín-Sanz, M., Comino, I., Segura, V., Vaquero, L., et al. (2021). Oral Consumption of Bread from an RNAi Wheat Line with Strongly Silenced Gliadins Elicits No Immunogenic Response in a Pilot Study with Celiac Disease Patients. *Nutrients* 13, 4548.
- Haro, C., Villatoro, M., Vaquero, L., Pastor, J., Giménez, M. J., Ozuna, C. V., et al. (2018). The dietary intervention of transgenic low-gliadin wheat bread in patients with non-coeliac gluten sensitivity (NCGS) showed no differences with gluten free diet (GFD) but provides better gut microbiota



- profile. *Nutrients* 10, 1964. doi:10.3390/nu10121964.
- Huang, S., Adams, W. R., Zhou, Q., Malloy, K. P., Voyles, D. A., Anthony, J., et al. (2004). Improving nutritional quality of maize proteins by expressing sense and antisense zein genes. *J. Agric. Food Chem.* 52, 1958–1964.
- Huo, N., Zhang, S., Zhu, T., Dong, L., Wang, Y., Mohr, T., et al. (2018a). Gene duplication and evolution dynamics in the homeologous regions harboring multiple prolamin and resistance gene families in hexaploid wheat. *Front. Plant Sci.* 9, 673. doi:10.3389/fpls.2018.00673.
- Huo, N., Zhu, T., Altenbach, S., Dong, L., Wang, Y., Mohr, T., et al. (2018b). Dynamic Evolution of α -Gliadin Prolamin Gene Family in Homeologous Genomes of Hexaploid Wheat. *Sci. Rep.* 8, 1–13. doi:10.1038/s41598-018-23570-5.
- Iida, S., Amano, E., and Nishio, T. (1993). A rice (*Oryza sativa* L.) mutant having a low content of glutelin and a high content of prolamine. *Theor. Appl. Genet.* 87, 374–378.
- Juhász, A., Belova, T., Florides, C. G., Maulis, C., Fischer, I., Gell, G., et al. (2018). Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci. Adv.* 4, eaar8602. doi:10.1126/sciadv.aar8602.
- Kawakatsu, T., and Takaiwa, F. (2010). Cereal seed storage protein synthesis: Fundamental processes for recombinant protein production in cereal grains. *Plant Biotechnol. J.* 8, 939–953. doi:10.1111/j.1467-7652.2010.00559.x.
- Kohajdová, Z., and Karovicova, J. (2007). Effect of incorporation of spelt flour on the dough properties and wheat bread quality. *Žywność Nauk. Technol. Jakość* 14.
- Larré, C., Lupi, R., Gombaudo, G., Brossard, C., Branlard, G., Moneret-Vautrin, D. A., et al. (2011). Assessment of allergenicity of diploid and hexaploid wheat genotypes: Identification of allergens in the albumin/globulin fraction. *J. Proteomics* 74, 1279–1289. doi:10.1016/j.jprot.2011.03.014.
- Levy, A. A., and Feldman, M. (2022). Evolution and origin of bread wheat. *Plant Cell*.
- Liu, S., Geng, S., Li, A., Mao, Y., and Mao, L. (2021). RNAi technology for plant protection and its application in wheat. *aBIOTECH*, 1–10.
- Longin, C. F. H., Ziegler, J., Schweiggert, R., Koehler, P., Carle, R., and Würschum, T. (2016). Comparative study of hulled (einkorn, emmer, and spelt) and naked wheats (durum and bread wheat): Agronomic performance and quality traits. *Crop Sci.* 56, 302–311. doi:10.2135/cropsci2015.04.0242.
- Ludvigsson, J. F., Leffler, D. A., Bai, J. C., Biagi, F., Fasano, A., Green, P. H. R., et al. (2013). The Oslo definitions for coeliac disease and related terms. *Gut* 62, 43–52. doi:10.1136/gutjnl-2011-301346.
- Maiuri, L., Ciacci, C., Ricciardelli, I., Vacca, L., Raia, V., Auricchio, S., et al. (2003). Association between innate response to gliadin and activation of pathogenic T cells in coeliac disease. *Lancet* 362, 30–37.
- Mojica, F. J. M., Díez-Villaseñor, C. sar, García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60, 174–182.
- Orman-Ligeza, B., Borrill, P., Chia, T., Chirico, M., Doležel, J., Drea, S., et al. (2020). LYS3 encodes a prolamin-box-binding transcription factor that controls embryo growth in barley and wheat. *J. Cereal Sci.* 93, 102965. doi:10.1016/j.jcs.2020.102965.
- Ozuna, C. V., Iehisa, J. C. M., Giménez, M. J., Alvarez, J. B., Sousa, C., and Barro, F. (2015). Diversification of the celiac disease α -gliadin complex in wheat: A 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *Plant J.* 82, 794–805. doi:10.1111/tbj.12851.
- Payne, P. I. (1987). Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Annu. Rev. Plant Physiol.* 38, 141–153.
- Payne, P. I., Nightingale, M. A., Krattiger, A. F., and Holt, L. M. (1987). The relationship between HMW glutenin subunit composition and the bread-making quality of British-grown wheat varieties. *J. Sci.*



- Pistón, F., Gil-Humanes, J., and Barro, F. (2013). Integration of promoters, inverted repeat sequences and proteomic data into a model for high silencing efficiency of coeliac disease related gliadins in bread wheat. *BMC Plant Biol.* 13, 136. doi:10.1186/1471-2229-13-136.
- Pistón, F., Gil-Humanes, J., Rodríguez-Quijano, M., and Barro, F. (2011). Down-Regulating γ -Gliadins in bread wheat leads to Non-Specific increases in other gluten proteins and has no major effect on dough gluten strength. *PLoS One* 6, e24754. doi:10.1371/journal.pone.0024754.
- Pont, C., Leroy, T., Seidel, M., Tondelli, A., Duchemin, W., Armisen, D., et al. (2019). Tracing the ancestry of modern bread wheats. *Nat. Genet.* 51, 905–911.
- Salentijn, E. M. J., Esselink, D. G., Goryunova, S. V., van der Meer, I. M., Gilissen, L. J. W. J., and Smulders, M. J. M. (2013). Quantitative and qualitative differences in celiac disease epitopes among durum wheat varieties identified through deep RNA-amplicon sequencing. *BMC Genomics* 14, 1–16.
- Salentijn, E. M. J., Mitea, D. C., Goryunova, S. V., van der Meer, I. M., Padioleau, I., Gilissen, L. J. W. J., et al. (2012). Celiac disease T-cell epitopes from gamma-gliadins: immunoreactivity depends on the genome of origin, transcript frequency, and flanking protein variation. *BMC Genomics* 13, 1–12.
- Samanta, M. K., Dey, A., and Gayen, S. (2016). CRISPR/Cas9: an advanced tool for editing plant genomes. *Transgenic Res.* 25, 561–573.
- Sánchez-León, S., Gil-Humanes, J., Ozuna, C. V., Giménez, M. J., Sousa, C., Voytas, D. F., et al. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant Biotechnol. J.* 16, 902–910. doi:10.1111/pbi.12837.
- Scaramuzza, A. E., Mantegazza, C., Bosetti, A., and Zuccotti, G. V. (2013). Type 1 diabetes and celiac disease: The effects of gluten free diet on metabolic control. *World J. Diabetes* 4, 130.
- Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., et al. (2002). Structural basis for gluten intolerance in Celiac Sprue. *Science (80-.)*. 297, 2275–2279. doi:10.1126/science.1074129.
- Shan, L., Qiao, S.-W., Arentz-Hansen, H., Molberg, Ø., Gray, G. M., Sollid, L. M., et al. (2005). Identification and analysis of multivalent proteolytically resistant peptides from gluten: implications for celiac sprue. *J. Proteome Res.* 4, 1732–1741.
- Shevkani, K., Singh, N., Bajaj, R., and Kaur, A. (2017). Wheat starch production, structure, functionality and applications—a review. *Int. J. Food Sci. Technol.* 52, 38–58. doi:10.1111/ijfs.13266.
- Shewry, P. (2019). What is gluten—why is it special? *Front. Nutr.*, 101.
- Shewry, P. R. (2009). Wheat. *J. Exp. Bot.* 60, 1537–1553.
- Shewry, P. R., and Halford, N. G. (2002). Cereal seed storage proteins: Structures, properties and role in grain utilization. *J. Exp. Bot.* 53, 947–958. doi:10.1093/jexbot/53.370.947.
- Shewry, P. R., and Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food Energy Secur.* 4, 178–202. doi:10.1002/FES3.64.
- Singh, P., Arora, A., Strand, T. A., Leffler, D. A., Catassi, C., Green, P. H., et al. (2018). Global Prevalence of Celiac Disease: Systematic Review and Meta-analysis. *Clin. Gastroenterol. Hepatol.* 16, 823–836.e2. doi:10.1016/j.cgh.2017.06.037.
- Sollid, L. M., Qiao, S.-W., Anderson, R. P., Gianfrani, C., and Koning, F. (2012). Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics* 64, 455–460.
- Sollid, L. M., Tye-Din, J. A., Qiao, S. W., Anderson, R. P., Gianfrani, C., and Koning, F. (2020). Update 2020: nomenclature and listing of celiac disease-relevant gluten epitopes recognized by CD4+ T cells. *Immunogenetics* 72, 85–88. doi:10.1007/s00251-019-01141-w.
- Sun, F., Liu, X., Wei, Q., Liu, J., Yang, T., Jia, L., et al. (2017). Functional characterization of TaFUSCA3, a



B3-superfamily transcription factor gene in the wheat. *Front. Plant Sci.* 8, 1133. doi:10.3389/fpls.2017.01133.

Triboï, E., Martre, P., and Triboï-Blondel, A. (2003). Environmentally-induced changes in protein composition in developing grains of wheat are related to changes in total protein content. *J. Exp. Bot.* 54, 1731–1742.

Tye-Din, J. A., Galipeau, H. J., and Agardh, D. (2018). Celiac disease: a review of current concepts in pathogenesis, prevention, and novel therapies. *Front. Pediatr.* 6, 350.

Wieser, H., Koehler, P., and Scherf, K. A. (2023). Chemistry of wheat gluten proteins: Quantitative composition. *Cereal Chem.* 100, 36–55.

Xiong, F., Yu, X., Zhou, L., Zhang, J., Jin, Y., Li, D., et al. (2014). Effect of nitrogen fertilizer on distribution of starch granules in different regions of wheat endosperm. *Crop J.* 2, 46–54.

Xynias, I. N., Mylonas, I., Korpetis, E. G., Ninou, E., Tsaballa, A., Avdikos, I. D., et al. (2020). Durum wheat breeding in the Mediterranean region: Current status and future prospects. *Agronomy* 10, 432.

Yu, X., Chen, X., Wang, L., Yang, Y., Zhu, X., Shao, S., et al. (2017). Novel insights into the effect of nitrogen on storage protein biosynthesis and protein body development in wheat caryopsis. *J. Exp. Bot.* 68, 2259–2274.

Zhang, Z., Zheng, X., Yang, J., Messing, J., and Wu, Y. (2016). Maize endosperm-specific transcription factors O2 and PBF network the regulation of protein and starch synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 113, 10842–10847. doi:10.1073/pnas.1613721113.

Zu-Jun, Y., Guang-Rong, L. I., Chang, L. I. U., Juan, F., Jian-Ping, Z., and Zheng-Long, R. E. N. (2006). Molecular characterization of a HMW glutenin subunit allele providing evidence for silencing of x-type gene on Glu-B1. *Acta Genet. Sin.* 33, 929–936.

The images in the introduction section were made by BioRender (<https://biorender.com/>).





CHAPTER II

Prolamin Content and Grain Weight in RNAi Silenced Wheat Lines Under Different Conditions of Temperature and Nitrogen Availability

Published in:

Marín-Sanz, M., Giménez, M. J., Barro, F., and Savin, R. (2020). Prolamin content and grain weight in RNAi silenced wheat lines under different conditions of temperature and nitrogen availability. *Front. Plant Sci.* 11, 314. doi.org/10.3389/fpls.2020.00314



Abstract

Temperature and nitrogen (N) availability are two important environmental factors that may produce important changes in grain composition during grain filling of bread wheat. In this study, four wheat lines with the down-regulation of gliadins by means of RNA interference (RNAi) have been characterized to determine the effect of thermal stress and N availability on grain weight and quality; with focus on gliadin and glutenin protein fractions. Grain weight was reduced with heat stress (HS) in all RNAi lines, whereas gliadin content was increased in the wild-types. With respect to gliadin content, RNAi lines responded to HS and N availability differently from their respective wild-types, except for ω -gliadin content, indicating a very clear stability of silencing under different environmental conditions. In a context of increased temperature and HS events, and in environments with different N availability, the RNAi lines with down-regulated gliadins seem well suited for the production of wheat grain with low gliadin content.

Keywords: gluten proteins, heat stress, transgenic lines, celiac disease, *Triticum aestivum*



Resumen

La temperatura y la disponibilidad de nitrógeno (N) son dos factores ambientales importantes que pueden producir grandes cambios en la composición del grano durante el llenado de este en trigo harinero. En este estudio, cuatro líneas de trigo con la regulación a la baja de las gliadinas mediante ARN de interferencia (ARNi) han sido caracterizadas para determinar el efecto del estrés térmico y la disponibilidad de N en el peso del grano y su calidad; focalizándonos en las fracciones de proteínas de las gliadinas y gluteninas. El peso del grano se redujo con estrés térmico (ET) en todas las líneas ARNi, mientras que el contenido de gliadinas se incrementó en las líneas control. Con respecto al contenido de gliadinas, las líneas ARNi respondieron al ET y a la disponibilidad de N de forma diferente a sus respectivas líneas control, excepto para el contenido de las ω -gliadinas, indicando una gran estabilidad del silenciamiento bajo diferentes condiciones ambientales. En un contexto de incremento de temperatura y eventos de ET, y en ambientes con distintas disponibilidades de N, las líneas ARNi con regulación a la baja de gliadinas parecen muy adecuados para la producción de grano de trigo con bajo contenido de gliadinas.

Palabras clave: proteínas del gluten, estrés térmico, líneas transgénicas, enfermedad celíaca, *Triticum aestivum*



1. Introduction

Grain yield and quality are critical for wheat breeding and management. Both traits are determined during the grain-filling phase. Understanding the processes affecting grain weight and quality during grain filling is important for improving breeding and management strategies. Grain filling is commonly partitioned into three phases: the lag phase, the effective grain filling period, and the maturation drying phase (Egli, 1998). The lag phase is a period of active cell division, characterized by a rapid increase in water content with almost no dry matter accumulation. Grain dry weight then increases rapidly during the effective grain filling period until the maximum dry weight is attained, after which it remains approximately stable while the grain dries. During the effective grain filling period, starch and proteins are deposited in the endosperm (Jenner et al., 1991). It has been shown that the rate of their deposition is controlled by the source-sink balance (Fischer et al., 1977). Approximately 80% of total proteins in wheat grain are gluten (also termed prolamins) whereas the remaining 20% is composed of non-gluten proteins (NGPs) – mainly albumins and globulins (van den Broeck et al., 2009; Wen et al., 2012). Wheat gluten is able to form a network responsible for the viscoelastic properties of wheat flour since it allows the retention of carbon dioxide released during fermentation (Shewry, 2009). Gluten proteins can be further divided into two fractions: glutenins and gliadins (Lafiandra and Kasarda, 1985; Shewry, 2019). The glutenins form polymers linked by inter-chain disulfide bonds, they are insoluble in alcohol solutions, and can be divided according to their mobility in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDSPAGE) into high molecular weight (HMW) and low molecular weight (LMW) glutenins. The gliadins are monomeric proteins, soluble in alcohol, and divided into three groups according to their mobility by electrophoresis in polyacrylamide gels at acidic pH (A-PAGE); ω , α and γ -gliadins. The glutenins are responsible for wheat dough elasticity and strength, while gliadins are important for viscosity and extensibility (Shewry and Halford, 2002; Shewry et al., 2003). Gluten proteins, particularly the gliadin fraction, are the primary factors responsible for triggering celiac disease (CD), since they contain the most immunogenic CD epitopes (Arentz-Hansen et al., 2000, 2002; Shan et al., 2002; Molberg et al., 2003; Ludvigsson et al., 2013; Box 1).

RNA interference-based (RNAi) techniques are ideal for the down-regulation of specific protein fractions related to CD. Using this technology, γ -gliadins were silenced in two lines of bread wheat, providing reductions of up to 80% in this gliadin fraction (Gil-Humanes et al., 2008). Subsequently, the same workers used chimeric interference RNAs capable of silencing the genes from all the three groups of ω , γ and α -gliadins, to obtain several lines of two wheat genotypes with major reductions (in some cases up to 90%) in total gliadin content (Gil-Humanes et al., 2010). The crossing of the silenced lines with commercial varieties of wheat has allowed the obtention of new lines that effectively express the fragment of silencing in different genetic backgrounds, both to silence γ -gliadins (Gil-Humanes et al., 2012) and total gliadin fractions.



Climate model projections suggest that higher temperatures and heat stress events will become commonplace in most regions where grain crops are produced (Meehl and Tebaldi, 2004). Deleterious effects of high temperature on crop yield and quality are well documented in the literature (e.g. Barnabás et al., 2008). It is also well known that temperate species, such as wheat, maximize their vegetative growth during the period of colder temperatures, and the grain develops as temperature rises. Much work has studied the effects of very high temperature – moderate and short periods – during grain filling in wheat (Wardlaw et al., 2002; Farooq et al., 2011; Nuttall et al., 2018). Typically, high temperature during the grain filling linear phase results in the reduction of grain weight, mainly due to the decrease of soluble starch synthase activity under heat stress (Hawker and Jenner, 1993), reducing starch accumulation (e.g. Bhullar and Jenner, 1986). Simultaneously, heat stress (HS) not only increases protein percentage (Stone, 2001; Wardlaw et al., 2002), but also affects the synthesis of the different prolamin fractions and their ratios, which are responsible of the bread quality (Blumenthal et al., 1993; Stone, 2001).

Nitrogen (N) fertilization is one of the most common management practices used by farmers to improve yields. Consequently, there have been many studies analyzing yield (Foulkes et al., 1998; Barraclough et al., 2014) and protein content (Fischer et al., 1993) in response to soil N availability in wheat. In addition, several studies reported the effects of N on the types of proteins being synthesized during grain filling (Pechanek et al., 1997; Daniel and Triboi, 2000; Johansson et al., 2013), indicating that the synthesis of proteins in cereals is clearly influenced by temperature and N condition under which grain filling proceeds. However, genotypic variability can be found in all these responses (Saint Pierre et al., 2008; Elbashir et al., 2017); and even in the response to interactions between heat and N (Elía et al., 2018; Slafer and Savin, 2018). Therefore, determining the effects of heat and N availability on wheat genotypes with contrasting protein composition are of particular interest in the understanding grain protein distribution and its influence on grain weight and quality.

RNAi lines with low gluten content were subjected to various N and sulfur treatments to study the stability of the gliadin silencing under different fertilization conditions (García-Molina and Barro, 2017). In relation to N, this study showed that the RNAi lines had consistently lower gliadin levels than the wild-type across different N-fertilization regimes, but also that the level of gliadins in RNAi lines was sometimes significantly increased when N availability increased. In that study, N was applied when it would strongly affect grain number and yield and, therefore, may have diluted the availability of N-compounds during grain filling (i.e. more N available for absorption had to be shared between much higher grain numbers). As late N fertilization can be used to maintain green tissues during grain filling and to increase overall N content of the grains (Blandino et al., 2015), it may be relevant to explore whether the response of the RNAi lines would be even more marked. As mentioned before, grain filling is significantly affected by HS which would also favor the synthesis of proteins compared with that of starch (Barlow et al., 2015). It would thus be of interest to determine whether



the synthesis of proteins in general, and gliadins in particular, in these RNAi lines is affected by the combinations of high temperatures and N availability during grain filling.

In this context, the objective of the present work was to determine the effects of contrasting temperature and N availability conditions on the silencing of gliadins. Thus, grain weight, total protein content and gluten protein distribution were studied in a set of RNAi lines and their respective wild-types. The aim was to determine to what degree the silencing of the synthesis of gliadins depends on environmental conditions, which is important for progressing in the development of low-gliadin wheat varieties suitable for new dietary approaches for gluten-related disorders.

BOX 1 Celiac disease (CD) is a chronic enteropathy that results from the ingestion of gluten proteins present in wheat, and other similar proteins in barley and rye (Trier, 1998; Sollid, 2002). After ingestion of gluten, lesions form in the small intestine, characterized by flattening of the microvilli, hyperplasia of crypt cells, and infiltration of leukocytes (Sollid, 2002). As a result, symptoms such as diarrhea and malabsorption of food appear among others, since the spectrum of symptoms can be very broad. The immune response is triggered by the activation of CD4 T cells when they recognize the gluten peptides presented by serotypes HLA-DQ2 and HLA-DQ8. The presence of gluten peptides can be detected by the activity of the tissue transglutaminase 2 enzyme from the intestinal mucosa (Sollid, 2002; Sollid et al., 2012; Gayathri and Rashmi, 2014). CD is present throughout the world and the prevalence in the United States is around 1%, as in Europe, with the highest estimates in Finland and Sweden, and the lowest in Germany (Catassi et al., 2014). Gluten is present in many food products as the main element or as an additive. So far, the only possible treatment for CD is to follow a gluten-free diet for life (Sollid, 2002). The increase in the incidence of the disease was associated with the duration of exposure to gluten (Ventura et al., 1999), which increases the need to obtain wheat with a reduced content of proteins immunogenic for celiac sufferers. In addition to CD, there are other pathologies related to wheat: (i) allergies as wheat-dependent exercise-induced anaphylaxis (WDEIA) – induced by the ω -5 gliadins and the HMW- (Morita et al., 2007; Morita et al., 2009), or baker's asthma associated with non-specific lipid transfer proteins (Brant, 2007; Palacin et al., 2007); and (ii) non-celiac wheat sensitivity (NCWS) (Gibson et al., 2017), with an estimated prevalence ranging from 0.6 to 13% of global population (Aziz et al., 2016). Most of the allergens and proteins related to wheat pathologies have been mapped to the bread wheat Chinese Spring reference genome (RefSeq v1.0, International Wheat Genome Sequencing Consortium) (Appels et al., 2018) contributing to the knowledge of these diseases (Juhász et al., 2018). Moreover, there is a broad study on wheat allergens and CD peptides that allows their identification and composition for diagnostic assays by liquid chromatography-tandem mass spectrometry (Lexhaller et al., 2019).



2. Materials and Methods:

2.1. Plant Material, Chamber Experiment and Treatments

Six lines of bread wheat were used: BW208, D770, D793, Gazul, J631, and M959. BW208 is a line from CIMMYT and Gazul is a commercial variety, and both were used as wild-types. D770 and D793 are lines derived from BW208 with RNA interference (RNAi) silencing of all gliadin fractions (Gil-Humanes et al., 2010). J631 and M959 are derived from crossing the Gazul genotype and line D770. Lines J631 and M959 were backcrossed at least four times with Gazul, always selecting the silencing character and the high and low molecular weight glutenins of Gazul, so that both silenced lines maintain the glutenin profile of this parent line.

We carried out a chamber experiment involving six wheat lines (two wild-type cultivars and four RNAi lines), two temperature treatments (control and heat stress, HS) during the linear phase of grain filling period, and two nitrogen (N) availabilities with three replicates, each replicate was composed of 6 plants (all in all there were 18 plants per genotype \times N \times HS; i.e. 216 plants per chamber). Two seeds were sown in pots (270 cm³) filled with a mixture of 30% peat and 70% soil. After emergence, one plant was left in each pot.

Plants were grown outdoors until heading when all pots were placed in a growth chamber set at 20/15°C. Different temperature treatments were imposed from 10 days after anthesis (DAA) during 10 consecutive days (Supplementary Figure 2.1). The control was set at 25/18°C in a chamber and the HS treatment to 40/18°C in another chamber. Minimum and maximum temperatures of 18 and 25 or 40°C were maintained for 8 and 6 h, respectively (Supplementary Figure 2.1). After the 10 days of treatment, temperatures were set to 25/18°C until maturity.

Chambers were set under long-day conditions (16 h). Pots inside the chambers were rearranged approximately once a week to minimize the effects of possible differences in microenvironment at different positions within each chamber. Pots were watered regularly to avoid water stress. N (9 mg per pot) was applied as urea diluted in all pots at flag leaf appearance (DC 4.5, Zadoks et al., 1974). At heading (DC 5.9), half of the plants received another dose of N (21 mg per pot).

2.2. Grain Weight and Total Protein Determination

At maturity, 18 plants per treatment were sampled. Mature grain weight was determined as the average of all grains from the main spikes harvested. Total grain protein content was determined by Dumas methodology (Dumas, 1831).

2.3. Prolamins Quantification by RP-HPLC

For gliadin and glutenin extraction, two grains from three different plants of each line and treatment were weighed and ground using a ball mill, and sequentially extracted following a previous protocol (Pistón et al., 2011) adapted to small samples. Briefly, gliadins were extracted stepwise three times with up to 400 μ L of 60% (v/v)



ethanol. Samples were centrifuged, and the supernatants collected, mixed together and filtered. The insoluble pellet was re-suspended in 50% (v/v) 1-propanol, 2 M urea, 0.05 M Tris-HCl (pH 7.5), and 2% (w/v) DTT for glutenin extraction, incubated for 30 min at 60°C and centrifuged stepwise three times. For each sample, the three collected supernatants were mixed together and filtered. The protein extracts were used for gliadin and glutenin quantification by Reverse-Phase High-Performance Liquid Chromatography (RP-HPLC, 1200 Series Quaternary LC System liquid chromatography from Agilent Technologies) with a DAD UV-V detector at 210 nm. A 25 cm long column LiChrospher R 100 RP8 (5 µm) (Merck) was used at 50°C and a sample volume equivalent to 2 mg of flour was injected. The flow rate was 0.5 mL·min⁻¹. Mobile phase consisted in a mixture of Acetonitrile (ACN 0.1% TFA) and 0.1% aqueous TFA in a linear gradient (0 min 26% ACN, 60 min 54% ACN). The absolute amount of protein was calculated using bovine serum albumin protein as standard (BSA; BSA ≥ 98%, fraction V. Sigma-Aldrich, St Louis, MO, United States cat. no. A3294) (Supplementary Figure 2.2). The intervals of retention time used for the separation of prolamin fractions peaks are indicated in Supplementary Figure 2.3 according to Wieser et al. (1998). The integration of the peaks was performed automatically by RP-HPLC software with minor modifications if necessary.

2.4. Non-gluten Proteins (NGPs) Determination

The NGPs content was calculated by the difference between the total protein and prolamin content (glutenins and gliadins) for each line. The total protein content (µg/mg) was calculated from the percentage of N obtained by Dumas using the wheat conversion factor (5.83) (Merrill and Watt, 1973).

2.5. Data Processing

The retention time (min) and area (mAU) output of the RPHPLC software was imported into a house developed software made in Python v2.7 (github.com/MiriamMarinS/prolaminsQuantification) to obtain the average values from the transformed technical repeats using the following formulas, that processes the hundreds of output files in a single run. The output of the software is a file with Microsoft Excel format.

$$Protein (\mu g(mg \text{ of flour})^{-1}) = 0.0005 \cdot Area (mAU) \frac{V_{extraction}(\mu L)}{V_{injection}(\mu L) \cdot Grain \text{ weight}(mg)}$$

The integration of the profiles, to obtain the area of each peak, and the subsequent transformation using the formula described, allow estimating the amount of protein for the samples. The arithmetic mean of the three technical repetitions was used for the variance analysis.

2.6. Statistical Analysis

To determine the effect of HS and N multifactorial univariates ANOVA were performed. Two variants of this model were tested: in the first, genotype, temperature, nitrogen and their interactions were independent variables, while grain weight and



protein fractions were dependent variables. It was used to determine the general effect of the treatments on all genotypes. The second, has the same factors and variables, but it was performed for wild-types and RNAi lines separately to determine the effect of the treatments on each of these groups. Principal Components Analysis, PCA, was carried out with grain weight, total gliadin and its fractions, total glutenin and its fractions, and total prolamin as variables to evaluate their contribution to the model variance. The software used for the statistical analysis was R v 3.5.1 (R Core Team, 2018).

3. Results

3.1. Heat Stress and Nitrogen Treatment Effects on Grain Weight and Total Protein

Grain weight was significantly decreased in both wild-types and RNAi lines by heat stress (HS) (Figure 2.1A, Table 2.1, and Supplementary Table 2.1). Additional applications of nitrogen (N) had no significant effect on the RNA interference (RNAi) lines or wild-types (Figure 2.1B and Table 2.1). No significant differences were found in total grain protein content among all genotypes (Table 2.1). HS for a short period did not significantly modify the total protein content for both wild-types and RNAi lines (Figure 2.1C), but the late application of N (N1) resulted in a significant increase of the total protein for both the wild-types and the RNAi lines (Figure 2.1D).

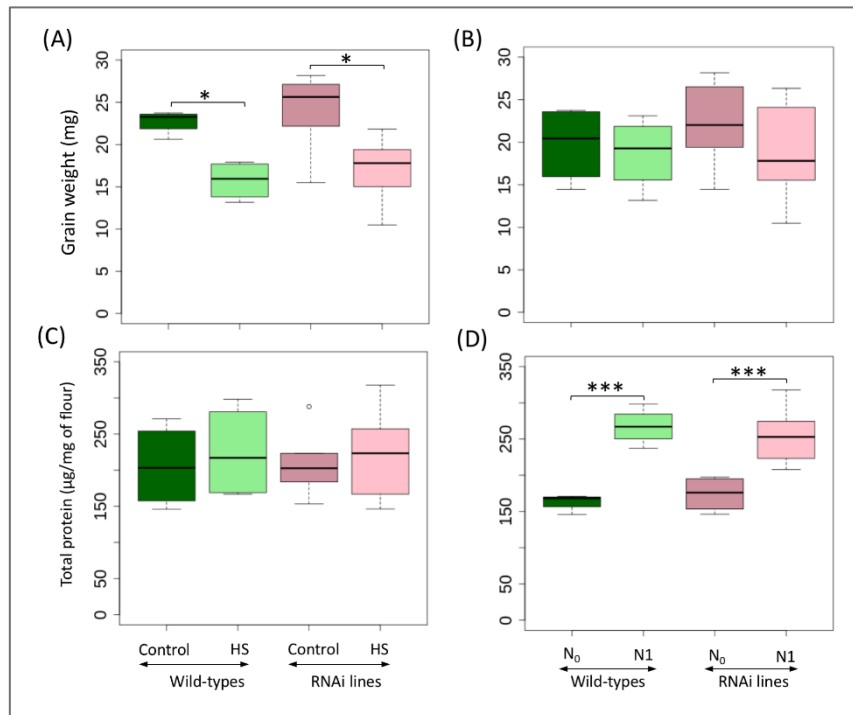


Figure 2.1 Grain weight and total protein content for wild-types and RNAi lines under control and heat stress temperature treatments (A,C) and nitrogen availability (B,D). N₀: no N application after heading, N₁: N application after heading; control: 25/18°C during whole grain filling period, Heat stress (HS): 40/18°C for 10 days during grain filling period. The black line represents the median value. * above the bars indicates significant difference ($*P \leq 0.05$; $***P \leq 0.001$) between treatments according to the variance analysis.



Table 2.1. Significance of the variance of effects of genotype (6 lines: BW208, Gazul, D770, D793, J631 and M959), temperature (2 levels: Control temperature and heat stress), nitrogen (2 levels: No and N1) and their interactions for each variable studied (grain weight, protein fractions and protein ratios).

Variables	Factors	P-value	Variables	Factors	P-value
Grain weight	G	0.002082 **	HMW	G	0.16009
	T	3.25E-05 ***		T	0.09833
	N	0.003541 **		N	0.05425
	GxT	0.604707		GxT	0.55593
	GxN	0.092625		GxN	0.84124
	TxN	0.678413		TxN	0.1672
Total protein	G	0.6512291	LMW	G	0.005985 **
	T	0.2856815		T	0.142497
	N	0.0008094 ***		N	0.104511
	GxT	0.9285325		GxT	0.051446
	GxN	0.5782825		GxN	0.333028
	TxN	0.2003991		TxN	0.929684
ω -gliadins	G	6.27E-05 ***	Total glutenins	G	0.39305
	T	0.07043		T	0.08163
	N	3.81E-05 ***		N	0.04616 *
	GxT	0.00102 **		GxT	0.34164
	GxN	0.06651		GxN	0.65589
	TxN	0.059		TxN	0.24875
α -gliadins	G	1.30E-05 ***	Ratio GLI/GLU	G	0.0006601 ***
	T	0.075602		T	0.0555856
	N	0.001394 **		N	0.3908054
	GxT	0.014791 *		GxT	0.1579138
	GxN	0.010881 *		GxN	0.2167633
	TxN	0.583088		TxN	0.9445682
γ -gliadins	G	3.01E-07 ***	Ratio GLI/TP	G	2.79E-05 ***
	T	0.1145912		T	0.87811
	N	0.1861917		N	0.009226 **
	GxT	0.0005413 ***		GxT	0.024892 *
	GxN	0.180311		GxN	0.072216
	TxN	0.9548613		TxN	0.242135
Total gliadins	G	1.88E-06 ***	Ratio GLU/TP	G	0.23479
	T	0.1784208		T	0.06773
	N	0.0004081 ***		N	0.08808
	GxT	0.0016167 **		GxT	0.30135
	GxN	0.0127328 *		GxN	0.45733
	TxN	0.8348028		TxN	0.17623

The P-value is presented for significant factors of each variable. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$. G, genotype; T, temperature; N, nitrogen; HMW, high molecular weight; LMW, low molecular weight; ratio GLI/GLU, ratio total gliadin content/total glutenin content; ratio GLI/TP, ratio total gliadin content/total protein content; ratio GLU/TP, ratio total glutenin content/total protein content; total proteins, total protein content in percent of nitrogen by Dumas. The degree of freedom (d. f.) of the variance analysis (N - 1, N: number of observations) for the factors are: G, 5; T, 1; N, 1; GxT, 5; GxN, 5; TxN, 1. P < 0.05 are in bold.

3.2. Heat Stress and Nitrogen Treatment Effects on Gliadins and Glutenins

We confirmed that total gliadin content was significantly higher in the wild-types than in the RNAi lines (Figure 2.2A and Table 2.1).



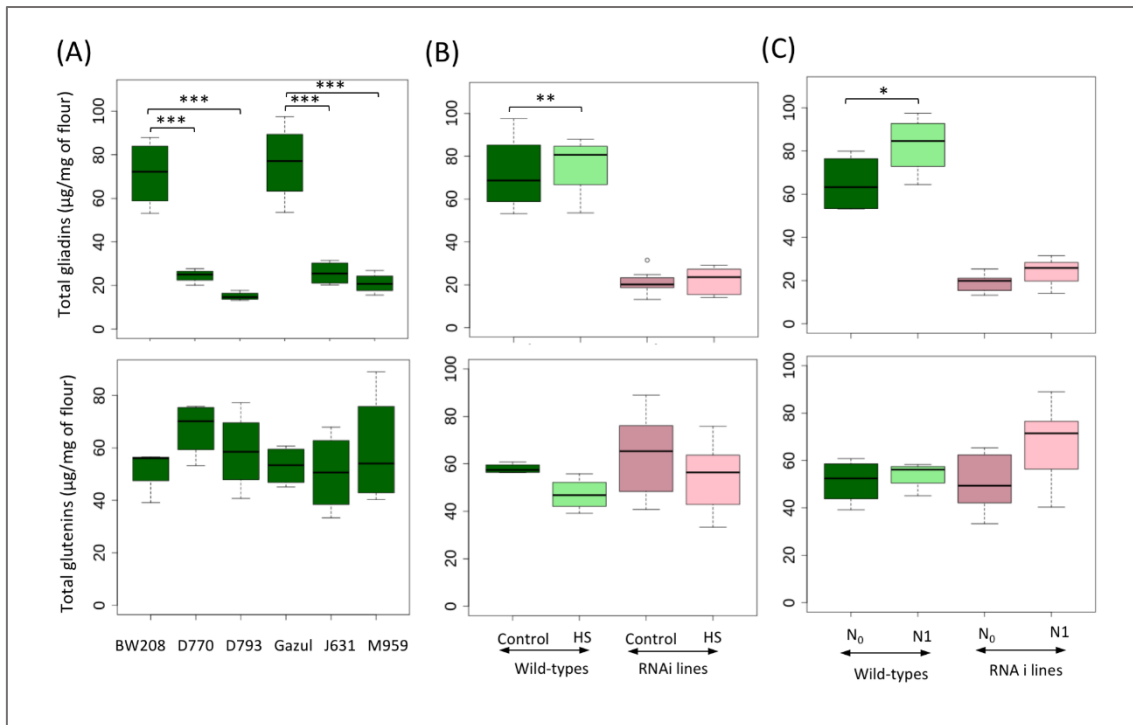


Figure 2.2. Content of total gliadin and total glutenin content for different genotypes with comparisons between RNAi lines and their wild-type by Dunnett's test (A), wild-types and RNAi lines under different temperature treatments (B) and nitrogen availability (C). N0: no N application after heading, N1: application after heading; control: 25/18°C during whole grain filling period, Heat stress HS: 40/18°C for 10 days during grain filling period. The black line represents the median value. * above the bars indicates significant difference ($*P \leq 0.05$; $**P \leq 0.01$) between treatments according to the variance analysis.

Among RNAi lines, D793 had lower content of gliadins than that of the rest of RNAi lines (Supplementary Table 2.1). Both HS and the late application of N (N1) resulted in a significant increase in the total gliadin content for the wild-type lines, whereas no significant variation was observed for the RNAi lines (Figures 2.2B,C).

The degree of silencing of ω -gliadins was lower than that of the rest of the gliadin fractions (Supplementary Figure 2.4A and Supplementary Table 2.1). A significant increase in the content of ω -gliadins due to supplementary N was found on both wild-types and RNAi lines, while HS only had an effect on the wild-types (Supplementary Figures 2.4B,C and Table 2.1). An overall effect of N level on the α -gliadin content (Supplementary Figure 2.4C and Table 2.1), as well as on the total gliadin content (Figure 2.2C), was observed in the wild-types, but the RNAi lines did not show this effect. HS treatment had no effect on the α -gliadin content for wild-types and RNAi lines (Supplementary Figure 2.4B). Conversely, γ -gliadin content was not affected by N availability, but a decrease in the amount of this fraction under HS was seen in wild-types, but not in the RNAi lines (Supplementary Figure 2.4B and Supplementary Table 2.1). Overall, the RNAi line D793 showed the highest reduction in α - and γ -gliadins (Supplementary Figure 2.4A and Supplementary Table 2.1).



The content of glutenins at grain maturity (Supplementary Table 2.1) was not significantly higher for most RNAi genotypes than that of the wild-type lines (Figure 2.2A). The content of HMW was not statistically different between wild-types and RNAi lines (Supplementary Figure 2.5A and Table 2.1). HS had no effect on the HMW fraction, and for N treatment, RNAi lines tend to have a higher amount of HMW. In contrast, the LMW fraction was decreased in most of the RNAi lines in comparison to that of the wild-types (Supplementary Figure 2.5A, Table 2.1, and Supplementary Table 2.1). In addition, LMW content was affected by HS and N availability; in the wild-types LMW proteins decreased under HS; in RNAi lines LMW fractions increased at higher N availability (Supplementary Figures 2.5B,C and Supplementary Table 2.1).

The total gliadin/total protein ratio (GLI/TP) confirms that RNAi lines have lower gliadin content than the wild-types, particularly line D793 whatever experimental conditions. However, a lower total glutenin/total protein ratio (GLU/TP), was seen in wild-types than in RNAi lines without additional N supply, and in BW208 under HS with N1 than in RNAi lines (Figure 2.3 and Table 2.1).

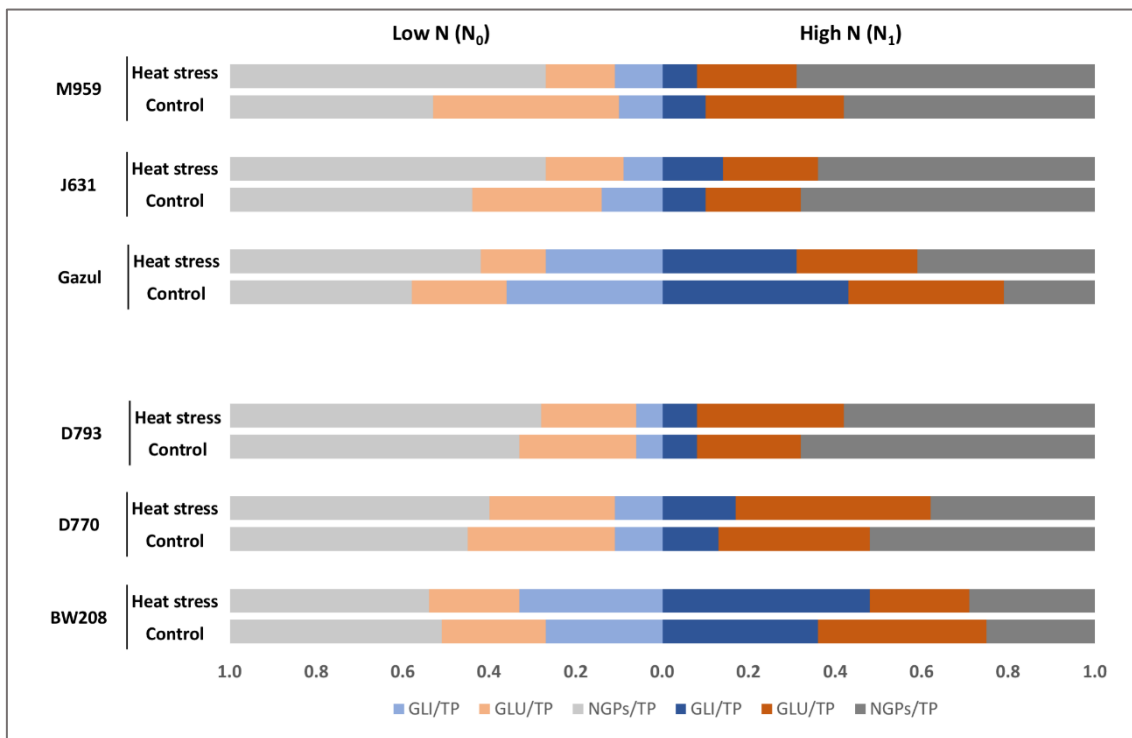


Figure 2.3. Total gliadin/total grain protein ratio (GLI/TP), total glutenin/total grain protein ratio (GLU/TP) and NGPs/total grain protein ratio (NGPs/TP) between treatments and genotypes. The ratios were obtained with mean values for protein content. Low N: no N application after heading, High N: N application after heading; control: 25/18°C during whole grain filling period, Heat stress HS: 40/18°C for 10 days during grain filling period.

HS and N availability modified the GLI/TP ratio in the wild-type lines in different ways: it increased with N1 and HS in BW208 and decreased in Gazul (Figure 2.3). In contrast, minor effects were found in the GLI/TP ratio for the RNAi lines (Figure 2.3).



Regarding the GLU/TP ratio, it was also modified by HS and N availability; it was strongly decreased in Gazul wild-type and RNAi derived lines under HS treatment and N0, while only a minor effect on lines with BW208 background was observed under those conditions; GLI/TP ratio was increased in BW208 RNAi lines under HS and N1, with only but minor effects in Gazul RNAi lines (Figure 2.3).

A Principal Component Analysis (PCA) was carried out, considering the effect of genotypes, temperature and N availability treatments on the variation of the protein fractions and grain weight (Figure 2.4A). Among the gliadin fractions, ω -gliadin and α -gliadin fractions contribute less and more, respectively, to the variance of the model. HMW proteins were the glutenin fraction that contributed most (Figure 2.4A). Gliadin and glutenin contents varied in perpendicular directions, indicating an independent behavior of both families of proteins (Figure 2.4A). The wild-types are separated from the RNAi lines (Figure 2.4B). The ellipses of 95% confidence level of each genotype indicated that there was a strong association between the variation of the glutenins and the silenced genotypes, and on the other hand, between the variation in the prolamins and the wild-types. D770 and J631 tend toward the direction of variation of the LMW proteins, and D793 and M959 toward that of the total glutenins (Figure 2.4). It is interesting to note that the variation in grain weight was independent of variations in the different protein fractions.

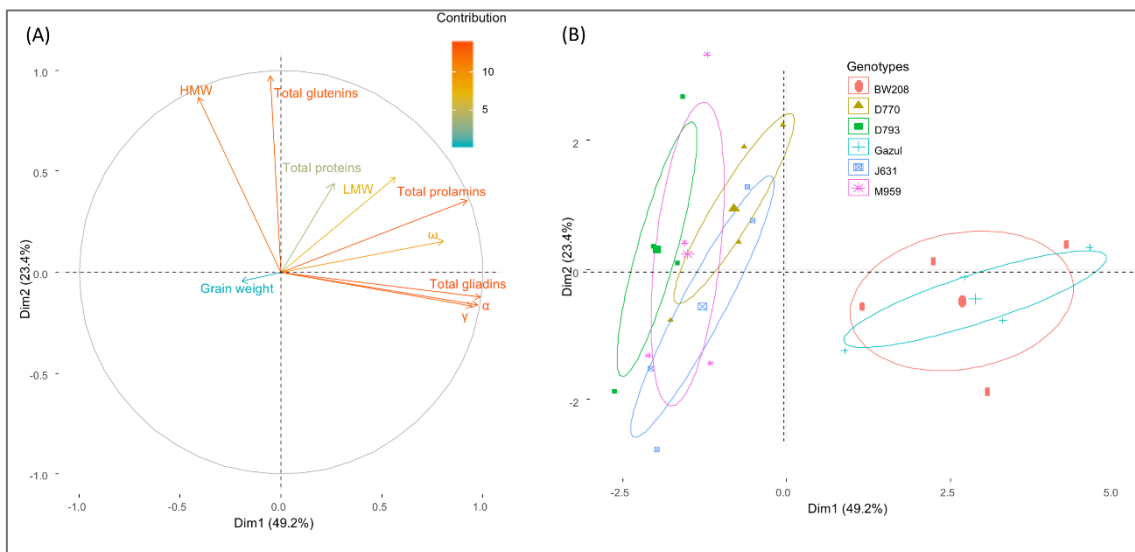


Figure 2.4. Principal Components Analysis (PCA). Effect of genotypes, temperature and N availability treatments on the variation of the protein fractions and grain weight (A). The high values in the color scale indicates a high contribution to the PCA. The direction and the size of the vectors indicate the relationship between all variables and their contribution to each axis. (B) Individuals are represented on the PCA axes with the 95% confidence ellipses showed for each genotype. The largest point for each genotype indicates the intersection of ellipse axes.

4. Discussion

Grain weight, total grain protein, and prolamins content under control temperature and N0 (control conditions) resulted in values similar to those previously reported for



gliadin downregulated lines (Gil-Humanes et al., 2010; Pistón et al., 2013). The decrease of the gliadin/total protein (GLI/TP) ratio in the RNA interference (RNAi) lines could be explained by the increase of the NGPs, as the glutenin/total protein (GLU/TP) ratio was higher in RNAi lines. In previous studies with these and other RNAi lines, protein compensation was observed (Altenbach et al., 2014; García-Molina and Barro, 2017) with increments in nongluten proteins (NGPs) such as serpins, triticins and globulins (Gil-Humanes et al., 2011; Barro et al., 2016).

Brief heat stress (HS) events during the grain filling period generally result in a decrease in grain weight (Bhullar and Jenner, 1986; Savin et al., 1999). The range of variation depends on the genotype, timing and severity of HS (Balla et al., 2019). In the present study, we found a reduction in grain weight of 30% on average under HS and different availability of nitrogen (N). Grain weight was reduced by 35% for lines with BW208 genetic background, and about 24% for Gazul and its RNAi derived lines. This reduction could be mainly due to an extreme dependence on the temperature of starch synthesis, and an irreversible effect of HS on starch production after only a few days or even a few hours per day under control (Daniel and Triboui, 2000; Triboui et al., 2003; Spiertz et al., 2006; DuPont et al., 2006b; Liu et al., 2011; Hurkman et al., 2013) or field conditions (Savin et al., 1996; García et al., 2016; Elía et al., 2018). The percentage of grain protein generally increases under moderately high and very high temperatures (Stone, 2001; Wardlaw et al., 2002), either by a reduction of starch greater than the accumulation of protein, or by a reduction in starch without no change in protein accumulation. However, this response may not always occur for different genetic and environmental backgrounds (Graybosch et al., 1995). Interestingly, in the present study, HS resulted in a higher proportion of gliadin fraction in the wild-types. Other studies reported no effect of HS on total prolamins or even found a decrease with HS (DuPont et al., 2006b; Hurkman et al., 2013). However, Daniel and Triboui (2000) studied each fraction of gliadins and found that the proportion of ω - and α -gliadins increased with HS while γ -gliadins decreased, as found in the wild-types in the present study with the exception of α -gliadins. Also, in this work, grain protein content was increased under HS when post-anthesis N availability was higher (N1), whereas the total prolamins content of the wild-types differed in the response to HS. The RNAi lines, regardless of their genetic background, did not respond to temperature treatments for the total gliadin content, and for the gliadin fractions. This is an indication that these lines have robust gliadin silencing, independently of the temperature environment. Several authors have described a slight increase in LMW and HMW glutenin fractions with HS (DuPont et al., 2006a,b). Evidence has been also reported that the effect of HS is to cause a reduction in the size of glutenin polymers (e.g. Naeem et al., 2012). Nevertheless, in the present work, total glutenin and their fractions content were not modified significantly with HS in any of the genotypes, except for LMW in wild-types, but there is a non-significant decrease in all fractions in wild-types and RNAi lines.

Under higher N availability (N1), both wild-type and RNAi lines increase total grain protein, as previously described for other wheat genotypes (Daniel and Triboui, 2000;



Triboï et al., 2003). Moreover, when increasing and splitting N doses, grain protein concentration increases and protein composition changes by increasing glutenin fractions (Xue et al., 2016). The response to N1 in wild-type lines is greater and in RNAi lines in lesser proportion, also confirmed that gliadin to glutenin ratio increase with N1 (Triboï et al., 2003). The content of ω -gliadins under N1 was increased, in comparison to N0, in the wild-types and RNAi lines, whereas the α -gliadin fraction was increased only in the wild-types as previously reported in D793 and other RNAi lines (Gil-Humanes et al., 2010; García-Molina and Barro, 2017). Total gliadin content did not increase in RNAi lines when additional N was supplied (N1), and this increase of ω -gliadins under N1 in RNAi lines has to be considered in further designing silencing constructs to improve their effectiveness since the ω -5 gliadins are related to wheat-dependent exercise-induced anaphylaxis (WDEIA) (Inomata, 2009; Morita et al., 2009) and ω -1,2 gliadins to CD (Tye-Din et al., 2010). However, the α -gliadins are reported as the major immunogenic complex in wheat, they contain three major celiac disease (CD) immunogenic peptides (Ozuna et al., 2015), and active peptides from this gliadin fraction were responsible for most of the immune response in patients with CD after eating wheat (Tye-Din et al., 2010). The α -gliadins were strongly reduced in the RNAi lines, and this was not affected either by HS or N application. Although some authors have indicated that LMW proteins decrease with high N availability at moderate temperatures (DuPont et al., 2006b; Hurkman et al., 2013), we found that LMW content increased in RNAi lines when additional N was supplied (N1). In contrast, the response in the HMW fraction and total glutenin content under N1, was not statistically significant in any of the genotypes.

5. Conclusion

Wheat grain proteins are important for the breadmaking quality of wheat, but they are also related to human pathologies as celiac disease (CD) and other gluten intolerances. RNA interference (RNAi) technology has provided wheat lines with all the gliadin fractions strongly down-regulated. Heat stress (HS) and nitrogen (N) availability could affect the synthesis and deposition of proteins during grain filling. Wild-types and RNAi lines studied in this work responded similarly for total grain protein and the content of ω -gliadins to additional N supply, as well as for the grain weight under HS. While the wild-types increase their total gliadin content under HS or high N availability the RNAi lines did not. Interestingly, the α -gliadin content, the most CD immunogenic fraction, is unaffected in the RNAi lines under additional N supply, but it was increased in wild-types. Therefore, under the specific scenario of brief events of temperature increase or additional application of N, studied in this work, the RNAi lines demonstrated a high stability of down-regulation of gliadins. However, further evaluations under field conditions will be necessary to confirm that the silencing of gliadin fractions in RNAi lines can be maintained under different abiotic stress environments.



6. References

- Altenbach, S.B., Tanaka, C.K., and Seabourn, B.W. (2014). Silencing of omega-5 gliadins in transgenic wheat eliminates a major source of environmental variability and improves dough mixing properties of flour. *BMC plant biology* 14(1), 393.
- Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361(6403), eaar7191.
- Arentz-Hansen, E., McAdam, S., Molberg, Ø., Kristiansen, C., and Sollid, L. (2000). Production of a panel of recombinant gliadins for the characterisation of T cell reactivity in coeliac disease. *Gut* 46(1), 46-51.
- Arentz-Hansen, H., Mcadam, S.N., Molberg, Ø., Fleckenstein, B., Lundin, K.E., Jørgensen, T.J., et al. (2002). Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. *Gastroenterology* 123(3), 803-809.
- Aziz, I., Dwivedi, K., and Sanders, D.S. (2016). From coeliac disease to noncoeliac gluten sensitivity; should everyone be gluten free? *Current opinion in gastroenterology* 32(2), 120-127.
- Barlow, K., Christy, B., O'leary, G., Riffkin, P., and Nuttall, J. (2015). Simulating the impact of extreme heat and frost events on wheat crop production: A review. *Field Crops Research* 171, 109-119.
- Barnabás, B., Jäger, K., Fehér, A. (2008). The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell & Environment*, 31: 11–38
- Barraclough, P.B., López-Bellido, R., and Hawkesford, M.J. (2014). Genetic variation in the uptake, partitioning and remobilisation of nitrogen during grain-filling in wheat. *Field Crops Research* 156, 242-248.
- Barro, F., Iehisa, J.C., Giménez, M.J., García-Molina, M.D., Ozuna, C.V., Comino, I., et al. (2016). Targeting of prolamins by RNAi in bread wheat: effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant biotechnology journal* 14(3), 986-996.
- Balla, K., Karsai, I., Bónis, P., Kiss, T., Berki, Z., Horváth, Á., et al. (2019). Heat stress responses in a large set of winter wheat cultivars (*Triticum aestivum* L.) depend on the timing and duration of stress. *PloS one* 14(9).
- Bhullar, S., and Jenner, C. (1986). Effects of temperature on the conversion of sucrose to starch in the developing wheat endosperm. *Functional Plant Biology* 13(5), 605-615.
- Blandino, M., Vaccino, P., and Reyneri, A. (2015). Late-season nitrogen increases improver common and durum wheat quality. *Agronomy journal* 107(2), 680-690.
- Blumenthal, C., Barlow, E., and Wrigley, C. (1993). Growth environment and wheat quality: the effect of heat stress on dough properties and gluten proteins. *Journal of Cereal Science* 18(1), 3-21.
- Brant, A. (2007). Baker's asthma. *Current opinion in allergy and clinical immunology* 7(2), 152-155.
- Catassi, C., Gatti, S., and Fasano, A. (2014). The new epidemiology of celiac disease. *Journal of pediatric gastroenterology and nutrition* 59, S7-S9.
- Daniel, C., and Triboi, E. (2000). Effects of temperature and nitrogen nutrition on the grain composition of winter wheat: effects on gliadin content and composition. *Journal of Cereal Science* 32(1), 45-56.
- Dumas, J. B. A. (1831). *Procédes de l'analyse organique*. *Ann. Chim. Phys.* 2(47), 198–213.
- DuPont, F., Hurkman, W., Vensel, W., Chan, R., Lopez, R., Tanaka, C., et al. (2006a). Differential accumulation of sulfur-rich and sulfur-poor wheat flour proteins is affected by temperature and mineral nutrition during grain development. *Journal of Cereal Science* 44(1), 101-112.



- Dupont, F.M., Hurkman, W.J., Vensel, W.H., Tanaka, C., Kothari, K.M., Chung, O.K., et al. (2006b). Protein accumulation and composition in wheat grains: effects of mineral nutrients and high temperature. *European Journal of Agronomy* 25(2), 96-107.
- Egli, D.B. (1998) Seed biology and the yield of grain crops. CAB International, New York, 178 p
- Elbashir, A.A., Gorafi, Y.S., Tahir, I.S., Elhashimi, A.M., Abdalla, M.G., and Tsujimoto, H. (2017). Genetic variation in heat tolerance-related traits in a population of wheat multiple synthetic derivatives. *Breeding science* 67(5), 483-492.
- Elía, M., Slafer, G.A., and Savin, R. (2018). Yield and grain weight responses to post-anthesis increases in maximum temperature under field grown wheat as modified by nitrogen supply. *Field crops research* 221, 228-237.
- Farooq, M., Bramley, H., Palta, J.A., and Siddique, K.H. (2011). Heat stress in wheat during reproductive and grain-filling phases. *Critical Reviews in Plant Sciences* 30(6), 491-507.
- Fischer, R., Aguilar, I., and Laing, D. (1977). Post-anthesis sink size in a high-yielding dwarf wheat: yield response to grain number. *Australian Journal of Agricultural Research* 28(2), 165-175.
- Fischer, R.A., Howe, G.N., Ibrahim, Z. (1993). Irrigated spring wheat and timing and amount of nitrogen fertilizer. I. Grain yield and protein content. *Field Crops Research* 33, 37-56.
- Foulkes, M.J., Sylvester-Bradley, R., and Scott, R.K. (1998). Evidence for differences between winter wheat cultivars in acquisition of soil mineral nitrogen and uptake and utilization of applied fertilizer nitrogen. *Journal of Agricultural Science* 130, 29-44.
- García-Molina, M.D., and Barro, F. (2017). Characterization of changes in gluten proteins in low-gliadin transgenic wheat lines in response to application of different nitrogen regimes. *Frontiers in Plant Science* 8, 257.
- García, G.A., Serrago, R.A., Dreccer, M.F., and Miralles, D.J. (2016). Post-anthesis warm nights reduce grain weight in field-grown wheat and barley. *Field crops research* 195, 50-59.
- Gayathri, D., and Rashmi, B. (2014). Development of Celiac disease; pathogenesis and strategies to control: A molecular approach. *Journal of Nutrition & Food Sciences* 4(6), 1.
- Gibson, P.R., Skodje, G.I., and Lundin, K.E. (2017). Non-coeliac gluten sensitivity. *Journal of gastroenterology and hepatology* 32, 86-89
- Gil-Humanes, J., Pistón, F., Giménez, M.J., Martín, A., and Barro, F. (2012). The introgression of RNAi silencing of γ -gliadins into commercial lines of bread wheat changes the mixing and technological properties of the dough. *PLoS One* 7(9), e45937.
- Gil-Humanes, J., Pistón, F., Hernando, A., Alvarez, J.B., Shewry, P.R., and Barro, F. (2008). Silencing of γ -gliadins by RNA interference (RNAi) in bread wheat. *Journal of Cereal Science* 48(3), 565-568.
- Gil-Humanes, J., Pistón, F., Shewry, P.R., Tosi, P., and Barro, F. (2011). Suppression of gliadins results in altered protein body morphology in wheat. *Journal of experimental botany* 62(12), 4203-4213.
- Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L.M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proceedings of the National Academy of Sciences* 107(39), 17023-17028.
- Graybosch, R.A., Peterson, C., Baenziger, P.S., and Shelton, D. (1995). Environmental modification of hard red winter wheat flour protein composition. *Journal of Cereal Science* 22(1), 45-51.
- Hawker, J., and Jenner, C. (1993). High temperature affects the activity of enzymes in the committed pathway of starch synthesis in developing wheat endosperm. *Functional Plant Biology* 20(2), 197-209.



- Hurkman, W.J., Tanaka, C.K., Vensel, W.H., Thilmony, R., and Altenbach, S.B. (2013). Comparative proteomic analysis of the effect of temperature and fertilizer on gliadin and glutenin accumulation in the developing endosperm and flour from *Triticum aestivum* L. cv. Butte 86. *Proteome science* 11(1), 8.
- Inomata, N. (2009). Wheat allergy. *Current opinion in allergy and clinical immunology* 9(3), 238-243.
- Jenner, C., Ugalde, T., and Aspinall, D. (1991). The physiology of starch and protein deposition in the endosperm of wheat. *Functional Plant Biology* 18(3), 211-226.
- Johansson, E., Malik, A.H., Hussain, A., Rasheed, F., Newson, W.R., Plivelic, T., et al. (2013). Wheat gluten polymer structures: The impact of genotype, environment, and processing on their functionality in various applications. *Cereal Chemistry* 90(4), 367-376.
- Juhász, A., Belova, T., Florides, C.G., Maulis, C., Fischer, I., Gell, G., et al. (2018). Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Science advances* 4(8), eaar8602
- Lafiandra, D., and Kasarda, D.D. (1985). One- and two-dimensional (two-pH) polyacrylamide gel electrophoresis in a single gel: separation of wheat proteins. *Cereal Chem* 62(5), 314-319.
- Lexhaller, B., Colgrave, M.L., and Scherf, K.A. (2019). Characterization and Relative Quantitation of Wheat, Rye, and Barley Gluten Protein Types by Liquid Chromatography–Tandem Mass Spectrometry. *Frontiers in Plant Science* 10, article 1530.
- Liu, P., Guo, W., Jiang, Z., Pu, H., Feng, C., Zhu, X., et al. (2011). Effects of high temperature after anthesis on starch granules in grains of wheat (*Triticum aestivum* L.). *The Journal of agricultural science* 149(2), 159-169.
- Ludvigsson, J.F., Leffler, D.A., Bai, J.C., Biagi, F., Fasano, A., Green, P.H., et al. (2013). The Oslo definitions for coeliac disease and related terms. *Gut* 62(1), 43-52.
- Meehl, G.A., Tebaldi, C. (2004). More intense, more frequent, and longer lasting heat waves in the 21st century. *Science* 305: 994-997.
- Merrill, A., and Watt, B. (1973). Energy value of foods: basis and derivation. *Agric Handb US Dep*
- Molberg, Ø., Solheim flÆte, N., Jensen, T., Lundin, K.E., Arentz-Hansen, H., Anderson, O.D., et al. (2003). Intestinal T-cell responses to high-molecular-weight glutenins in celiac disease. *Gastroenterology* 125(2), 337-344.
- Morita, E., Kunie, K., and Matsuo, H. (2007). Food-dependent exercise-induced anaphylaxis. *Journal of dermatological science* 47(2), 109-117.
- Morita, E., Matsuo, H., Chinuki, Y., Takahashi, H., and Dahlström, J. (2009). Food-dependent exercise-induced anaphylaxis—importance of omega-5 gliadin and HMW-glutenin as causative antigens for wheat-dependent exercise-induced anaphylaxis—. *Allergy International* 58(4), 493-498.
- Naeem, H., Paulon, D., Irmak, S., and MacRitchie, F. (2012). Developmental and environmental effects on the assembly of glutenin polymers and the impact on grain quality of wheat. *Journal of Cereal Science* 56(1), 51-57.
- Nuttall, J.G., Barlow, K.M., Delahunty, A.J., Christy, B.P., and O’Leary, G.J. (2018). Acute High Temperature Response in Wheat. *Agronomy Journal* 110(4), 1296-1308.
- Ozuna, C.V., Iehisa, J.C., Giménez, M.J., Alvarez, J.B., Sousa, C., and Barro, F. (2015). Diversification of the celiac disease α -gliadin complex in wheat: a 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *The Plant Journal* 82(5), 794-805.



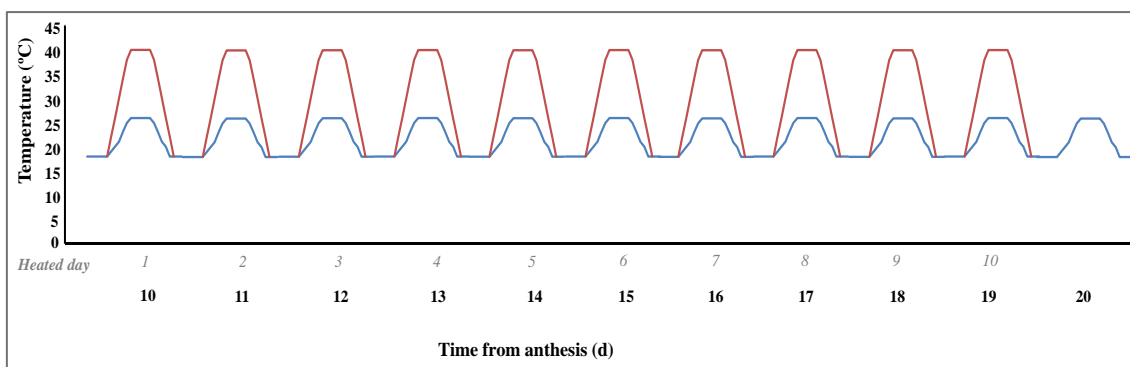
- Palacin, A., Quirce, S., Armentia, A., Fernández-Nieto, M., Pacios, L.F., Asensio, T., et al. (2007). Wheat lipid transfer protein is a major allergen associated with baker's asthma. *Journal of allergy and clinical immunology* 120(5), 1132-1138.
- Pechanek, U., Karger, A., Gröger, S., Charvat, B., Schöggel, G., and Lelley, T. (1997). Effect of nitrogen fertilization on quantity of flour protein components, dough properties, and breadmaking quality of wheat. *Cereal chemistry* 74(6), 800-805.
- Pistón, F., Gil-Humanes, J., and Barro, F. (2013). Integration of promoters, inverted repeat sequences and proteomic data into a model for high silencing efficiency of coeliac disease related gliadins in bread wheat. *BMC plant biology* 13(1), 136.
- Pistón, F., Gil-Humanes, J., Rodríguez-Quijano, M., and Barro, F. (2011). Down-regulating γ -gliadins in bread wheat leads to non-specific increases in other gluten proteins and has no major effect on dough gluten strength. *PLoS One* 6(9), e24754.
- R Core Team (2018). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Saint Pierre, C., Peterson, C.J., Ross, A.S., Ohm, J.B., Verhoeven, M.C., Larson, M., Hofer, B., (2008). Winter wheat genotypes under different levels of nitrogen and water stress: Changes in grain protein composition. *Journal of Cereal Science* 47, 407-416.
- Savin, R., Calderini, D.F., Slafer, G.A., and Abeledo, L.G. (1999). Final grain weight in wheat as affected by short periods of high temperature during pre-and post-anthesis under field conditions. *Functional Plant Biology* 26(5), 453-458.
- Savin, R., Stone, P.J., and Nicolas, M.E. (1996). Responses of grain growth and malting quality of barley to short periods of high temperature in field studies using portable chambers. *Australian Journal of Agricultural Research* 47(3), 465-477.
- Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G.M., et al. (2002). Structural basis for gluten intolerance in celiac sprue. *Science* 297(5590), 2275-2279.
- Shewry, P.R. (2019). What is gluten-why is it special? *Frontiers in Nutrition* 6, 101.
- Shewry, P.R. (2009). Wheat. *Journal of experimental botany* 60(6), 1537-1553.
- Shewry, P.R., and Halford, N.G. (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. *Journal of experimental botany* 53(370), 947-958.
- Shewry, P.R., Halford, N.G., and Lafiandra, D. (2003). "Genetics of wheat gluten proteins," in *Advances in genetics*. Elsevier, 111-184.
- Slafer G.A., Savin, R., (2018). Can N management affect the magnitude of yield loss due to heat waves in wheat and maize? *Current Opinion in Plant Biology* 45:276-283
- Sollid, L.M. (2002). Coeliac disease: dissecting a complex inflammatory disorder. *Nature Reviews Immunology* 2(9), 647.
- Sollid, L.M., Qiao, S.-W., Anderson, R.P., Gianfrani, C., and Koning, F. (2012). Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics* 64(6), 455-460.
- Spiertz, J., Hamer, R., Xu, H., Primo-Martin, C., Don, C., and Van Der Putten, P. (2006). Heat stress in wheat (*Triticum aestivum* L.): Effects on grain growth and quality traits. *European Journal of Agronomy* 25(2), 89-95.
- Stone, P. (2001). The effects of heat stress on cereal yield and quality. *Crop responses and adaptations to temperature stress*, 243-291.



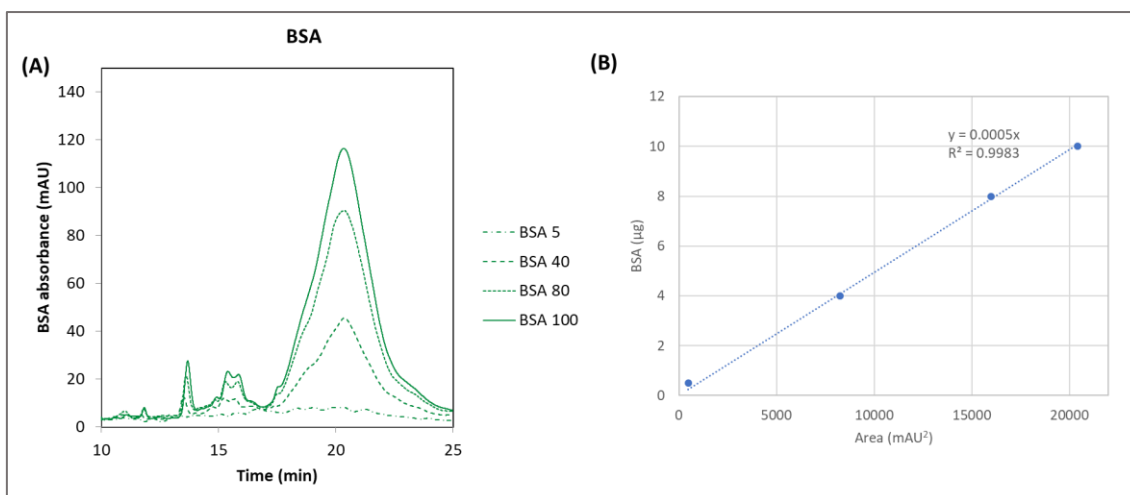
- Triboï, E., Martre, P., and Triboï-Blondel, A.M. (2003). Environmentally-induced changes in protein composition in developing grains of wheat are related to changes in total protein content. *Journal of experimental botany* 54(388), 1731-1742.
- Trier, J.S. (1998). Diagnosis of celiac sprue. *Gastroenterology* 115(1), 211-216.
- Tye-Din, J.A., Stewart, J.A., Dromey, J.A., Beissbarth, T., van Heel, D.A., Tatham, A., et al. (2010). Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Science translational medicine* 2(41), 41ra51-41ra51.
- van den Broeck, H.C., America, A.H., Smulders, M.J., Bosch, D., Hamer, R.J., Gilissen, L.J., et al. (2009). A modified extraction protocol enables detection and quantification of celiac disease-related gluten proteins from wheat. *Journal of Chromatography B* 877(10), 975-982.
- Ventura, A., Magazzù, G., and Greco, L. (1999). Duration of exposure to gluten and risk for autoimmune disorders in patients with celiac disease. *Gastroenterology* 117(2), 297-303.
- Wardlaw, I.F., Blumenthal, C., Larroque, O., and Wrigley, C.W. (2002). Contrasting effects of chronic heat stress and heat shock on kernel weight and flour quality in wheat. *Functional Plant Biology* 29(1), 25-34.
- Wen, S., Wen, N., Pang, J., Langen, G., Brew-Appiah, R.A., Mejias, J.H., et al. (2012). Structural genes of wheat and barley 5-methylcytosine DNA glycosylases and their potential applications for human health. *Proceedings of the National Academy of Sciences* 109(50), 20543-20548.
- Wieser, H., Antes, S., and Seilmeier, W. (1998). Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography. *Cereal chemistry* 75(5), 644-650.
- Xue, C., Rossmann, A., Schuster, R., Koehler, P., and Mühling, K.-H. (2016). Split nitrogen application improves wheat baking quality by influencing protein composition rather than concentration. *Frontiers in plant science* 7, 738.
- Zadoks, J.C., Chang, T.T., and Konzak, C.F. (1974). A decimal code for the growth stages of cereals. *Weed research* 14(6), 415-421.



7. Supplementary Materials

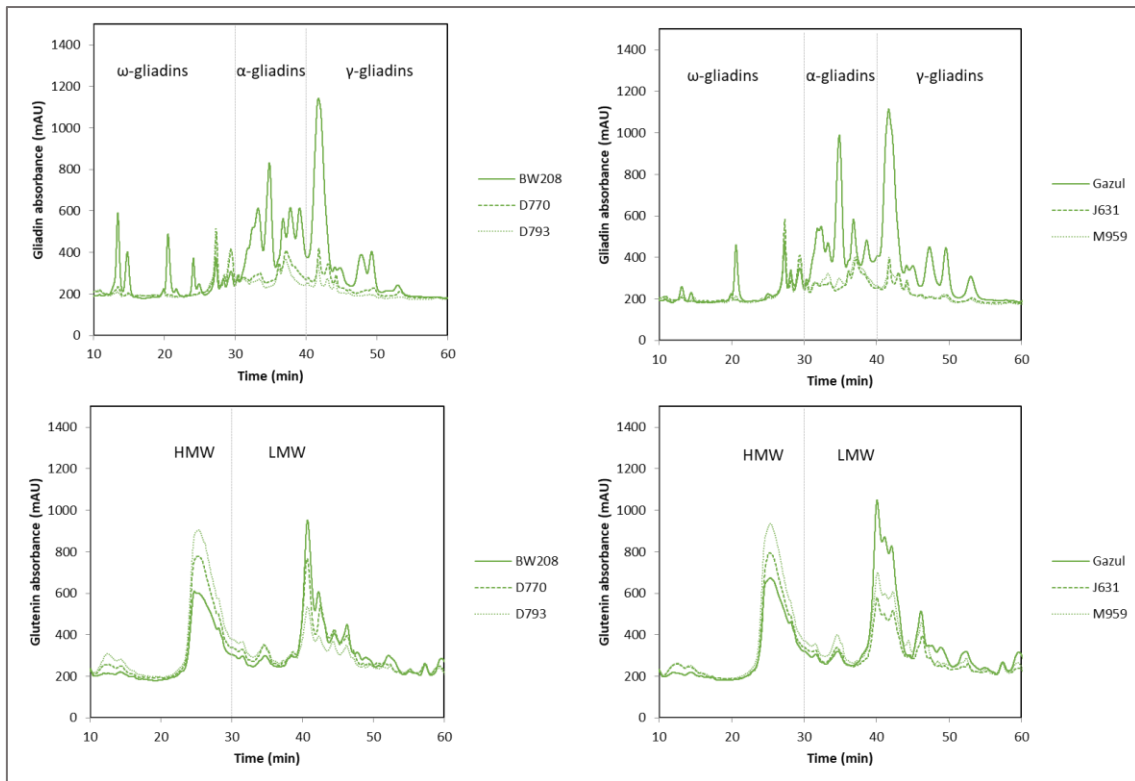


Supplementary Figure 2.1. Temperatures during the day and period of treatment imposition for the unheated control (blue, bottom lines) and the heated treatments (red, upper lines). Temperatures were determined hourly in control and heated experimental chambers for all wheat lines.

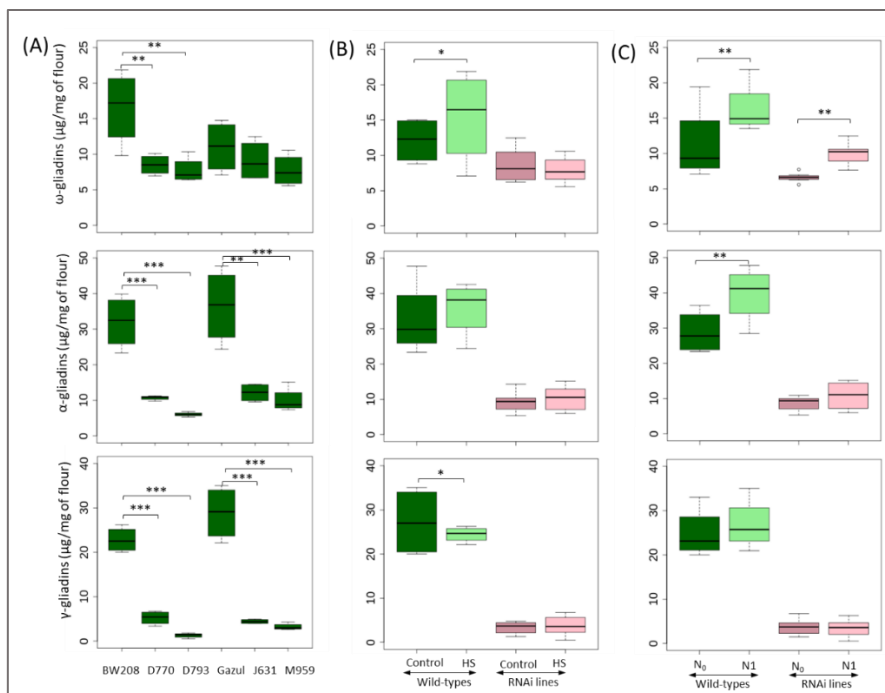


Supplementary Figure 2.2. RP-HPLC chromatograms of BSA used as standards. Three different volumes of solution of $0.1 \mu\text{g}/\mu\text{L}$ BSA protein diluted in water were quantified ($5 \mu\text{L}$, $40 \mu\text{L}$, $80 \mu\text{L}$, $100 \mu\text{L}$) (A). The linear regression between the BSA content and the area of the peak at 20 min has a R^2 coefficient closer to 1 and the slope of the linear equation is the coefficient used for the data transformation formula (B).





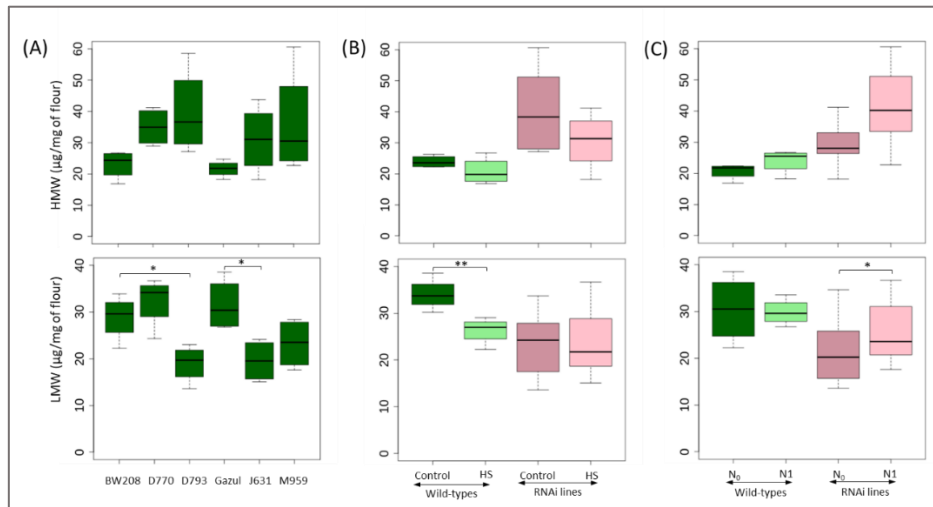
Supplementary Figure 2.3. RP-HPLC chromatograms of gliadin and glutenin fractions of all genotypes and control treatment conditions. No: no N application after heading; control: 25/18 °C during whole grain filling period. The signals were obtained by RP-HPLC software and represented with Microsoft Excel. The intervals of retention time used for the separation of prolamin fractions peaks are indicated according to Wieser et al. (1998).



Supplementary Figure 2.4. Content of ω -, α - and γ -gliadins under Heat stress (HS) and N availability. (A) Comparisons between RNAi lines and their wild-type by Dunnett's test, and for wild-types and RNAi lines under HS (B) and nitrogen availability (C). No: no N application after



heading, N₁: N application after heading; control: 25/18°C during whole grain filling period, HS: 40/18 °C for ten days during grain filling period. The black line represents the median value. * above the bars indicates significant difference (*, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$) between treatments according to the variance analysis.



Supplementary Figure 2.5. Content of HMW and LMW under Heat stress (HS) and N availability. (A) Comparisons between RNAi lines and their wild-type by Dunnett's test, and for wild-types and RNAi lines under HS (B) and nitrogen availability (C). N₀: no N application after heading, N₁: N application after heading; control: 25/18 °C during whole grain filling period, HS: 40/18 °C for ten days during grain filling period. The black line represents the median value. * above the bars indicates significant difference (*, $P \leq 0.05$; **, $P \leq 0.01$) between treatments according to the variance analysis.



Supplementary Table 2.1. Mean values for total grain protein ($\mu\text{g}/\text{mg}$ of flour), total gliadins and glutenins ($\mu\text{g}/\text{mg}$ of flour), and their fractions ($\mu\text{g}/\text{mg}$ of flour), and grain weight (mg) for each genotype under low (N_0) or high nitrogen (N_1), and under control temperature and Heat Stress.

Genotype	Temperature	Nitrogen	ω -gliadin	α -gliadin	γ -gliadin	Total gliadins	HMW	LMW	Total glutenins	Total prolamins	Total grain protein	Grain weight
BW208	Control	N_0	9.8	23.3	20.0	53.2	22.4	33.9	56.3	109.5	146.0	23.7
D770	Control	N_0	7.0	9.8	3.4	20.1	28.9	24.3	53.3	73.3	153.3	27.9
D793	Control	N_0	6.4	5.3	1.5	13.2	27.2	13.6	40.7	54.0	172.0	22.2
Gazul	Control	N_0	8.8	31.1	33.0	72.9	22.2	38.5	60.8	133.7	169.2	23.4
J631	Control	N_0	6.7	9.6	4.0	20.3	27.2	16.3	43.5	63.8	195.6	28.2
M959	Control	N_0	6.3	9.1	4.3	19.7	35.4	27.3	62.7	82.4	197.5	25.2
BW208	Control	N_1	15.0	28.5	21.0	64.5	26.3	30.2	56.5	120.9	237.4	20.6
D770	Control	N_1	9.3	10.9	4.6	24.8	41.3	33.7	75.0	99.8	223.4	22.1
D793	Control	N_1	10.4	6.1	1.3	17.7	58.6	18.7	77.3	95.0	288.0	15.5
Gazul	Control	N_1	14.8	47.7	35.0	97.5	24.7	33.6	58.3	155.8	270.9	23.1
J631	Control	N_1	12.5	14.3	4.8	31.5	43.8	24.2	67.9	99.4	223.4	26.1
M959	Control	N_1	10.6	8.4	2.8	21.8	60.6	28.4	89.0	110.8	208.1	26.4
BW208	Heat stress	N_0	19.4	36.4	24.1	79.9	16.9	22.2	39.1	119.1	167.0	14.5
D770	Heat stress	N_0	7.7	10.9	6.7	25.3	30.7	34.6	65.4	90.7	146.4	19.0
D793	Heat stress	N_0	6.6	6.8	1.8	15.2	41.3	20.7	61.9	77.2	180.2	14.5
Gazul	Heat stress	N_0	7.1	24.4	22.1	53.6	21.3	27.2	48.5	102.1	171.1	17.5
J631	Heat stress	N_0	6.7	10.2	5.0	21.9	18.2	15.1	33.3	55.2	153.8	21.8
M959	Heat stress	N_0	5.6	7.4	2.7	15.6	25.7	19.8	45.4	61.0	194.7	19.9
BW208	Heat stress	N_1	21.9	39.9	26.2	88.0	26.8	29.1	55.8	143.8	263.2	13.2
D770	Heat stress	N_1	10.1	11.3	6.3	27.7	39.2	36.7	75.8	103.5	261.0	15.6
D793	Heat stress	N_1	7.6	6.0	0.5	14.1	32.0	23.0	55.0	69.1	252.2	10.5
Gazul	Heat stress	N_1	13.5	42.5	25.3	81.3	18.3	26.8	45.1	126.4	298.2	17.9
J631	Heat stress	N_1	10.6	14.5	4.0	29.1	35.0	22.8	57.7	86.8	318.0	16.8
M959	Heat stress	N_1	8.6	15.1	3.2	26.9	22.7	17.6	40.3	67.2	253.5	18.9





CHAPTER III

New transcriptomic insights in two RNAi wheat lines with the gliadins strongly down-regulated by two endosperm specific promoters

Published in:

Marín-Sanz, M., Iehisa, J. C. M., and Barro, F. (2022). New transcriptomic insights in two RNAi wheat lines with the gliadins strongly down-regulated by two endosperm specific promoters. *The Crop Journal* 10(1), 194-203. doi.org/10.1016/j.cj.2021.04.009



Abstract

The gluten proteins of wheat grain are responsible for visco-elastic properties of flour, but they also trigger the immune-response of celiac disease. In this work, two low-gliadin RNA interference (RNAi) wheat lines that differ for the promoter driving the silencing (D-hordein and γ -gliadin promoters for D783 and D793 lines, respectively), were characterized at transcriptomic, and protein fraction levels in the grain. Silencing of gliadins provides a readjustment in the grain protein fractions that also affects to the nongluten proteins (NGP), which were increased in both RNAi lines. Determination of wheat gluten by means of the R5 monoclonal antibody also showed a strong reduction in the content of gluten in both RNAi lines. Moreover, fructans, an oligosaccharide linked with the development of non-celiac wheat sensitivity (NCWS) were also significantly decreased in RNAi lines. The down-regulation of gliadins fractions also impacts to other metabolic processes, particularly on carbohydrate metabolism, enzyme regulator activity and response to stress. Genes and transcription factors regulated by ABA were up-regulated, which could suggest the implication of this phytohormone on the stress response observed in the RNAi lines.

Keywords: celiac disease, gliadin, protein compensation, RNAi lines, *Triticum aestivum*



Resumen

Las proteínas del gluten del grano de trigo son responsables de las propiedades viso-elásticas de la harina, pero también desencadenan la respuesta inmune de la enfermedad celíaca. En este trabajo, dos líneas de trigo ARN de interferencia (ARNi) con bajo contenido en gliadinas que difieren en el promotor que conduce el silenciamiento (promotores de D-hordeína y γ -gliadina para las líneas D783 y D793, respectivamente), fueron caracterizadas a nivel transcriptómica y de fracción proteica en el grano. El silenciamiento de gliadinas provoca un reajuste en las fracciones de proteína del grano que también afecta a las proteínas no pertenecientes al gluten (NGPs), que se incrementaron en ambas líneas ARNi. La determinación del gluten de trigo mediante el anticuerpo monoclonal R5 también mostró una fuerte reducción del contenido de gluten en ambas líneas ARNi. Además, los fructanos, oligosacáridos ligados con el desarrollo de la sensibilidad al trigo no celíaca (STNC) se redujeron significativamente en estas líneas. La regulación a la baja de las fracciones de gliadinas también tuvo impacto en otros procesos metabólicos, particularmente en el metabolismo de los carbohidratos, la regulación de la actividad enzimática y la respuesta a estrés. Genes y factores de transcripción regulados por el ABA estuvieron sobreexpresados, lo que podría sugerir la implicación de esta fitohormona en la respuesta a estrés observada en las líneas ARNi.

Palabras clave: enfermedad celíaca, gliadinas, compensación de proteínas, líneas ARNi, *Triticum aestivum*



1. Introduction

The grain storage proteins of wheat are formed by monomeric and polymeric proteins that largely determine the viscoelastic properties of wheat flour. Polymeric proteins are denoted as glutenins, they are insoluble in alcohol as they form large polymers linked by interchain disulfide bonds, and are divided into two fractions; the high molecular weight (HMW) glutenins and the low molecular weight (LMW) glutenins. In contrast, gliadins are mainly monomeric proteins with alcoholic solubility, and they are further divided into three structural groups; ω -, α/β -, and γ -gliadins (Shewry and Halford, 2002). Despite their role in breadmaking quality, gluten proteins also contain epitopes responsible for triggering the immuno-response in celiac disease (CD) (Ludvigsson *et al.*, 2013) after ingestion of gluten-containing foods. CD is a chronic enteropathy characterized for lesions in the small intestine, flattening of the microvilli, hyperplasia of crypt cells and infiltration of leukocytes. The response is triggered by human leukocyte antigen (HLA) HLA-DQ2 and HLA-DQ8-presented peptides that are recognized by CD4 T cells (cluster of differentiation 4 T cells) (Sollid, 2002). In addition to CD, other pathologies are also associated with wheat intake; allergies, which comprises the wheat dependent exercise-induced anaphylaxis (WDEIA) and baker's asthma, and non-celiac wheat sensitivity (NCWS) (Aziz *et al.*, 2016). Wheat gluten is poorly digested in the human intestine and, when an individual eats food containing gluten, non-digested peptides cross the mucosa in the small intestine where the human transglutaminase 2 (tTG2) deaminates glutamine residues present in gluten peptides into glutamic acid, increasing the binding affinity of these peptides to HLA-DQ2/8 molecules. After activation of T-lymphocytes by interaction with antigen-presenting cells, a spectrum of proinflammatory cytokines is released, damaging the enterocytes and producing the intestinal lesions typical of CD. The prevalence of CD is about 1% with a higher prevalence (1.5%) in Northern European countries (Ludvigsson *et al.*, 2014). The prevalence of NCWS is still unknown due to cases of self-diagnoses and the lack of standardized diagnostic criteria or biomarkers, but globally it can affect a large population ranging from 0.6% to 13% (Aziz *et al.*, 2016).

The down-regulation of the immunogenic gliadin genes by RNA interference (RNAi) (Gil-Humanes *et al.*, 2010) provided wheat lines with low toxicity that could be used to prepare foodstuff with excellent sensory properties and health benefits for NCWS (Haro *et al.*, 2018) and CD patients. These RNAi lines showed pronounced shifts in the different grain protein fractions that affect the quality, agronomic, and immunogenic properties (Haro *et al.*, 2018; Gil-Humanes *et al.*, 2014; García-Molina *et al.*, 2017). Lines D793 and D783 showed a strong reduction in the gliadin content of the grain, offering a reduced response to T-cell (Gil-Humanes *et al.*, 2010; Gil-Humanes *et al.*, 2014). These lines contain the same RNAi silencing fragment but driven by two different endosperm specific promoters. Interestingly, in response to the increment in nitrogen supply in both of them, silencing fragment was highly effective in the downregulation of α - and γ -gliadins, at whatever nitrogen supplied (García-Molina and Barro, 2017; Marín-Sanz *et al.*, 2020). Despite the strong down-regulation of gliadins, the total grain



protein content was not significantly affected in comparison to the wild type, indicating a protein compensation with other protein fractions (García-Molina and Barro, 2017) and specifically, increments in non-gluten proteins (NGPs) like serpins, triticins and globulins were observed (Barro *et al.*, 2016). However, the effects of RNAi gliadin silencing on other metabolic and physiological processes in the grain, which in the end, could regulate the compensation mechanisms is still unknown.

In this study, we report new insights in the regulation, transcriptome and trafficking of grain storage and non-storage protein genes in two wheat lines with all gliadin genes down-regulated by RNAi. Results reported are important for understanding the processes of protein regulation in the grain and for undertaking new strategies to redesign the wheat grain proteins by RNAi or CRISPR/Cas towards obtaining new varieties suitable for people suffering any gluten intolerance.

2. Material and Methods

2.1. Plant material

Two RNAi lines, D793 and D783, derived from the wild type BW208 were used in this work. Lines D793 and D783 contain the plasmids pGhp- ω/α and pDhp- ω/α , respectively, harboring inverted repeat sequences of the most conserved regions from ω -, α/β - and γ -gliadin genes, driven by endosperm-specific promoters; γ -gliadin promoter for line D793 and D-hordein promoter for D783 (Pistón *et al.*, 2009; Pistón *et al.*, 2008). The RNAi lines and the wild type were grown in a greenhouse with supplementary lights providing a day/night regime of 12 h/12 h at 24 °C/16 °C.

2.2. Prolamins determination

The gliadin and glutenin fractions were quantified by Reverse Phase High-Performance Liquid Chromatography (RP-HPLC) (1200 Series Quaternary LC System liquid chromatography from Agilent Technologies) with the RP8 column (5 mm in diameter and 250 mm in length, Merck). Three biological replicates for each genotype were analyzed following the protocol established in Pistón *et al.* (2011).

2.3. Total protein, NGPs, gluten content by ELISA, fructans and starch determination

The protein content of whole flour was determined using the Kjeldahl (%N \times 5.7) method according to the standard ICC no. 105/2 (ICC, 1994), while the starch content was determined according to the standard ICC method no. 123/1. (ICC, 1994). Both components were expressed on a 14% moisture basis. The NGPs content was calculated as the difference of total protein and total prolamins content for each line as described in previous work (Marín-Sanz *et al.*, 2020). Gluten content was determined by competitive ELISA assays using the monoclonal antibody R5 as described (Valdés *et al.*, 2003).

The fructan content per dry weight was calculated with Megazyme commercial kit following the fructan assay procedure (KFRUC, www.megazyme.com) using 150 mg of two biological samples and two duplicates per biological sample. After removing



polysaccharides, the hydrolysis of fructan was performed to measure D-fructose content by spectrophotometric at 410 nm (iMark Microplate Absorbance Reader, Bio-Rad). To determine fructan content per dry weight, the moisture content was determined.

2.4. RNA extraction

For RNA sequencing (RNA-seq) data analysis, total RNA was extracted from seeds collected at 23 and 26 days after anthesis (DAA) from three different plants and bulked. For Real-Time quantitative Polymerase Chain Reaction (qPCR) analysis, seeds were collected at 8, 14, 18, 23 and 29 DAA from three different plants as biological replicates. After seed grinding, samples were placed in a tube containing extraction buffer (100 mmol L^{-1} Tris-HCl, pH 9.5, 150 mmol L^{-1} NaCl, 1.0% sarkosyl, 5 mmol L^{-1} DTT (Dithiothreitol), vortex mixed and centrifuged. Acid phenol (pH 4.3)- chloroform extraction was performed using the supernatant. RNA was precipitated with 1/10 vol of 3 mol L^{-1} sodium acetate pH 5.2 and one volume of isopropanol. Resuspended RNA was extracted again using TRIzol reagent (Thermo Fisher Scientific, USA) and chloroform, and then with acid phenol–chloroform mixed with 1/3 vol of cold 3 mol L^{-1} potassium acetate pH 5.5. The supernatant was mixed with salt precipitation solution (1.2 mol L^{-1} NaCl and 0.8 mol L^{-1} sodium citrate tribasic) and 0.25 vol of ethanol for RNA precipitation.

2.5. cDNA synthesis and qPCR experiments

cDNA was synthesized using 1 mg of total RNA and qScript cDNA SuperMix (Quantabio, USA) for RNA-seq and qPCR analyses. qPCR was carried out using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, USA) on CFX Connect Real-time PCR Detection System (Bio-Rad). PCR efficiency of each primer pair was determined for all samples with LinRegPCR version 2012.0 (Ruijter *et al.*, 2009). Normalization factor was calculated for each sample with geNorm (Vandesompele *et al.*, 2002) based on the expression levels of *RLL*, *CDC*, and *ADP-RF* genes. The list of primers used is presented in Supplementary Table 3.1.

2.6. RNA-seq data and bioinformatic analysis

cDNA for mRNA-seq data analysis was high-throughput sequenced by Fundación Parque Científico de Madrid (FPCM) (S/ Faraday 7, Madrid, Spain) with Illumina Genome Analyzer (GAxII) platform. The quality control of total RNA (Agilent's 2100 BioAnalyzer) and the library preparation of poly-A tails-mediated selection (Illumina TruSeq RNA kit) were performed by FPCM. The library type and sequencing details are in Supplementary Table 3.2. Sequencing data can be found in the NCBI Sequence Read Archive under the following accession number: PRJNA593367. The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. Bioinformatic analysis of the RNA-seq was performed on a server with 64 cores and 128 Gb of RAM (Random Access Memory) with the operative system GNU/Linux Ubuntu version 18.04 (www.ubuntu.com, United Kingdom). The preprocessing of the RNA-seq data was carried out by removing the adapters with



cutadapt (cutadapt v. 1.16) (Martin, 2011) and trimming low quality reads, i.e. (<Q30) 5' and 3' reads, and minimum length of 25 bp, with Trimmomatic (Trimmomatic-0.36) (Bolger *et al.*, 2014). Then, 0.1% of reads of all three samples were discarded (Supplementary Table 3.2; Supplementary Figure 3.1). Reads were mapped to the annotated wheat reference genome (International Wheat Genome Sequencing Consortium, IWGSC RefSeq annotation v1.0, [https://wheat-urgi.versailles.inra.fr/Seq-Repository/Assemblies]) (IWGSC, 2014) with STAR mapper (STAR 2.6) (Supplementary Figure 3.1) (Dobin *et al.*, 2013). Quality evaluation of mapped reads and conversion of SAM to BAM formats were performed with Samtools (samtools 1.3.1) (Handsaker *et al.*, 2009). Quantification of genes expression was carried out with StringTie (stringtie 1.3.5) (Pertea *et al.*, 2015). De novo assembly of reads that were not mapped in the known transcripts loci of gtf from wheat genome reference was carried out and further included in the gtf file (Supplementary Table 3.2). Bioconductor package edgeR (edgeR-3.22.5) was used for differential expression analysis in R (R-3.5.1) (Supplementary Figure 3.1) (Robinson *et al.*, 2010). Differential gene expression (DGE) analysis was carried out with Trimmed-Mean of M-values (TMM) normalization method in genes expressed in the transcriptome (genes with counts per million (CPM) row sum in each pair-wise comparison >1). The common dispersion ($d = 0.05096259$) of a subset of housekeeping genes (Kiarash *et al.*, 2018) was calculated. The final value used for DGE with edgeR was increased to 0.1, established for genetically identical model organisms (Chen *et al.*, 2021) to reduce false positives in the data analysis.

2.7. Functional annotation

Functional annotation of genes was performed for known genes and de novo assembled genes. The last ones were annotated by blastx (blast-2.7.1+) against the non-redundant protein database of the National Center for Biotechnology Information (NCBI) and UniProtKB/Swiss-Prot knowledgebase (U. Consortium, 2018). Genes that have been identified as uncharacterized genes by blastx method were subjected to a blastp search against UniProtKB/Swiss-Prot database, by obtaining protein sequence from transcripts using TransDecoder (TransDecoder 5.3.0) (Haas and Papanicolaou, 2021). In addition, gene ontology (GO) enrichment analysis was performed to study differential expressed (DE) genes in cell processes. Custom Python v2.7 scripts were implemented in RNA-seq workflow to facilitate data compatibility between software and annotation, and expression genes matrix SQLite database management.

3. Results

3.1. Grain composition, quality and agronomic features of RNAi lines

In this work, RNAi lines D783 and D793 with a strong reduction in α/β -, ω -, and γ -gliadins were characterized in comparison to their BW208 wild type. Both lines harbor the same RNAi silencing fragment but driven by different endosperm specific promoters; a D-hordein promoter for D783 and a γ -gliadin promoter for D793. A general overview of grain components and other quality and agronomic features between RNAi lines and the wild type is showed in Figure 3.1. All three gliadins fractions were



decreased in both RNAi lines, particularly the γ - and α/β -gliadins. In contrast, the HMW glutenin subunits were increased in both RNAi lines but not the LMW, which was also decreased for D793. This readjustment in the grain protein fractions also affects the NGP, which were significantly increased for line D793. Overall, the total protein and starch contents, the major components of wheat grain, were slightly decreased in both RNAi lines in comparison to the wild type, although it was significant only for starch content in line D793. Determination of wheat gluten by means of the R5 monoclonal antibody also showed a strong reduction in the content of gluten in both RNAi lines. Moreover, fructans, an oligosaccharide linked with the development NCWS (Skodje *et al.*, 2018) were also significantly decreased in RNAi lines. Kernel weight, test weight and anthesis were not affected by the down-regulation of gliadins. In contrast, sodium dodecyl sulphate sedimentation (SDSS) test, a parameter used as predictor for wheat quality, was significantly affected in one RNAi line (D793) (Figure 3.1).

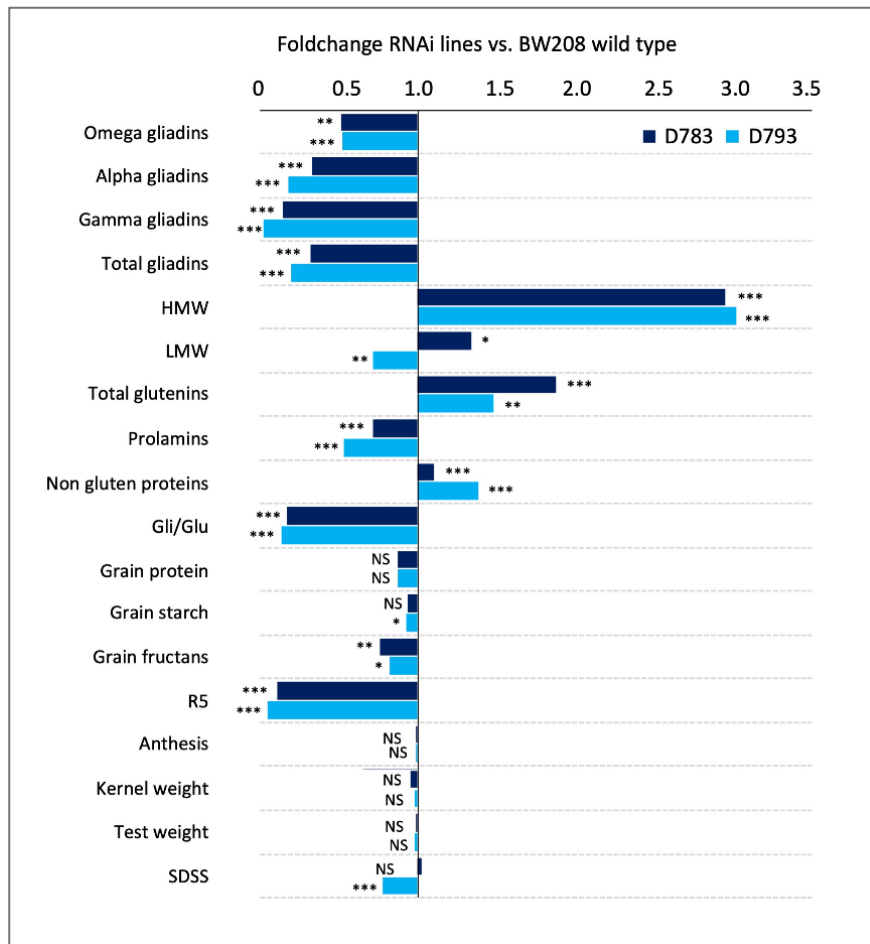


Figure 3.1. Fold change of grain components and other quality and agronomic features for RNAi lines in comparison to that of the wild type. HMW, high molecular weight glutenin subunits; LMW, low molecular weight glutenin subunits; Gli/Glu, gliadins to glutenins ratio; SDSS, sodium dodecyl sulphate sedimentation; R5, gluten content (mg kg^{-1}) determined by R5 monoclonal antibody. ANOVA was performed to determine statistically differences between genotypes for each component and feature. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. NS, Non-significant.



3.2. Transcriptomic analysis of grains from RNAi lines

An RNA-seq data analysis was carried out for the two RNAi lines using a pool of grains samples at 23–26 DAA. Read libraries from the three samples were mapped to the wheat reference genome and DGE analysis was performed to identify DE genes between the three lines. In the pipeline, 17.4, 19.2, and 18.8 million of reads from BW208, D783, and D793, respectively, were uniquely mapped to the reference genome (Supplementary Table 3.2). A total of 129,963 genes were considered: 110,790 known genes and 19,173 de novo assembly genes (identified as MSTRG name), of which 55,069, 75,773, and 52,144 were considered for DGE analysis in BW208 vs. D783, BW208 vs. D793 and D783 vs. D793 pair-wise comparisons, respectively (Supplementary Table 3.2). There were a higher number of DE genes for the BW208 vs. D783 comparison than for the BW208 vs. D793 one (Supplementary Figure 3.2; Supplementary Table 3.2). The GO enrichment analysis revealed that carbohydrate metabolic, response to stress or nutrient reservoir activity, among others, concentrate the higher number of DE genes, particularly for the BW208 vs. D783 comparison. In fact, carbohydrate metabolism term makes a remarkable difference between the RNAi lines in the number of DE genes, being this GO more enriched in D783 compared to the wild type (Figure 3.2). In contrast, nutrient reservoir activity was more enriched in D793 compared to the wild type (Figure 3.2). We also analyzed which GOs were enriched between up- and down-regulated genes in both RNAi lines compared to the wild type, to find out common responses between RNAi lines: (i) Nutrient reservoir activity was enriched in the set of up-regulated and down-regulated genes. (ii) However, response to stress was enriched in both RNAi lines among up-regulated genes while (iii) carbohydrate metabolic process was enriched among down-regulated genes (Supplementary Figure 3.3A, C). In the same way, we studied DE genes between D783 and D793 which were also DE in the BW208 vs. D793 comparison; nutrient reservoir activity was enriched among down-regulated genes because of the stronger silencing of gliadin genes in D793 line in comparison to that of D783 (Supplementary Figure 3.3B, D). A summary of DE genes in both RNAi lines with a high expression in almost one of the genotypes is provided in Supplementary Table 3.3. Among them, serine-type endopeptidase inhibitor (Serpins) genes were over-represented among up-regulated genes, and some non-specific lipid transfer proteins (ns-LTPs) genes were DE in both RNAi lines (Supplementary Table 3.3). According to the wheat reference genome, the major differences for gene expression (up or down-regulation) between RNAi lines and the wild type were observed for genes mapping in group 1, 6, 7 and Unassigned (Un) chromosomes, where gliadins and glutenins, among other genes, are located (Supplementary Figure 3.4). In contrast, the expression of serpins and ns-LTPs, mapping in group 4 and 5 chromosomes, was always higher in both RNAi lines (Supplementary Figure 3.4; Supplementary Table 3.4).



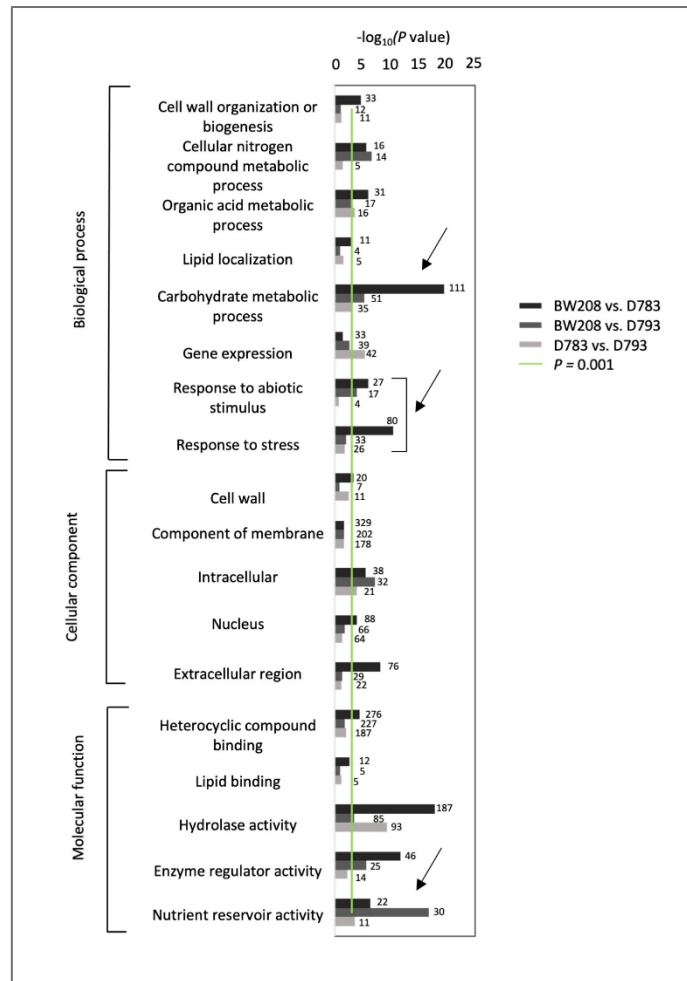


Figure 3.2. GO enrichment analysis of differentially expressed (DE) genes in RNAi lines. All domains (biological process, cellular component and molecular function) are represented in three pair-wise comparisons: BW208 vs. D783, BW208 vs. D793, and D783 vs. D793. First column shows the P -value of pair-wise comparisons based on hypergeometric distribution. Green line represents the threshold for statistical significance ($P < 0.001$) of GO terms. Numbers on each column represent the number of DE genes in the comparison. FC, Fold-change.

Next, hierarchical clustering of all DE genes allowed their gathering into four subclusters (Supplementary Figure 3.5), and the comparison of the expression patterns between the RNAi lines and the wild type. As shown, the wild type had clear differences for the expression pattern of DE genes to the RNAi lines. However, there are groups of genes in all subclusters with expression patterns notably different between the two RNAi lines, which could be due to different silencing pattern as they have different promoter driving the same RNAi fragment. Expression pattern in subcluster 1 showed an increased gene expression in D783 and D793, while a decreased gene expression was shown in subcluster 2 for both RNAi lines. Interestingly, GO enrichment analysis of genes grouped in subcluster 1 showed that they were mostly related to stress response, hydrolase activity and enzyme regulator activity, while in subcluster 2, nutrient reservoir activity was enriched.



3.3. Expression of prolamins and other grain protein genes

The expression of three different HMW genes (*1Bx7*, *1Dx5*, and *1Ax2*), encompassing the three bread wheat genomes, and three NGP genes were performed throughout grain development by qPCR analysis (Figure 3.3). The expression of all three HMW genes were higher in both RNAi lines from 14 to 26 DAA, particularly for line D793 with the maximum peak at 18 DAA. Data from RNA-seq analysis at 23–26 DAA also showed higher expression of the HMW genes for the RNAi lines (Supplementary Figure 3.6A, C). In addition, DE analysis of RNA-seq data confirmed the down-regulation of ω -, α/β -, and γ -gliadin genes in the RNAi lines (Supplementary Figure 3.6C).

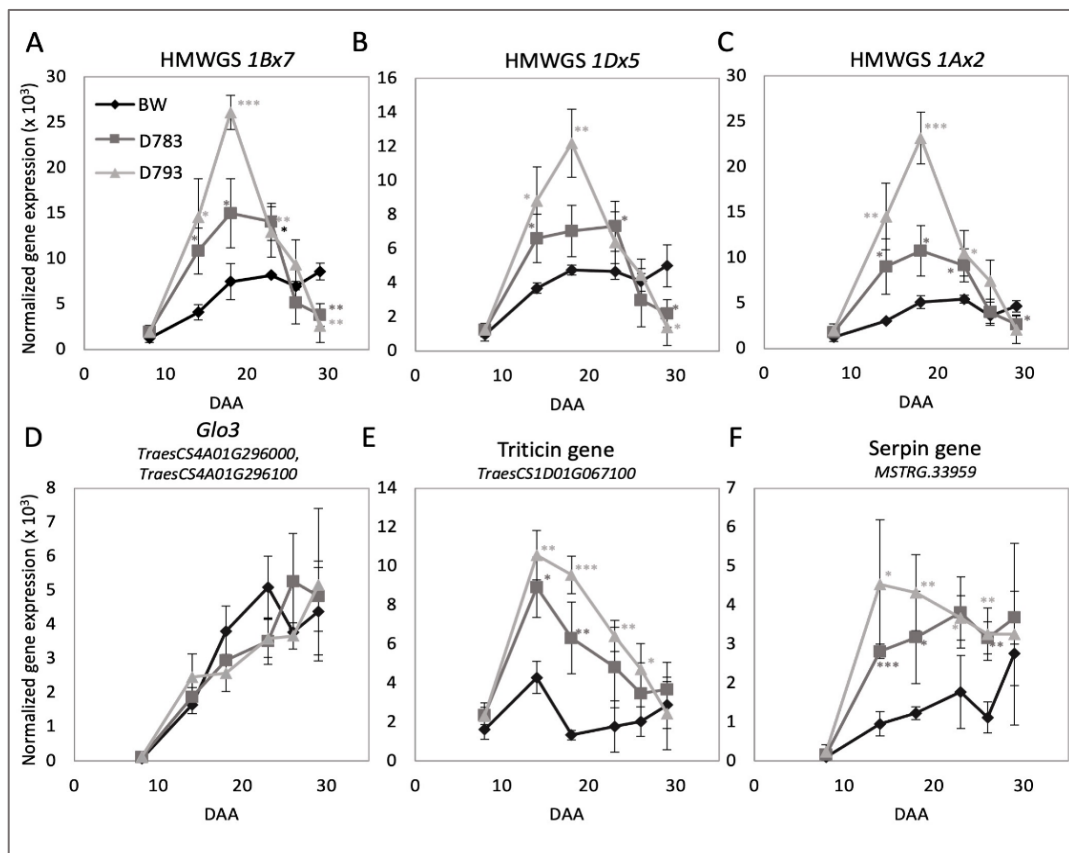


Figure 3.3. qPCR normalized expression values for three HMW glutenin subunit genes (*1Bx7*, *1Dx5*, and *1Ax2*), *Glo3* (*TraesCS4A01G296000*, *TraesCS4A01G296100*), triticins (*TraesCS1D01G067100*) and serpin (*MSTRG.33959*) genes for BW208, D783, and D793 genotypes through grain development. Mean values are represented, and standard deviation was calculated with three biological replicates. DAA, days after anthesis; HMW-GS, high molecular weight glutenin subunit; BW, BW208. ANOVA was performed for each DAA between each RNAi line and the wild type. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

On the other hand, the expression of triticin and serpin genes was higher in both RNAi lines from 14 to 26 DAA while the globulin gene expression was similar for all three lines (Figure 3.3D). Interestingly, the expression pattern of all three genes was



different throughout grain development. This increase in the expression of serpin and triticin genes observed by qPCR was confirmed in the RNA-seq data analysis in the case of serpin (Supplementary Table 3.3; Supplementary Figure 3.6A, C).

There are other genes encoding proteins of interest involved in triggering wheat-related human diseases as α -amylase/trypsin inhibitors (ATIs). Although there were no DE ATI genes in D783 and D793 lines compared to BW208 (Supplementary Figure 3.6C), it can be appreciated that there are differences in gene expression between genotypes (Supplementary Figure 3.6C).

Next, we examined the expression of transcription factors (TFs) related to prolamin genes. In Supplementary Table 3.5 a comprehensive list of TFs related to seed storage proteins found in the RNA-seq and their homologs/homeologs in barley, wheat or rice is provided. There was only one TF (*TraesCS2D01G117000*), homeolog to OsGZF1 from rice, that was significantly up-regulated in D783 in comparison to the wild type. In rice, this OsGZF1 TF participate in the regulation of the seed storage protein GluB-1 (Chen *et al.*, 2014).

3.4. Expression of genes related to stress response in RNAi lines

As showed previously, many genes related to stress response were enriched in the RNAi lines (Figure 3.2). For example, dehydrins and Rab genes (response to abiotic stimulus) were highly expressed in both RNAi lines (Supplementary Table 3.3). Among TFs related with defense response, or abiotic and biotic stress response with high fold-change (FC) in the RNAi lines, 12 were associated with response to abiotic, biotic and/or abscisic acid (ABA) (Table 3.1), and some of them were significantly up-regulated in both RNAi lines but, particularly in line D783.

Table 3.1. Putative transcription factors (TFs) related with defense response, and abiotic and biotic stress response.

Gene ID ^a	Homolog with known function		TMM normalized expression value ^b		
	Gene	Reported function	BW208	D783	D793
<i>TraesCS1A01G244800</i>	<i>OsDREB2C/OsERF44</i>	Expressed under drought and salinity conditions (Herath, 2016).	6.40	24.88	19.77
<i>TraesCS1B01G256000</i>	(<i>Os08g0565200</i>)		5.25	22.59	21.11
<i>TraesCS1D01G244500</i>			3.96	18.44	14.42
<i>TraesCS2D01G198100</i>	<i>TaWRKY78</i> (HM013818.1)	Induces the expression of PR4 genes involved in defense response (Proietti and Bertini, 2010).	2.20	5.43	6.74
<i>TraesCS6B01G060700</i>	<i>OsHHLH59</i> (<i>Os02g0116600</i>)	Induces the expression of xylanase inhibitors involved in defense	10.96	28.10	36.42



Gene ID ^a	Homolog with known function		TMM normalized expression value ^b		
	Gene	Reported function	BW208	D783	D793
		response (Zhan <i>et al.</i> , 2017).			
<i>TraesCS7A01G488600</i>	<i>OsbZIP20/RITA-</i>	Probably involved in	2.08	14.47	10.23
<i>TraesCS7B01G391800</i>	<i>1/RISBZ3</i> (<i>Os02g0266800</i>)	seed development and ABA-dependent stress response (Izawa <i>et al.</i> , 1994; Guo <i>et al.</i> , 2017).	0	31.07	18.89
<i>TraesCS5A01G468300</i>	<i>OsNAC19</i>	Induced by blast	2.00	8.95	5.39
<i>TraesCS5B01G480900</i>	(<i>Os03g0815100</i>)	fungus (<i>Magnaporthe grisea</i>), methyl jasmonate and ABA (Lin <i>et al.</i> , 2007).	1.28	4.88	3.61
<i>MSTRG.36868</i>	<i>SNAC1</i> (<i>Os03g0815100</i>)	Overexpression of <i>SNAC1</i> in rice enhances drought and salt tolerance, and ABA sensitivity (Hu <i>et al.</i> , 2006).	1.60	7.08	5.04
<i>TraesCS3A01G371900</i>	<i>OsABI5</i> (<i>Os01g0859300</i>)	Induced by ABA and	2.88	9.78	8.56
<i>TraesCS3A01G372200</i>		high salinity (Zou <i>et al.</i> , 2008).	4.96	30.82	13.82
<i>MSTRG.19588</i>	<i>TaVP1</i> (ABF48529.1) <i>HvVP1</i> (CAD24413.1)	Component of ABA signaling and regulates expression of seed maturation genes (Vasil <i>et al.</i> , 1995). Down-regulates seed storage proteins induced by HvGAMYB and BPBF (Abraham <i>et al.</i> , 2016).	1.09	7.68	8.52

^a Homoeologous genes are grouped. The requirement to be included was: log₂ FC >1.5 or <-1.5 in any of the RNAi lines compared to the wild type.

^b Genes with TMM-normalized expression value ≥5 in one of the RNAi lines are shown. Statistically significant expression values (P <0.001) respect to the wild type are indicated in bold. All of them present log₂ FC >1.5 in both RNAi lines compared to the wild type. Homoeologous in other species with their known function are indicated. TMM, Trimmed-Mean of M-values unit; FC, Fold-Change.

Next, we set up qPCR analysis for some of the DE genes related to stress throughout grain development (Figure 3.4). The expression of dehydrin (*Dhdn4*) gene was significantly higher in D793 for some of the last days of grain development and, from 23 DAA, D783 showed no significantly higher expression than that of the wild type (Figure 3.4 A). The expression of the ABA INSENSITIVE 5 (*ABI5*) TF was also analyzed because of its relation to ABA-dependent signaling pathway and its role in abiotic stress response (Zou *et al.*, 2008; Skubacz *et al.*, 2016). The expression of this gene was higher in both



RNAi lines compared to the BW208 line, but only significant for D793 (Figure 3.4 B). Finally, the basic Leucine Zipper (bZIP) TF (homolog to rice OsbZIP20/RITA-1/ RISBZ3) was reported as implicated in the ABA-dependent stress response (Izawa *et al.*, 1994; Guo *et al.*, 2017) and in the transcriptional activation of prolamin genes and other storage protein genes as globulin in rice (Onodera *et al.*, 2001). The RNA-seq and qPCR data analysis confirmed that this TF was up-regulated in both RNAi lines. This bZIP TF present higher expression levels in the RNAi lines than the wild type for most of the grain development stages, significantly higher for both RNAi lines at 18 and 26 DAA (Figure 3.4 C).

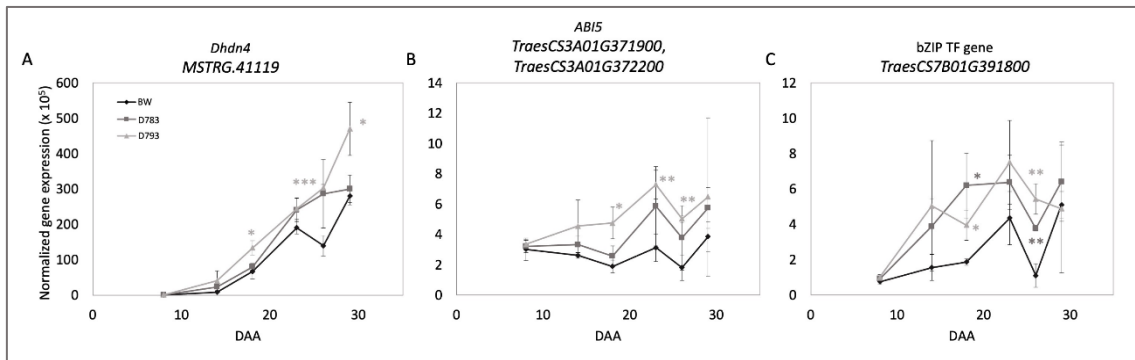


Figure 3.4. qPCR normalized expression values for *Dhdn4* (*dehydrin 4*) (*MSTRG.41119*), *ABI5* (*TraesCS3A01G371900*, *TraesCS3A01G372200*) and bZIP (*TraesCS7B01G391800*), ortholog to rice *OsZIP20/RITA-1/RISBZ3* through grain development. Mean values are represented, and standard deviation was calculated with three biological replicates. DAA, days after anthesis; *Dhdn4*, dehydrin 4; *ABI5*, ABA INSENSITIVE 5; BW, BW208. ANOVA was performed for each DAA between each RNAi line and the wild type. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

3.5. Expression of genes related to carbohydrate metabolism and other metabolic process in RNAi lines

Both RNAi lines show a slight decrease in the grain starch content, only significant for line D793 compared to the wild type (Figure 3.1). Carbohydrate metabolism process was enriched for both RNAi lines, especially in D783 with up to 111 DE genes (Figure 3.2). Some of the genes implicated in starch synthesis were up- or down-regulated in both RNAi lines, as starch synthase, beta-amylase, starch branching enzyme and sucrose synthase among others (Supplementary Table 3.3).

We also monitored the expression of three NAC TF genes for the RNAi and the wild type lines throughout seed development. In addition to response to stress, NAC TFs are also related to many important metabolic processes as seed development, starch synthesis and grain nitrogen concentration (Guérin *et al.*, 2019; Zhao *et al.*, 2015; Mathew *et al.*, 2016; Liu *et al.*, 2020). We designed the primers based on *TraesCS7B01G094000*, the homeologs *TraesCS7D01G154200* and *TraesCS7B01G056300*, and *TraesCS7A01G569100*. These genes are orthologs to *NAC92* (XM_020326243.1 and XM_020341606.1) from



Aegilops tauschii, to ANAC092 (AT5G39610.1) from *Arabidopsis thaliana*, and to ONAC20 (*Os01g0104500*) from rice. ONAC20 is related to seed development and seed size/weight (Mathew *et al.*, 2016) and NAC92 participates in many processes as regulation of seed germination, anther development, age-related resistance and stress response among others (Balazadeh *et al.*, 2010; Patil *et al.*, 2014). The expression of these NAC TFs is shown in Figure 3.5. For all of them, higher expression in the RNAi lines in comparison to that of the wild type was found during grain maturation, significant in many of the stages for both RNAi lines. In the RNA-seq data, the expression of three of them (*TraesCS7D01G154200*, *TraesCS7B01G056300*, and *TraesCS7A01G569100*) was also higher in the RNAi lines, although it was only significant for *TraesCS7D01G154200* in D783 compared to the wild type.

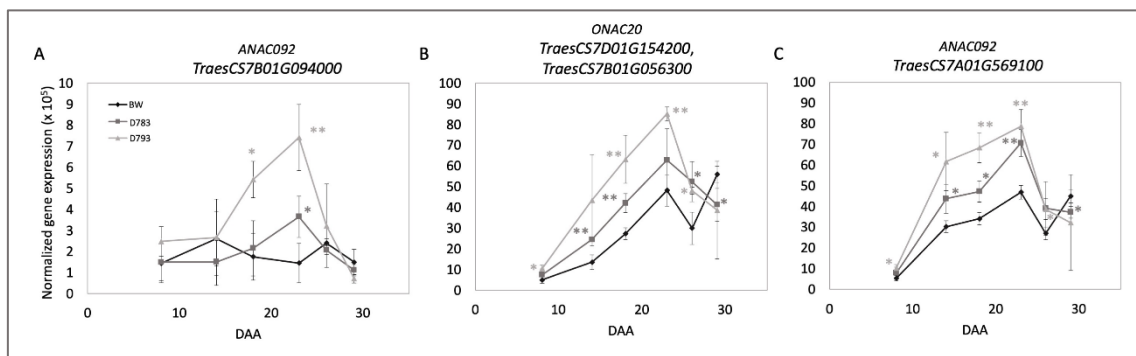


Figure 3.5. qPCR normalized expression values for three NAC TFs through grain development, orthologs to NAC92 (XM_020326243.1 and XM_020341606.1) from *Aegilops tauschii*, to ANAC092 (AT5G39610.1) from *Arabidopsis thaliana*, and to ONAC20 (*Os01g0104500*) from rice. Mean values are represented, and standard deviation was calculated with three biological replicates. DAA, days after anthesis; BW, BW208. ANOVA was performed for each DAA between each RNAi line and the wild type. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

4. Discussion

Obtaining wheat lines lacking the immunogenic component of gluten is a very appealing goal as an alternative to produce foods for people that need to follow a gluten-free diet. The RNAi technology has demonstrated to be highly effective in the downregulation of wheat gliadin genes (Gil-Humanes *et al.*, 2010). In addition to RNAi, CRISPR/Cas9 technology was also used for the reduction of wheat gliadins (Sánchez-León *et al.*, 2018). In this work, two RNAi lines, D783 and D793, with the same silencing fragment but different endosperm specific promoter, were evaluated for protein fractions, and transcriptomic features. Results reported here, confirmed the strong down-regulation of gliadin proteins in both RNAi lines (Gil-Humanes *et al.*, 2010; Gil-Humanes *et al.*, 2014). However, this down-regulation was lower for ω -gliadins, indicating that for this gliadin fraction the silencing construct is not as effective as for the α/β - and γ -gliadins. In fact, in a nitrogen fertilization study using these two RNAi lines, α/β - and γ -gliadins were silenced in both lines, even at high nitrogen concentrations.



However, the ω -gliadins were increased when increasing the nitrogen, confirming the lower effectiveness of silencing (García-Molina and Barro, 2017; Marín-Sanz *et al.*, 2020).

Wheat lines with down-regulation of one or more gliadin fractions were reported to present a compensatory mechanism involving other grain proteins (Altenbach *et al.*, 2019; García-Molina *et al.*, 2017) providing nitrogen contents comparable to that of the wild types. In this work, the increment of HMW in both RNAi lines is the major responsible of the increasing of the total glutenin content as the LMW had a different behavior in D783 and D793 lines, with a significant decrease for line D793. Similar changes were previously reported for RNAi lines with all the gliadin fractions down-regulated (García-Molina and Barro, 2017) where the readjustment of the protein composition is not the same among the RNAi lines. In the qPCR data, the increase of HMW transcripts abundance occurs at 18 DAA in both RNAi lines, and it could account for the higher HMW protein content observed in the grain for lines D783 and D793. This increment in the HMW glutenin content in both RNAi lines is important as they contribute to the viscoelastic properties of dough, and could be related to the unique bread-making quality features observed in these RNAi lines, particularly the tolerance to over-mixing (Gil-Humanes *et al.*, 2014). The synthesis of both LMW and HMW is essentially regulated at the transcriptional level, and therefore, differences for the HMW gene expression between RNAi lines and the wild type could be attributed to the transcriptional regulation. In the stage of development in which we have carried out the RNA-Seq, we have found that gene *TraesCS2D01G117000*, ortholog to *OsGZF1* in rice, and described as regulator of the seed storage protein GluB-1 (Chen *et al.*, 2014) was up-regulated in RNAi lines, particularly D783. In addition, we obtained higher expression of a bZIP wheat genes orthologs to *OsZIP20/RITA-1/RISBZ3* gene from rice, which is involved in transcriptional activation of prolamin genes and other storage protein genes as globulins (Onodera *et al.*, 2001).

For NGPs, such as globulins, serpins, and triticins, their increase was reported to compensates the down-regulation of gliadins to maintain the total grain protein content (Barro *et al.*, 2016; Gil-Humanes *et al.*, 2011). An increase in the number and size of the inclusions at the surface of protein bodies was also reported for the RNAi low gluten wheat lines (Gil-Humanes *et al.*, 2011). These inclusions contain triticins that raise the 5% content of total grain protein (Gil-Humanes *et al.*, 2011; Bechtel *et al.*, 1991). García-Molina *et al.* (2017) reported that serpin spots were increased in the RNAi low gluten wheat lines. The qPCR and RNA-seq analysis agreed to protein reported data and confirmed that the higher content in the NGPs is related to a higher gene expression during grain development. Serpins are related to wheat allergy as they contain Immunoglobulin E (IgE)-binding epitopes, and therefore, the increment in this protein fraction in the RNAi lines may indicate the increment in wheat allergy properties. Serpins have also a role in response to abiotic and biotic stresses (Roberts and Hejgaard, 2008) and they have been described as possible prolamin proteases inhibitors (Østergaard *et al.*, 2000) so it could be possible that a regulation mechanism against the decrease of gliadins has been activated.



In addition to grain proteins, there are other wheat grain components also responsible for allergies and intolerances. One of them are fructans (oligo- or polysaccharides with short chain of fructose units and terminal glucose), which are hydrolyzed partially in the intestine causing symptoms as bloating and abdominal pain. These belong to fermentable oligosaccharides, disaccharides, monosaccharides and polyols (FODMAPs) and are related to NCWS (Biesiekierski *et al.*, 2013). Fructans are used as reserve carbohydrates with similar function as starch and sucrose. Although genes encoding fructan degrading enzymes were not DE in both RNAi lines in comparison to that of the wild type, at least between 23 and 26 DAA, the fructan content in mature grains of RNAi lines was lower than the wild type, in the same way as total starch decreased. In the RNA-seq data analysis, carbohydrate metabolic process was enriched for both RNAi lines, and was more significant for D783 compared with the wild type, involving some DE carbohydrate synthesis enzymes genes such as sucrose synthase (SS), starch branching enzyme (SBE) and starch synthase (SSS). Interestingly, D783 and D793 have different behavior in terms of up- or down-regulated status of these genes compared to the wild type. This indicates that the variation in fructan and starch contents during kernel maturation could be part of signaling pathways to regulate carbohydrate metabolism and storage (Cimini *et al.*, 2015).

Other genes related to wheat pathologies are ns-LTPs and ATIs which show high reactivity with IgE related to baker's asthma (Juhász *et al.*, 2018). In this work, ns-LTPs were up-regulated in both RNAi lines while the expression of ATIs genes at 23–26 DAA was not affected by the silencing of gliadins. However, proteomic studies showed that ATIs were over-accumulated in RNAi lines, particularly in D793 line (García-Molina *et al.*, 2017). This clearly indicates that the expression data obtained by RNA-Seq at 23–26 DAA for ATIs are not predictive of the protein readjustments that finally operate in the grain. The increase in ATIs and serpins, both related to allergies, make necessary further studies, particularly in vivo studies to determine the possible impact of these lines on allergies.

The down-regulation of gliadins fractions also impacts to other metabolic processes. Two of the top enriched GOs in our analysis were response to abiotic and biotic stresses. Both of them comprise many genes as dehydrin, rab protein gene, the rRNA N-glycosidase, wheatwin-1 and subtilisin-chymotrypsin inhibitor WSCI among others, related to drought stress and plant defense against pests, and bacteria and fungi pathogens (Shakirova *et al.*, 2016; Endo and Tsurugi, 1998; Hamel *et al.*, 1998; Poerio *et al.*, 2003; Caporale *et al.*, 2004; Nawrot *et al.*, 2014; Zhu *et al.*, 2018). These genes were upregulated in both RNAi lines, mainly in D783. Moreover, some TFs that regulate the stress response, were also up-regulated in the RNAi lines, as DREB TFs that regulate dehydrin genes (Herath, 2016). Among these TFS, the ABI5 stands out as having a key role in the stress response in presence of ABA (Skubacz *et al.*, 2001) and it has higher gene expression in both RNAi lines, particularly in D793 line as revealed also by qPCR. ABA participates in the stress response by regulating the expression of many genes which have a key role in this response, being part of ABA-dependent stress response



process (Nakashima and Yamaguchi-Shinozaki, 2013). In the present work, some genes regulated by ABA were upregulated, which could suggest the implication of this phytohormone on the enrichment of stress response metabolic processes observed in the RNAi lines. For example, dehydrin genes has ABA-responsive elements (ABRE) motifs in their promoter sequence in wheat (Nordin *et al.*, 1993); the pathogen defense Wheatwin-1 (Caruso *et al.*, 1999; Bertini *et al.*, 2003) a class II chitinase of PR-4 family with a high sequence similarity with ZmPR4 protein from *Zea mays*, could be up-regulated by ABA, as it was reported that ABA can induce the expression of ZmPR4 (Bravo *et al.*, 2003); ABI5 and NAC TFs have a role in ABA-mediated response to abiotic stresses (Nakashima and Yamaguchi-Shinozaki, 2013; Skubacz *et al.*, 2016); the NAC TF coding gene *TraesCS5A01G468300*, up-regulated in RNAi lines, could be also involved in ABA-dependent response as its homolog in rice (*OsNAC19*) is induced by this phytohormone, providing resistance to blast fungus (Lin *et al.*, 2007). In addition, other grain components as ns-LTPs and HMW have been studied as genes transcriptionally regulated by this phytohormone in barley and wheat (Federico *et al.*, 2005; Wang *et al.*, 2013). In a recent RNA-seq analysis in wheat under heat stress condition, γ -gliadin and LMW were up-regulated across 13 and 30 DAA, suggesting a relation between these prolamin genes and the heat stress condition (Rangan *et al.*, 2020). However, in our experimental conditions we have not found an increment in the expression of 9-cis-epoxycarotenoid dioxygenase and ABA 8'-hydroxylase, main enzymes involved respectively in biosynthesis and catabolism of ABA (Nakashima and Yamaguchi-Shinozaki, 2013; Krochko *et al.*, 1998). Although activation of response to stress mechanisms could be a consequence of the down-regulation of gliadins by RNAi lines, more studies are required to determine the molecular and physiological basis of this response and the possible implication of ABA. However, the pre-activation of these stress response mechanisms in these lines is a very interesting aspect since it could provide them advantages in certain conditions where a rapid response to stress situations is needed.

5. Conclusions

Wheat lines D783 and D793 have a strong down-regulation in the gliadin fraction, which are compensated by NGPs and HMW glutenins since they are increased in D783 and D793 lines. This compensation with NGPs, particularly serpins and ATIs, could somehow increase allergies. However, these lines also had other transcriptomic changes compared to the wild type, involving enzyme regulator activity, carbohydrate metabolism, and response to stress. Carbohydrate metabolic process showed the downregulation of some key enzymes in starch biosynthesis that could provide lower starch content in RNAi lines. Finally, the silencing of wheat gliadins has a clear impact on the activation of genes and TFs related to the response to stress, leaving a much more complex metabolic restructuring, and that ultimately would also be responsible for the differences in agronomic traits such as kernel weight and/or yield. Many of these metabolic processes are regulated by ABA, which would indicate that this phytohormone may be involved.



6. References

- Abraham, Z., Iglesias-Fernández, R., Martínez, M., Rubio-Somoza, I., Díaz, I., Carbonero, P., et al. (2016). A developmental switch of gene expression in the barley seed mediated by HvVP1 (Viviparous-1) and HvGAMYB interactions. *Plant Physiol.* 170, 2146–2158. doi:10.1104/pp.16.00092.
- Altenbach, S. B., Chang, H. C., Yu, X. B., Seabourn, B. W., Green, P. H., and Alaedini, A. (2019). Elimination of omega-1,2 gliadins from bread wheat (*Triticum aestivum*) flour: Effects on immunogenic potential and end-use quality. *Front. Plant Sci.* 10, 580. doi:10.3389/fpls.2019.00580.
- Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., et al. (2004). UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* 32, D115–D119. doi:10.1093/nar/gkh131.
- Aziz, I., Dwivedi, K., and Sanders, D. S. (2016). From coeliac disease to noncoeliac gluten sensitivity; Should everyone be gluten free? *Curr. Opin. Gastroenterol.* 32, 120–127. doi:10.1097/MOG.0000000000000248.
- Balazadeh, S., Siddiqui, H., Allu, A. D., Matallana-Ramirez, L. P., Caldana, C., Mehrnia, M., et al. (2010). A gene regulatory network controlled by the NAC transcription factor ANAC092/AtNAC2/ORE1 during salt-promoted senescence. *PLANT J.* 62, 250–264. doi:10.1111/j.1365-313X.2010.04151.x.
- Barro, F., Iehisa, J. C. M., Giménez, M. J., García-Molina, M. D., Ozuna, C. V., Comino, I., et al. (2016). Targeting of prolamins by RNAi in bread wheat: Effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotechnol. J.* 14, 986–996. doi:10.1111/pbi.12455.
- BECHTEL, D. B., WILSON, J. D., and SHEWRY, P. R. (1991). IMMUNOCYTOCHEMICAL LOCALIZATION OF THE WHEAT STORAGE PROTEIN TRITICIN IN DEVELOPING ENDOSPERM TISSUE. *Cereal Chem.* 68, 573–577.
- Bertini, L., Leonardi, L., Caporale, C., Tucci, M., Cascone, N., Di Berardino, I., et al. (2003). Pathogen-responsive wheat PR4 genes are induced by activators of systemic acquired resistance and wounding. *PLANT Sci.* 164, 1067–1078. doi:10.1016/S0168-9452(03)00112-2.
- Biesiekierski, J. R., Peters, S. L., Newnham, E. D., Rosella, O., Muir, J. G., and Gibson, P. R. (2013). No effects of gluten in patients with self-reported non-celiac gluten sensitivity after dietary reduction of fermentable, poorly absorbed, short-chain carbohydrates. *Gastroenterology* 145, 320–328. doi:10.1053/j.gastro.2013.04.051.
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.
- Bravo, J. M., Campo, S., Murillo, I., Coca, M., and San Segundo, B. (2003). Fungus- and wound-induced accumulation of mRNA containing a class II chitinase of the pathogenesis-related protein 4 (PR-4) family of maize. *Plant Mol. Biol.* 52, 745–759. doi:10.1023/A:1025016416951.
- Caporale, C., Di Berardino, I., Leonardi, L., Bertini, L., Cascone, A., Buonocore, V., et al. (2004). Wheat pathogenesis-related proteins of class 4 have ribonuclease activity. *FEBS Lett.* 575, 71–76. doi:10.1016/j.febslet.2004.07.091.
- Caruso, C., Chilosi, G., Caporale, C., Leonardi, L., Bertini, L., Magro, P., et al. (1999). Induction of pathogenesis-related proteins in germinating wheat seeds infected with *Fusarium culmorum*. *PLANT Sci.* 140, 87–97. doi:10.1016/S0168-9452(98)00199-X.
- Chen, Y., Sun, A., Wang, M., Zhu, Z., and Ouwerkerk, P. B. F. (2014). Functions of the CCCH type zinc finger protein OsGZF1 in regulation of the seed storage protein GluB-1 from rice. *Plant Mol. Biol.* 84, 621–634. doi:10.1007/s11103-013-0158-5.



- Cimini, S., Locato, V., Vergauwen, R., Paradiso, A., Cecchini, C., Vandenpoel, L., et al. (2015). Fructan biosynthesis and degradation as part of plant metabolism controlling sugar fluxes during durum wheat kernel maturation. *Front. Plant Sci.* 6. doi:10.3389/fpls.2015.00089.
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635.
- Endo, Y., and Tsurugi, K. (1988). The RNA N-glycosidase activity of ricin A-chain – the characteristics of the enzymatic-activity of ricin A-chain with ribosomes and with ribosomal-RNA. *J. Biol. Chem.* 263, 8735–8739.
- Federico, M. L., Kaeppeler, H. F., and Skadsen, R. W. (2005). The complex developmental expression of a novel stress-responsive barley Ltp gene is determined by a shortened promoter sequence. *Plant Mol. Biol.* 57, 35–51. doi:10.1007/s11103-004-6769-0.
- Garcia-Molina, M. D., and Barro, F. (2017). Characterization of Changes in Gluten Proteins in Low-Gliadin Transgenic Wheat Lines in Response to Application of Different Nitrogen Regimes. *Front. Plant Sci.* 8. doi:10.3389/fpls.2017.00257.
- Garcia-Molina, M. D., Muccilli, V., Saletti, R., Foti, S., Masci, S., and Barro, F. (2017). Comparative proteomic analysis of two transgenic low-gliadin wheat lines and non-transgenic wheat control. *J. Proteomics* 165, 102–112. doi:10.1016/j.jprot.2017.06.010.
- Gil-Humanes, J., Pistón, F., Altamirano-Fortoul, R., Real, A., Comino, I., Sousa, C., et al. (2014). Reduced-gliadin wheat bread: An alternative to the gluten-free diet for consumers suffering gluten-related pathologies. *PLoS One* 9, e90898. doi:10.1371/journal.pone.0090898.
- Gil-Humanes, J., Pistón, F., Barro, F., and Rosell, C. M. (2014). The shutdown of celiac disease-related gliadin epitopes in bread wheat by RNAi provides flours with increased stability and better tolerance to over-mixing. *PLoS One* 9, e91931. doi:10.1371/journal.pone.0091931.
- Gil-Humanes, J., Pistón, F., Shewry, P. R., Tosi, P., and Barro, F. (2011). Suppression of gliadins results in altered protein body morphology in wheat. *J. Exp. Bot.* 62, 4203–4213. doi:10.1093/jxb/err119.
- Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L. M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proc. Natl. Acad. Sci. U. S. A.* 107, 17023–17028. doi:10.1073/pnas.1007773107.
- Guerin, C., Roche, J., Allard, V., Ravel, C., Mouzeyar, S., and Bouzidi, M. F. (2019). Genome-wide analysis, expansion and expression of the NAC family under drought and heat stresses in bread wheat (*T. aestivum* L.). *PLoS One* 14, e0213390.
- Guo, Y., Huang, R., Duan, L., and Wang, J. (2017). The APETALA2/ethylene-responsive factor transcription factor OsDERF2 negatively modulates drought stress in rice by repressing abscisic acid responsive genes. *J. Agric. Sci.* 155, 966–977. doi:10.1017/S0021859617000041.
- Hamel, F., Breton, C., and Houde, M. (1998). Isolation and characterization of wheat aluminum-regulated genes: Possible involvement of aluminum as a pathogenesis response elicitor. *Planta* 205, 531–538. doi:10.1007/s004250050352.
- Haro, C., Villatoro, M., Vaquero, L., Pastor, J., Giménez, M. J., Ozuna, C. V., et al. (2018). The dietary intervention of transgenic low-gliadin wheat bread in patients with non-celiac gluten sensitivity (NCGS) showed no differences with gluten free diet (GFD) but provides better gut microbiota profile. *Nutrients* 10, 1964. doi:10.3390/nu10121964.
- Herath, V. (2016). Small family, big impact: In silico analysis of DREB2 transcription factor family in rice. *Comput. Biol. Chem.* 65, 128–139. doi:10.1016/j.compbiolchem.2016.10.012.



- Hu, H., Dai, M., Yao, J., Xiao, B., Li, X., Zhang, Q., et al. (2006). Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proc. Natl. Acad. Sci. U. S. A.* 103, 12987–12992. doi:10.1073/pnas.0604882103.
- Izawa, T., Foster, R., Nakajima, M., Shimamoto, K., and Chua, N. H. (1994). The rice bZIP transcriptional activator RITA-1 is highly expressed during seed development. *Plant Cell* 6, 1277–1287. doi:10.2307/3869825.
- Juhasz, A., Belova, T., Florides, C. G., Maulis, C., Fischer, I., Gell, G., et al. (2018). Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci. Adv.* 4. doi:10.1126/sciadv.aar8602.
- Kiarash, J. G., Wilde, H. D., Amirmahani, F., Moemeni, M. M., Zaboli, M., Nazari, M., et al. (2018). Selection and validation of reference genes for normalization of qRT-PCR gene expression in wheat (*Triticum durum* L.) under drought and salt stresses. *J. Genet.* 97, 1433–1444. doi:10.1007/s12041-018-1042-5.
- Krochko, J. E., Abrams, G. D., Loewen, M. K., Abrams, S. R., and Cutler, A. J. (1998). (+)-abscisic acid 8'-hydroxylase is a cytochrome P450 monooxygenase. *Plant Physiol.* 118, 849–860. doi:10.1104/pp.118.3.849.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *BIOINFORMATICS* 25, 1754–1760. doi:10.1093/bioinformatics/btp324.
- Lin, R., Zhao, W., Meng, X., Wang, M., and Peng, Y. (2007). Rice gene OsNAC19 encodes a novel NAC-domain transcription factor and responds to infection by *Magnaporthe grisea*. *Plant Sci.* 172, 120–130. doi:10.1016/j.plantsci.2006.07.019.
- Liu, Y. C., Hou, J., Wang, X. L., Li, T., Majeed, U., Hao, C. Y., et al. (2020). The NAC transcription factor NAC019-A1 is a negative regulator of starch synthesis in wheat developing endosperm. *J. Exp. Bot.* 71, 5794–5807. doi:10.1093/jxb/eraa333.
- Ludvigsson, J. F., Bai, J. C., Biagi, F., Card, T. R., Ciacci, C., Ciclitira, P. J., et al. (2014). Diagnosis and management of adult coeliac disease: Guidelines from the British society of gastroenterology. *Gut* 63, 1210–1228. doi:10.1136/gutjnl-2013-306578.
- Ludvigsson, J. F., Leffler, D. A., Bai, J. C., Biagi, F., Fasano, A., Green, P. H. R., et al. (2013). The Oslo definitions for coeliac disease and related terms. *Gut* 62, 43–52. doi:10.1136/gutjnl-2011-301346.
- Marin-Sanz, M., Gimenez, M. J., Barro, F., and Savin, R. (2020). Prolamin Content and Grain Weight in RNAi Silenced Wheat Lines Under Different Conditions of Temperature and Nitrogen Availability. *Front. Plant Sci.* 11. doi:10.3389/fpls.2020.00314.
- Marín-Sanz, M., Iehisa, J. C. M., and Barro, F. (2022). New transcriptomic insights in two RNAi wheat lines with the gliadins strongly down-regulated by two endosperm specific promoters. *Crop J.* 10, 194–203.
- Mathew, I. E., Das, S., Mahto, A., and Agarwal, P. (2016). Three rice NAC transcription factors heteromerize and are associated with seed size. *Front. Plant Sci.* 7, 1638. doi:10.3389/fpls.2016.01638.
- Mayer, K., Rogers, J., Dole el, J., Wincker, P., Keller, B., and Praud, S. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* (80-.). 345, 1251788–1251788.
- Nakashima, K., and Yamaguchi-Shinozaki, K. (2013). ABA signaling in stress-response and seed development. *Plant Cell Rep.* 32, 959–970. doi:10.1007/s00299-013-1418-1.
- Nawrot, R., Barylski, J., Nowicki, G., Broniarczyk, J., Buchwald, W., and Gozdzicka-Jozefiak, A. (2014). Plant antimicrobial peptides. *Folia Microbiol. (Praha)*. 59, 181–196. doi:10.1007/s12223-013-0280-4.



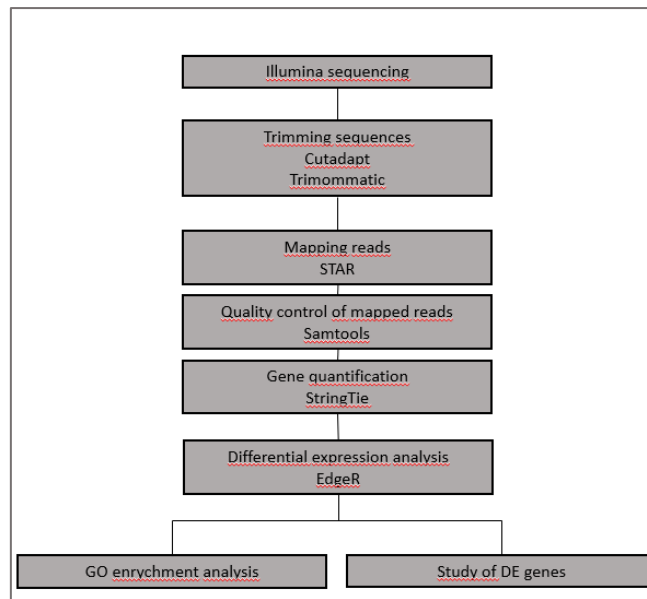
- Nordin, K., Vahala, T., and Palva, E. T. (1993). Differential expression of 2 related, low-temperature-induced genes in *Arabidopsis thaliana* (L) Heynh. *PLANT Mol. Biol.* 21, 641–653. doi:10.1007/BF00014547.
- Onodera, Y., Suzuki, A., Wu, C. Y., Washida, H., and Takaiwa, F. (2001). A Rice Functional Transcriptional Activator, RISBZ1, Responsible for Endosperm-specific Expression of Storage Protein Genes through GCN4 Motif. *J. Biol. Chem.* 276, 14139–14152. doi:10.1074/jbc.M007405200.
- Østergaard, H., Rasmussen, S. K., Roberts, T. H., and Hejgaard, J. (2000). Inhibitory serpins from wheat grain with reactive centers resembling glutamine-rich repeats of prolamin storage proteins - Cloning and characterization of five major molecular forms. *J. Biol. Chem.* 275, 33272–33279. doi:10.1074/jbc.M004633200.
- Patil, M., Ramu, S. V., Jathish, P., Sreevathsa, R., Reddy, P. C., Prasad, T. G., et al. (2014). Overexpression of AtNAC2 (ANAC092) in groundnut (*Arachis hypogaea* L.) improves abiotic stress tolerance. *Plant Biotechnol. Rep.* 8, 161–169. doi:10.1007/s11816-013-0305-0.
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi:10.1038/nbt.3122.
- Piston, F., Gil-Humanes, J., Rodriguez-Quijano, M., and Barro, F. (2011). Down-Regulating gamma-Gliadins in Bread Wheat Leads to Non-Specific Increases in Other Gluten Proteins and Has No Major Effect on Dough Gluten Strength. *PLoS One* 6. doi:10.1371/journal.pone.0024754.
- Piston, F., Leon, E., Lazzeri, P. A., and Barro, F. (2008). Isolation of two storage protein promoters from *Hordeum chilense* and characterization of their expression patterns in transgenic wheat. *EUPHYTICA* 162, 371–379. doi:10.1007/s10681-007-9530-3.
- Piston, F., Marin, S., Hernando, A., and Barro, F. (2009). Analysis of the activity of a gamma-gliadin promoter in transgenic wheat and characterization of gliadin synthesis in wheat by MALDI-TOF during grain development. *Mol. Breed.* 23, 655–667. doi:10.1007/s11032-009-9263-1.
- Poerio, E., Di Gennaro, S., Di Maro, A., Farisei, F., Ferranti, P., and Parente, A. (2003). Primary structure and reactive site of a novel wheat proteinase inhibitor of subtilisin and chymotrypsin. *Biol. Chem.* 384, 295–304. doi:10.1515/BC.2003.033.
- Proietti, S., Bertini, L., Van Der Ent, S., Leon-Reyes, A., Pieterse, C. M. J., Tucci, M., et al. (2011). Cross activity of orthologous WRKY transcription factors in wheat and *Arabidopsis*. *J. Exp. Bot.* 62, 1975–1990. doi:10.1093/jxb/erq396.
- Rangan, P., Furtado, A., and Henry, R. (2020). Transcriptome profiling of wheat genotypes under heat stress during grain-filling. *J. Cereal Sci.* 91. doi:10.1016/j.jcs.2019.102895.
- Roberts, T. H., and Hejgaard, J. (2008). Serpins in plants and green algae. *Funct. Integr. Genomics* 8, 1–27. doi:10.1007/s10142-007-0059-2.
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.
- Ruijter, J. M., Ramakers, C., Hoogaars, W. M. H., Karlen, Y., Bakker, O., van den hoff, M. J. B., et al. (2009). Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* 37, e45–e45. doi:10.1093/nar/gkp045.
- Sanchez-Leon, S., Gil-Humanes, J., Ozuna, C. V., Gimenez, M. J., Sousa, C., Voytas, D. F., et al. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *PLANT Biotechnol. J.* 16, 902–910. doi:10.1111/pbi.12837.



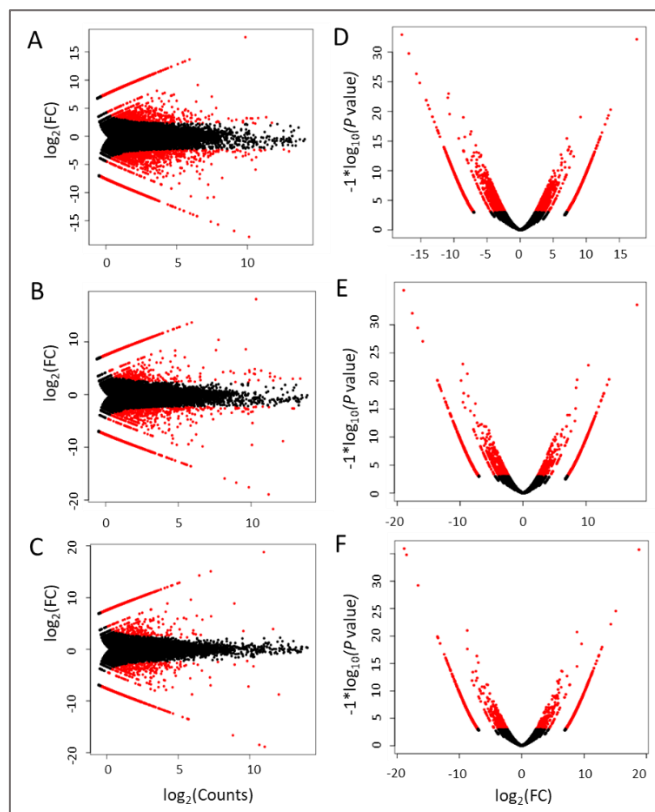
- Shakirova, F., Allagulova, C., Maslennikova, D., Fedorova, K., Yuldashev, R., Lubyanova, A., et al. (2016). Involvement of dehydrins in 24-epibrassinolide-induced protection of wheat plants against drought stress. *Plant Physiol. Biochem.* 108, 539–548. doi:10.1016/j.plaphy.2016.07.013.
- Shewry, P. R., and Halford, N. G. (2002). Cereal seed storage proteins: Structures, properties and role in grain utilization. *J. Exp. Bot.* 53, 947–958. doi:10.1093/jexbot/53.370.947.
- Skodje, G. I., Sarna, V. K., Minelle, I. H., Rolfsen, K. L., Muir, J. G., Gibson, P. R., et al. (2018). Fructan, Rather Than Gluten, Induces Symptoms in Patients With Self-Reported Non-Celiac Gluten Sensitivity. *Gastroenterology* 154, 529–539.e2. doi:10.1053/j.gastro.2017.10.040.
- Skubacz, A., Daszkowska-Golec, A., and Szarejko, I. (2016). The role and regulation of ABI5 (ABA-insensitive 5) in plant development, abiotic stress responses and phytohormone crosstalk. *Front. Plant Sci.* 7, 1884. doi:10.3389/fpls.2016.01884.
- Sollid, L. M. (2002). Coeliac disease: Dissecting a complex inflammatory disorder. *Nat. Rev. Immunol.* 2, 647–655. doi:10.1038/nri885.
- Valdes, I., Garcia, E., Llorente, M., and Mendez, E. (2003). Innovative approach to low-level gluten determination in foods using a novel sandwich enzyme-linked immunosorbent assay protocol. *Eur. J. Gastroenterol. Hepatol.* 15, 465–474. doi:10.1097/01.meg.0000059119.41030.df.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, research0034. 1. doi:10.1186/gb-2002-3-7-research0034.
- Vasil, V., Marcotte, W. R., Rosenkrans, L., Cocciolone, S. M., Vasil, I. K., Quatrano, R. S., et al. (1995). Overlap of viviparous1 (VP1) and abscisic acid response elements in the Em promoter: G-box elements are sufficient but not necessary for VP1 transactivation. *Plant Cell* 7, 1511–1518. doi:10.2307/3870139.
- Wang, K., Zhang, X., Zhao, Y., Chen, F., and Xia, G. (2013). Structure, variation and expression analysis of glutenin gene promoters from *Triticum aestivum* cultivar Chinese Spring shows the distal region of promoter 1Bx7 is key regulatory sequence. *Gene* 527, 484–490. doi:10.1016/j.gene.2013.06.068.
- Zhan, Y., Sun, X., Rong, G., Hou, C., Huang, Y., Jiang, D., et al. (2017). Identification of two transcription factors activating the expression of OsXIP in rice defence response. *BMC Biotechnol.* 17, 26. doi:10.1186/s12896-017-0344-7.
- Zhao, D., Derkx, A. P., Liu, D. C., Buchner, P., and Hawkesford, M. J. (2015). Overexpression of a NAC transcription factor delays leaf senescence and increases grain nitrogen concentration in wheat. *PLANT Biol.* 17, 904–913. doi:10.1111/plb.12296.
- Zhu, F., Zhou, Y. K., Ji, Z. L., and Chen, X. R. (2018). The plant ribosome-inactivating proteins play important roles in defense against pathogens and insect pest attacks. *Front. Plant Sci.* 9, 146. doi:10.3389/fpls.2018.00146.
- Zou, M., Guan, Y., Ren, H., Zhang, F., and Chen, F. (2008). A bZIP transcription factor, OsABI5, is involved in rice fertility and stress tolerance. *Plant Mol. Biol.* 66, 675–683. doi:10.1007/s11103-008-9298-4.



7. Supplementary Materials

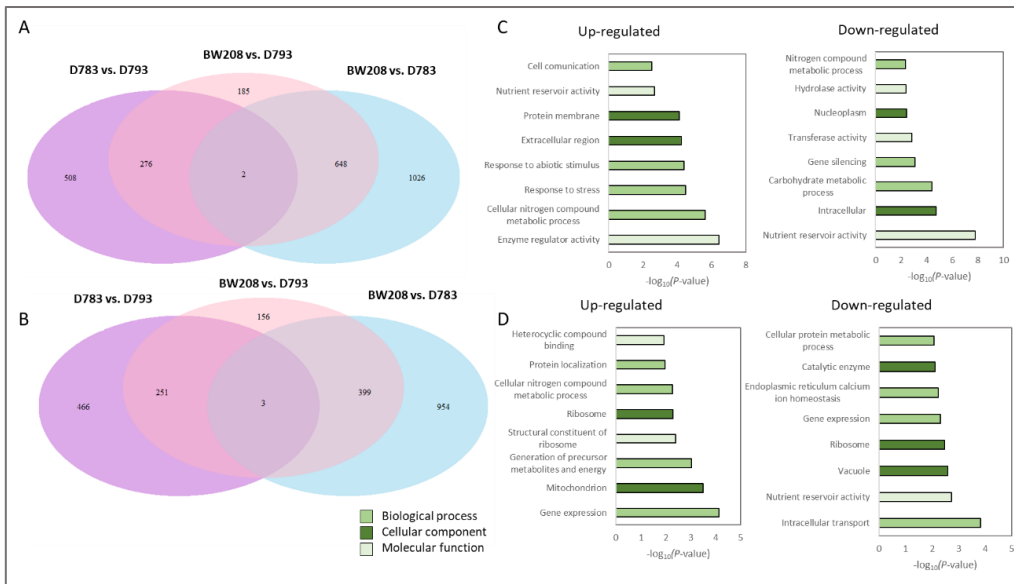


Supplementary Figure 3.1. Pipeline of RNA-seq data analysis. DE, differentially expressed; GO, gene ontology.

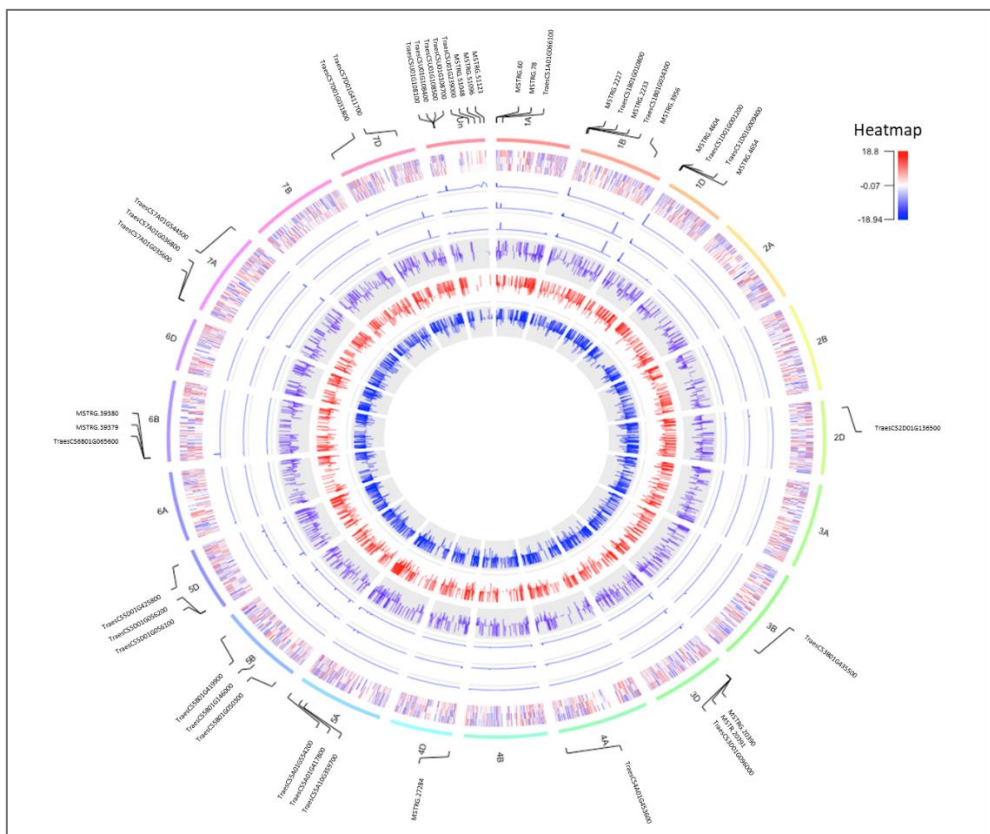


Supplementary Figure 3.2. MA (A–C) and volcano (D–F) plots of pair-wise comparisons between genes of BW208, D783, and D793 in RNA-seq data analysis. (A, D) BW208 vs. D783, (B, E) BW208 vs. D793, and (C, F) D783 vs. D793 comparisons. Red points represent differentially expressed (DE) genes ($P < 0.001$) between two genotypes. MA, plot shows the distribution of genes by the logarithm of the fold-change (FC) (M-values) and the mean of normalized counts (A-values) of the genotypes compared two by two.





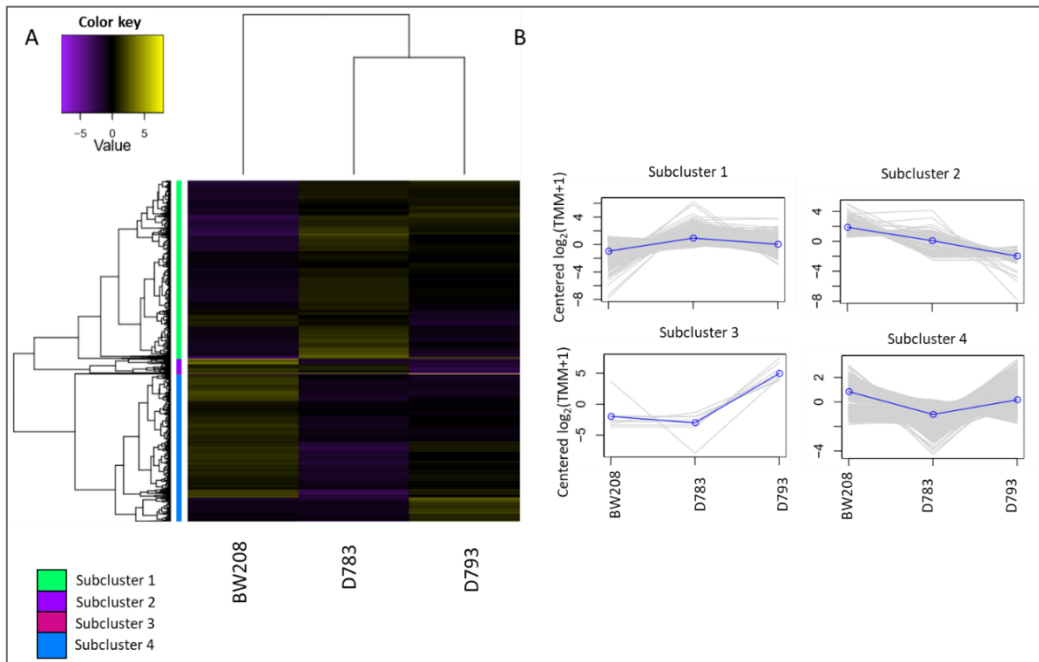
Supplementary Figure 3.3. Venn diagram of the number of genes A) up- and B) down-regulated between the three pair-wise comparisons and top 8 of most enriched gene ontology (GO) terms C) that are up- or down-regulated genes in both RNAi D783 and D793 compared to BW208, and D) up- and down-regulated genes in D793 compared to D783 and BW208.



Supplementary Figure 3.4. Transcriptome profiles throughout wheat genome for the three genotypes showing differences for gene expression between both RNAi lines and the wild type. From outside to inside: Labels of genes with expression greater than 1500 trimmed-mean of M-values (TMM)-normalized read counts in any of the three genotypes; chromosomes labels; heatmap of $\log_2(\text{FC})$ in three pair-wise comparisons (BW208 vs. D783, BW208 vs. D793, and D783 vs. D793).

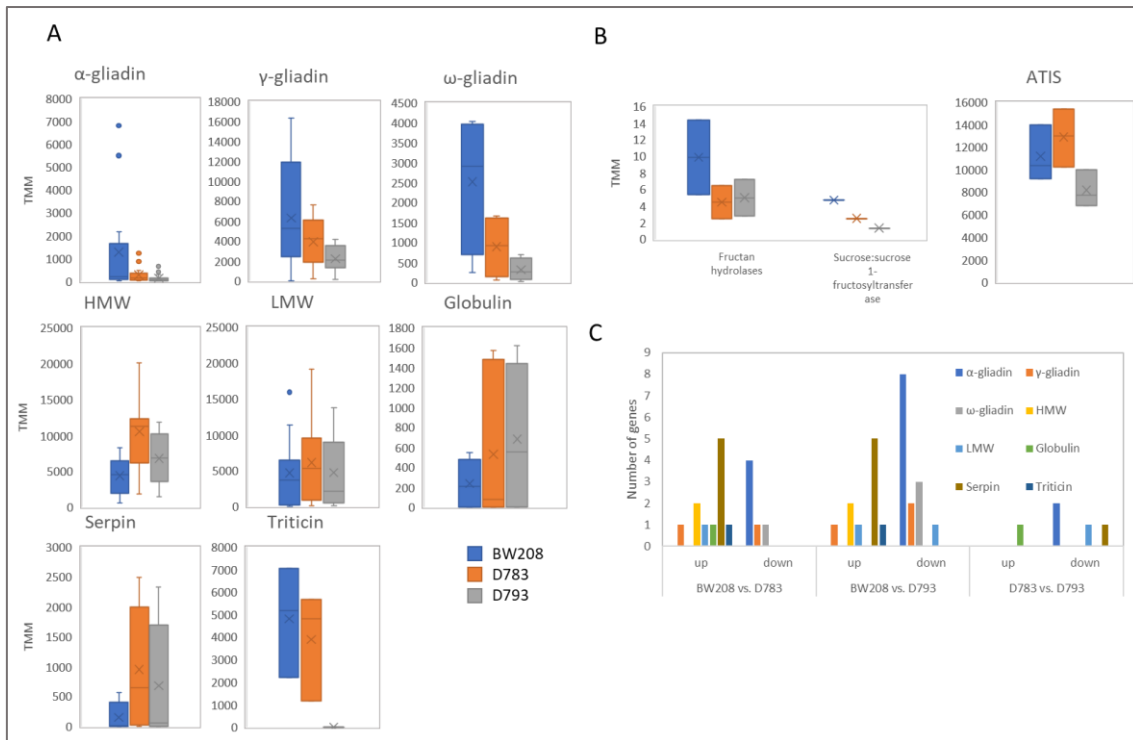


vs. D793); TMM-normalized gene expression values of BW208, D783 and D793; P -values of BW208 vs. D783 comparison with cut-off in $P = 0.001$ (no Differentially expressed (DE) gene: $P > 0.001$; DE gene: $P < 0.001$); P -value of genes in BW208 vs. D783 comparison; P -value D783 vs. D793 comparison. FC, Fold change.



Supplementary Figure 3.5. Heatmap of expression pattern of DE gene. (A) Heatmap of gene expression patterns of differentially expressed (DE) genes ($P < 0.001$ and $FC > 4$) for BW208, D783 and D793. Hierarchical clustering was performed through genotypes (up) and genes (left). DE genes were clustered into four subclusters according to their expression pattern among the three genotypes (left). Color Key substitutes the variation of gene-expression value respect to the average gene-expression value of the three genotypes. (B) Expression pattern of DE gene ($P < 0.001$ and $FC > 4$) between BW208, D783 and D793 in each subcluster. The grey lines represent centered log₂ (TMM + 1) expression values of genes in the genotypes and the blue line represents mean expression profile for that cluster. TMM, trimmed-mean of M-values; FC, Fold change.





Supplementary Figure 3.6. Box plots of TMM values for prolamin genes and other genes related to wheat allergies and intolerances annotated by blastx and blastp. (A) Gliadin fractions, glutenin fractions, globulins, serpins and triticins. (B) Fructan hydrolases and Sucrose: sucrose 1-fructosyltransferase and ATIS. (C) Number of genes up- and down-regulated in each pair-wise comparison for each proteins group ($P < 0.001$). TMM, trimmed-mean of M-values; HMW, high molecular weight; LMW, low molecular weight; ATIS, α -amylase/trypsin inhibitors; Up, up-regulated; Down, down-regulated.

Supplementary Table 3.1. List of primers used in this study for gene expression analysis by qPCR.

Gene	Description	Primer sequence (5'-3')	Reference
<i>RLI</i>	RNase L inhibitor	TGAGCAACTCATGGACCAG GCTTTCCAAGGCACAAACAT	Giménez et al. (2011)
<i>CDC</i>	Cell division control protein	CAGCTGCTGACTGAGATGGA ATGTCTGGCCTGTTGGTAGC	Giménez et al. (2011)
<i>ADP-RF</i>	ADP-ribosylation factor	TTCATGGTTGGTCTCGATG GGATGGTGGTGACGATCTCT	Giménez et al. (2011)
<i>MSTRG.41119</i>	Dehydrin	ACTGACACCGGCGAGAAG GCTCACGTTGTATGGTGAA	
<i>TraesCS3A01G372200</i>	ABI5 transcription factor	GGTGGTTTATGTTCCGGCTCA TGAAGTGTATGGGCGACAGA	
<i>TraesCS3A01G371900</i>	Triticin	TGACCCTGTACGGATTGTGA TATTGCCATCCGTCGTGTA	
<i>TraesCS4A01G296100</i>	Globulin	ATCGCCAAGATTCTCCACAC GGACGATGGAGATGGACTTC	
<i>TraesCS4A01G296000</i>	bZIP transcription factor homolog to rice	AGCGAGTCATTGGTCGAGAT GGGTTGCACTTCGCTTTTA	
<i>TraesCS7B01G391800</i>	<i>OsZIP20/RITA-1/RISBZ3</i>		
<i>HMWGS_Ax2</i>	HMWGS Ax2 allele	CGGGACAAGGGCAATCAG TGTCTGGTTGCTGTCTTTG	



Gene	Description	Primer sequence (5'-3')	Reference
<i>HMWGS_1Dx5</i>	HMWGS 1Dx5 allele	GCAACAAGGTCAGCAGGT GCTGATTGTTGCAATTGTCC	
<i>HMWGS_1Bx7</i>	HMWGS 1Bx7 allele	ACAAGGGCAACAAGATCAGC GTCCTGGCTGCTGTGAAGTT	
<i>HMWGS_y</i>	HMWGS y-type	GGGGAACATCTTCACAAACAG TGTCCCAGTTCTTGCCATTT	
<i>TraesCS7A01G569100</i>	NAC transcription factor	GCGATTGAGGATGTGAACCT TGGCTGGAAGATCTCCTTGT	
<i>TraesCS7B01G094000</i>	NAC transcription factor	GGGTGATAAGGGGCGATACT TCGTGCATGATCCAGTTTGT	
<i>TraesCS7D01G154200</i>	NAC transcription factor	ACAAAGCATTTCGACCCATA TTCCAGTGGCCTTCCAGTAG	
<i>TraesCS7B01G056300</i>	Serine Protease Inhibitor (Serpine)	GACACTTCCGGTGTGGTTCT TTGCATGGGTAGCATGATTG	

ABI5, ABA INSENSITIVE 5; HMWGS, high molecular weight glutenin subunit.

Reference:

Giménez, M.J., Pistón, F., Atienza, S.G. (2011). Identification of suitable reference genes for normalization of qPCR data in comparative transcriptomics analyses in the Triticeae. *Planta* 233, 163–173.

Supplementary Table 3.2. RNA-seq data.

Software	BW208	D783	D793
Cutadapt			
Total reads processed	28,216,500	29,022,596	27,632,977
Reads with adapters	16,807 (0.1%) ^a	18,065 (0.1%)	22,298 (0.1%)
Reads written (passing filters)	28,216,500 (100.0%)	29,022,596 (100.0%)	27,632,977 (100.0%)
Trimmomatic			
Input Reads	28,216,500	29,022,596	27,632,977
Surviving reads	27,787,302 (98.48%)	28,579,248 (98.47%)	27,218,750 (98.50%)
Dropped reads	429,198 (1.52%)	443,348 (1.53%) ^a	414,227 (1.50%) ^a
STAR			
Number of Uniquely mapped reads	17,491,200	19,297,932	18,868,256
Uniquely mapped reads (%)	62.95%	67.52%	69.32%
Number of reads mapped to multiple loci	8,859,108	7,771,443	7,102,236
Reads mapped to multiple loci (%)	31.88%	27.19%	26.09%
Reads unmapped: too many mismatches (%)	0.00%	0.00%	0.00%
Reads unmapped: too short (%)	2.25%	2.84%	2.48%
Reads unmapped: other (%)	2.92%	2.45%	2.10%
Number of chimeric reads	0	0	0
StringTie			
Number of genes in gtf file	110,790		
Number of novel genes merged	19,173		
edgeR			
Pair-wise comparisons	BW208 vs. D783	BW208 vs. D793	D783 vs. D793
Number of genes considered in DGE	55,069	75,773	52,144
Number of up-regulated genes	1676	1111	786
Number of down-regulated genes	1356	809	720

Results for trimming, filtering, mapping, number of genes identified, and number of genes differentially expressed (DE). Illumina sequencing was carried out using an Illumina GAIIx platform, providing 33 M 75 bp single-end (SE) per genotype.

^a Percentage of low quality reads removed.



Supplementary Table 3.3. List of differentially expressed (DE) genes in at least one genotype and one pair-wise comparison for gene ontology (GO) terms at three domains.

Up level	Intermedium level	Base level	BP	GO	Protein coded by DE genes	Related terms in other	BW208 vs. D783	BW208 vs. D793
Biological process (BP)								
Response to stimulus	Response to abiotic stimulus	Response to water			Dehydrin (Dehydrin 1) (Salt-induced YSK2 dehydrin 1)		Up-regulated***	Up-regulated***
Response to stimulus	Response to abiotic stimulus	Response to water			Rab protein		Up-regulated***	Up-regulated***
Response to stimulus	Response to stress	Defense response to bacterium; defense response to fungus; killing of cells of other organism; plant-type hypersensitive response			Wheatwin-1	MF: nucelase activity	Up-regulated***	Up-regulated***
Response to stimulus	Response to stress	Response to wounding			Subtilisin-chymotrypsin inhibitor WSCI	MF: serine-type endopeptidase inhibitor activity	Up-regulated***	Up-regulated***
Response to stimulus	Response to stress	Response to wounding			Proteins with serine-type endopeptidase inhibitor activity	MF: serine-type endopeptidase inhibitor activity	Up-regulated***	Up-regulated***
Response to stimulus	Response to stress	Defense response			Defensin		Up-regulated***	Up-regulated***
Response to stimulus	Response to stress	Defense response			Type V Thionin		Up-regulated***	Up-regulated***
Response to stimulus	Response to stress	Defense response; negative regulation of translation			rRNA N-glycosidase (EC 3.2.2.22)	MF: rRNA N-glycosylase activity; toxin activity	Up-regulated***	Up-regulated***
Metabolic process	Biosynthetic process	Biosynthetic process			Aspartate aminotransferase (EC 2.6.1.1)	MF: L-aspartate:2-oxoglutarate aminotransferase activity; pyridoxal phosphate binding	Up-regulated***	Up-regulated***
Cellular process	Organic acid metabolic process	Cellular amino acid metabolic process						



Up level	Intermedium level	Base level	BP	GO	Protein coded by DE genes	Related terms in other	BW208 vs. D783	BW208 vs. D793
Cellular process	Cellular compound metabolic process	nitrogen metabolic	Asparagine biosynthetic process	metabolic process	Asparagine synthetase-2 (EC 6.3.5.4)	MF: asparagine synthase (glutamine-hydrolyzing) activity; ATP binding	Up-regulated***	Up-regulated***
Cellular process	Cellular compound metabolic process	nitrogen metabolic	Positive regulation of translational elongation	positive regulation of translational termination; translational frameshifting	Eukaryotic translation initiation factor 5A (eIF-5A)	MF: ribosome binding; translation elongation factor activity; translation initiation factor activity	Non-significant	Up-regulated***
Cellular process	Cellular compound metabolic process	nitrogen metabolic	Defense response; negative regulation of translation		Protein synthesis inhibitor II	MF: rRNA N-glycosylase activity; toxin activity	Up-regulated***	Up-regulated***
Localization	Lipid localization		Lipid transport		Non-specific lipid-transfer protein	MF: lipid binding	Up-regulated***	Up-regulated***
Metabolic process	Gene expression		Translation		40S ribosomal protein S7	CC: ribosome MF: structural constituent of ribosome	Up-regulated***	Non-significant
Metabolic process	Gene expression		Translation		60S ribosomal protein L28-1	CC: ribosome MF: structural constituent of ribosome	Non-significant	Down-regulated***
Metabolic process	Gene expression		Translational elongation		Ribosomal protein P1	CC: ribosome MF: structural constituent of ribosome	Non-significant	Down-regulated***
Metabolic process	Carbohydrate metabolic process		Polysaccharide catabolic process		Beta-amylase (EC 3.2.1.2)	MF: amylopectin maltohydrolase activity; beta-amylase activity	Down-regulated***	Up-regulated***
Metabolic process	Carbohydrate metabolic process		Starch biosynthetic process		Starch synthase, chloroplast/amyloplastic (EC 2.4.1.-)	CC: amyloplast; chloroplast MF: glycogen (starch) synthase activity	Down-regulated***	Non-significant



Up level	Intermedium level	Base level BP	GO	Protein coded by DE genes	Related terms in other	BW208 vs. D783	BW208 vs. D793
Metabolic process	Carbohydrate metabolic process	Glycogen biosynthetic process;	starch biosynthetic process	Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) (ADP-glucose pyrophosphorylase)	CC: chloroplast MF: ATP binding; glucose-1-phosphate adenylyltransferase activity	Non-significant	Up-regulated***
Metabolic process	Carbohydrate metabolic process	Glycogen biosynthetic process		Starch branching enzyme (EC 2.4.1.18)	CC: plastid MF: 1,4-alpha-glucan branching enzyme activity; 1,4-alpha-glucan branching enzyme activity (using a glucosylated glycogenin as primer for glycogen synthesis); cation binding; hydrolase activity, hydrolyzing O-glycosyl compounds	Down-regulated***	Down-regulated***
Metabolic process	Carbohydrate metabolic process	Inositol catabolic process		Uncharacterized protein	CC: cytoplasm MF: inositol oxygenase activity; iron ion binding	Up-regulated***	Up-regulated*
Metabolic process	Carbohydrate metabolic process	Inositol biosynthetic process;	phospholipid biosynthetic process	Myo-inositol 1-phosphate synthase	MF: inositol-3-phosphate synthase activity	Up-regulated***	Non-significant
Metabolic process	Carbohydrate metabolic process	Sucrose metabolic process		Sucrose synthase (EC 2.4.1.13)	MF: sucrose synthase activity	Down-regulated***	Non-significant
Metabolic process	Carbohydrate metabolic process	Carbohydrate metabolic process		Beta-glucosidase 44	MF: hydrolase activity, hydrolyzing O-glycosyl compounds	Up-regulated***	Up-regulated***
Metabolic process	Carbohydrate metabolic process	Trehalose biosynthetic process		Uncharacterized protein	MF: catalytic activity	Up-regulated***	Up-regulated**
Cellular process	Cell wall organization or biogenesis	Cell wall organization		Pectin acetyltransferase (EC 3.1.1.-)	CC: cell wall; extracellular region; integral component of membrane MF: hydrolase activity	Up-regulated***	Up-regulated***



Up level	Intermedium level	Base level	BP	GO	Protein coded by DE genes	Related terms in other	BW208 vs. D783	BW208 vs. D793
Cellular process	Cell wall organization or biogenesis	Xylan catabolic process			Xylanase inhibitor TAXI-IV	MF: aspartic-type endopeptidase activity	Up-regulated***	Up-regulated***
Cellular process	Cell wall organization or biogenesis	Cell wall organization; lipid transport			Non-specific lipid-transfer protein 2G (LTP2G) (7 kDa lipid transfer protein 1) (Lipid transfer protein 2 isoform 1) (LTP2-1)	CC: cell wall; extracellular region MF: lipid binding	Down-regulated***	Down-regulated***
Cellular component (CC)								
Membrane	Component membrane	of	Integral component of membrane		ABA induced plasma membrane protein PM 19		Up-regulated***	Up-regulated***
Membrane	Component membrane	of	Integral component of membrane		Uncharacterized protein		Up-regulated*	Up-regulated***
Membrane	Component membrane	of	Integral component of membrane		Uncharacterized protein		Up-regulated***	Up-regulated***
Membrane	Component membrane	of	Integral component of membrane; intracellular		Uncharacterized protein	MF: GTP binding	Down-regulated***	Non-significant
Membrane	Component membrane	of	Integral component of membrane		Uncharacterized protein		Up-regulated**	Up-regulated***
Membrane	Component membrane	of	Integral component of membrane		Uncharacterized protein		Down-regulated***	Down-regulated***
Extracellular region	Extracellular region	Extracellular region			Wali6 protein	MF: serine-type endopeptidase activity	Up-regulated***	Up-regulated***
Extracellular region	Extracellular region	Extracellular region			Peroxidase (EC 1.11.1.7)	BP: Hydrogen peroxide catabolic process; response to oxidative stress MF: heme binding; metal ion binding; peroxidase activity	Up-regulated*	Non-significant
Extracellular region	Extracellular region	Extracellular space			Serpin-Z1C (TriaeZ1c) (WSZ1c)	MF: serine-type endopeptidase activity	Up-regulated***	Up-regulated***
Extracellular region	Extracellular region	Extracellular space			Serpin-Z2A (TriaeZ2a) (WSZ2a)	MF: serine-type endopeptidase activity	Up-regulated***	Up-regulated***



Up level		Intermedium level		Base level	BP	GO	Protein coded by DE genes	Related terms in other	BW208 vs. D783	BW208 vs. D793
Extracellular region		Extracellular region		Extracellular space			Serpin-N3.2		Up-regulated***	Up-regulated***
Extracellular region		Extracellular region		Extracellular space			Serpin-Z1A (TriaeZ1a) (WSZ1a) (WSZ1) (WSZCI)	MF: serine-type endopeptidase activity	Up-regulated***	Up-regulated***
Extracellular region		Extracellular region		Extracellular space			Serpin-Z2B		Up-regulated***	Up-regulated***
Cell		Cytoplasm		Cytoplasm; microtubule			Tubulin beta chain	BP: microtubule-based process MF: GTPase activity; GTP binding; structural constituent of cytoskeleton	Down-regulated***	Non-significant
Cell		Cytoplasm		Cytoplasm; microtubule			Uncharacterized protein	BP: microtubule-based process MF: GTPase activity; GTP binding; structural constituent of cytoskeleton	Down-regulated***	Non-significant
Cell		Cytoplasm		Cytoplasm			Uncharacterized protein	MF: GTPase activity; GTP binding; translation elongation factor activity	Up-regulated***	Non-significant
Cell		Cytoplasm		Cytoplasm			Elongation factor 1-alpha	MF: GTPase activity; GTP binding; translation elongation factor activity	Down-regulated***	Non-significant
Cell		Nucleus		Nucleosome; nucleus			Histone H1	BP: nucleosome assembly MF: DNA binding	Up-regulated***	Up-regulated***
Molecular function (MF)										
Nutrient activity	reservoir	Nutrient activity	reservoir	Nutrient activity	reservoir		ω -gliadin		Non-significant	Non-significant
Nutrient activity	reservoir	Nutrient activity	reservoir	Nutrient activity	reservoir		α -gliadin		Down-regulated***	Down-regulated***
Nutrient activity	reservoir	Nutrient activity	reservoir	Nutrient activity	reservoir		γ -gliadin		Down-regulated***	Down-regulated***



Up level	Intermedium level		Base level	BP	GO	Protein coded by DE genes	Related terms in other	BW208 vs. D783	BW208 vs. D793
Nutrient activity	reservoir	Nutrient activity	reservoir	Nutrient activity	reservoir	HMW		Up-regulated**	Up-regulated*
Nutrient activity	reservoir	Nutrient activity	reservoir	Nutrient activity	reservoir	LMW		Up-regulated***	Up-regulated***
Nutrient activity	reservoir	Nutrient activity	reservoir	Nutrient activity	reservoir	Globulin		Up-regulated*	Up-regulated*
Binding		Heterocyclic compound binding		DNA binding		Stress response protein		Up-regulated*	Up-regulated*
Binding		Ion binding		Iron ion binding; oxidoreductase activity		Uncharacterized protein	BP: lipid biosynthetic process	Up-regulated***	Up-regulated***
Binding		Ion binding		Zinc ion binding		EC protein III (Zinc metallothionein class II)		Up-regulated***	Up-regulated***
Binding		Ion binding		Zinc ion binding		EC protein I/II (Zinc metallothionein class II)		Up-regulated***	Up-regulated**
Binding		Ion binding		DNA binding; zinc ion binding		Stress response protein		Up-regulated***	Up-regulated***
Catalytic activity		Catalytic activity, acting on a protein		Peptidase activity		Vacuolar-processing enzyme		Down-regulated***	Non-significant

GO term at base-level and included in the two higher GO levels are indicated in bold. Proteins related to stress response are up-regulated in both RNAi lines, while those related to carbohydrate metabolic process have different behavior in the RNAi lines. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.



Supplementary Table 3.4. According to reference genome, high expressed genes at least in one genotype were identified among the chromosomes.

ID gene	Protein coded by gene	Chr.	FC and signification in pair-wise comparisons		
			BW208 vs. D783	BW208 vs. D793	D783 vs. D793
<i>MSTRG.60</i>	ω -gliadin	1A	-3.17	-4.83	-1.66
<i>MSTRG.78</i>	LMW glutenin subunit	1A	0.53	-3.39	-3.91
<i>gene:TraesCS1A01G066100</i>	Triticin	1A	2.35	2.63	0.29
<i>MSTRG.2227</i>	γ -gliadin	1B	-2.26	-2.71	-0.45
<i>gene:TraesCS1B01G010800</i>	γ -gliadin	1B	-1.82	-2.98	-1.16
<i>MSTRG.2233</i>	ω -gliadin B3	1B	-0.21	-2.29	-2.07
<i>gene:TraesCS1B01G034300</i>	Serine proteinase inhibitor-like allergen	1B	2.74	2.65	-0.08
<i>MSTRG.3956</i>	HMW	1B	2.67	2.66	-0.02
<i>MSTRG.4604</i>	γ -gliadin	1D	-1.92	-2.66	0.74
<i>gene:TraesCS1D01G001200</i>	γ -gliadin	1D	-1.55	-3.02	-1.46
<i>gene:TraesCS1D01G009400</i>	LMW glutenin subunit	1D	0.1	8.84	8.75
<i>MSTRG.4654</i>	γ -gliadin	1D	15.68	15.92	0.25
<i>gene:TraesCS2D01G136500</i>	β -glucosidase 4	2D	3.72	2.8	-0.92
<i>gene:TraesCS3B01G435500</i>	Thiamine thiazole synthase, chloroplastic	3B	2.42	1.84	0.58
<i>MSTRG.20390</i>	Dimeric α -amylase inhibitor	3D	-17.62	0.96	18.49
<i>MSTRG.20391</i>	β -amylase	3D	17.88	17.58	-0.29
<i>gene:TraesCS3D01G096000</i>	Dimeric α -amylase inhibitor	3D	-2.11	0.16	2.27
<i>gene:TraesCS4A01G453600</i>	ns-LTP	4A	1.95	2.39	0.45
<i>MSTRG.27284</i>	Serpin-Z2A	4D	2.58	1.34	-1.24
<i>gene:TraesCS5A01G359700</i>	Serpin-Z1C	5A	4.19	3.64	-0.55
<i>gene:TraesCS5A01G417800</i>	Serpin-Z2B	5A	5.92	5.42	0.51
<i>gene:TraesCS5A01G554200</i>	β -amylase	5A	NA	18.94	18.86
<i>gene:TraesCS5B01G050300</i>	Serpin-N3.2	5B	2.52	3.45	0.94
<i>gene:TraesCS5B01G146000</i>	ns-LTP	5B	3.39	2.31	-1.08
<i>gene:TraesCS5B01G419900</i>	Serpin-Z2A	5B	2.4	2.29	-0.11
<i>gene:TraesCS5D01G056100</i>	Uncharacterized gene	5D	1.72	3.24	1.52
<i>gene:TraesCS5D01G056200</i>	Defensin	5D	2.09	3.31	1.22
<i>gene:TraesCS5D01G425800</i>	Serpin-Z2B	5D	2.44	2.39	-0.04
<i>gene:TraesCS6B01G065600</i>	α -gliadin	6B	-2.06	-2.74	-0.68
<i>MSTRG.39379</i>	α -gliadin	6B	-3.19	-4.26	-1.07
<i>MSTRG.39380</i>	α -gliadin	6B	-2.77	-4.09	-1.32
<i>gene:TraesCS7A01G035600</i>	α/β -gliadin	7A	1.95	2.35	0.4
<i>gene:TraesCS7A01G036800</i>	AAI domain-containing protein	7A	1.03	2.58	1.54
<i>gene:TraesCS7A01G544500</i>	α -gliadin	7A	2.65	2.79	1.13
<i>gene:TraesCS7D01G031800</i>	α -gliadin	7D	0.8	-18.08	-18.8
<i>gene:TraesCS7D01G411700</i>	Defensin-like	7D	2.08	2.32	0.26
<i>gene:TraesCSU01G108100</i>	α -gliadin	Unassigned	-2.18	-3.06	-0.88
<i>gene:TraesCSU01G108400</i>	α/β -gliadin	Unassigned	-3.24	-4.53	-1.29
<i>gene:TraesCSU01G108500</i>	α -gliadin	Unassigned	-2.39	-3.87	-1.48
<i>gene:TraesCSU01G108700</i>	α -gliadin	Unassigned	-2.4	-3.38	0.19
<i>gene:TraesCSU01G239000</i>	α -gliadin	Unassigned	-3.24	-4.59	-1.35
<i>MSTRG.51048</i>	ω -gliadin	Unassigned	-3.4	-4.02	-0.62
<i>MSTRG.51096</i>	α -gliadin	Unassigned	-2.73	-3.91	1.18
<i>MSTRG.51123</i>	α -amylase inhibitor 0.19	Unassigned	-3.19	-3.96	-0.77

These labels are represented in Supplementary Figure 3.4. The FC in bold indicates that this gene was differential expressed ($P < 0.001$) for DEG analysis in each pair-wise comparison. HMW, high molecular weight; LMW, low molecular weight; ns-LTP, non-specific lipid transfer protein; AAI, α -amylase inhibitor; FC, fold-change; Chr., chromosome.



Supplementary Table 3.5. Trimmed-Mean of M-values (TMM)-normalized expression values in RNAi and BW208 the wild type lines for wheat genes related to Transcription Factors (TFs) (and their homologous barley, wheat or rice) involved in the regulation of the expression of seed storage proteins.

Gene ID ^a	Reported gene/species ^b	TMM-normalized expression value ^c		
		BW208	D783	D793
<i>TraesCS3A01G249100</i>	<i>HvFUACA3</i> (AM418838.1)/ <i>Hordeum</i>	5.56	1.44	3.85
<i>TraesCS3B01G278000</i>	<i>vulgare</i>	2.90	0.95	2.98
<i>TraesCS3D01G249100</i>		4.86	2.08	4.34
<i>TraesCS1A01G329900</i>	<i>TaSPA</i> (Y09013.1)/ <i>T. aestivum</i>	0.20	0.19	0.12
<i>TraesCS1B01G343500</i>		9.22	2.05	4.51
<i>TraesCS1D01G332200</i>		4.92	3.59	4.17
<i>TraesCS5A01G155900</i>	<i>BPBF</i> (AJ000991.1)/ <i>H. vulgare</i>	50.03	22.79	34.08
<i>TraesCS5B01G154100</i>		0.00	0.01	0.00
<i>TraesCS5D01G161000</i>		0.03	0.03	0.00
<i>TraesCS3A01G336500</i>	<i>HvGAMYB</i> (X87690.1)/ <i>H. vulgare</i>	18.75	15.96	14.17
<i>TraesCS3B01G367500</i>		0.00	0.00	0.44
<i>TraesCS3D01G329400</i>		7.02	7.71	6.35
<i>TraesCS5A01G440400</i>	<i>HvBLZ1</i> (CAA56374.1)/ <i>H. vulgare</i>	7.05	6.70	6.73
<i>TraesCS5B01G444100</i>		6.64	5.98	7.24
<i>TraesCS5D01G447500</i>		3.81	5.91	4.00
<i>TraesCS6A01G255500</i>	<i>HvSAD</i> (CAC85739.1)/ <i>H. vulgare</i>	0.00	0.32	0.16
<i>TraesCS6B01G270100</i>		0.29	0.30	0.20
<i>TraesCS6D01G236700</i>		0.25	0.70	0.51
<i>TraesCS1A01G218800</i>	<i>HvMCB1</i> (CAI84066.1)/ <i>H. vulgare</i>	40.77	27.11	30.45
<i>TraesCS1B01G232300</i>		21.67	12.39	14.61
<i>TraesCS1D01G220500</i>		18.87	12.45	11.20
<i>TraesCS1A01G219400</i>	<i>HvMYBS3</i> (CAJ53899.1)/ <i>H. vulgare</i>	2.42	1.56	1.56
<i>TraesCS1B01G232800</i>		4.63	5.76	3.42
<i>TraesCS1D01G221000</i>		2.71	2.64	2.28
<i>TraesCS6A01G335400</i>	<i>TaNfYA-1</i> (KM078747.1)/ <i>T. aestivum</i>	3.48	2.30	2.73
<i>TraesCS6B01G366100</i>		2.86	2.02	1.55
<i>TraesCS6D01G315300</i>		1.50	1.83	2.16
<i>TraesCS6A01G108300</i>	<i>TaNAM-1</i> (DQ869675.1)/ <i>T. aestivum</i>	1.26	1.07	0.60
<i>TraesCS2B01G228900</i>		0.65	0.32	0.55
<i>TraesCS6D01G096300</i>		0.98	1.42	1.12
<i>TraesCS2A01G118100</i>	<i>OsGZF1</i> (Os07g0668600)/ <i>Oryza sativa</i>	22.75	14.88	24.44
<i>MSTRG.9826</i>		9.95	11.27	9.39
<i>TraesCS2B01G138100</i>		25.45	23.15	26.57
<i>TraesCS2D01G116900</i>		43.70	44.44	41.20
<i>TraesCS2A01G118000</i>		9.06	35.11	16.77
<i>TraesCS2B01G138000</i>		2.99	10.62	7.31
<i>TraesCS2D01G117000</i>		5.28	<u>30.00</u>	14.82

^a Putative homoeologous genes are grouped.

^b Accession number or locus ID of the known genes are indicated in the parenthesis.

^c Statistically significant expression values ($P < 0.001$) with respect to the wild type are indicated with underline.





CHAPTER IV

RNAi silencing of wheat gliadins alters the network of transcription factors that regulate the synthesis of seed storage proteins toward maintaining grain protein levels

Published in:

Marín-Sanz, M., and Barro, F. (2022). RNAi silencing of wheat gliadins alters the network of transcription factors that regulate the synthesis of seed storage proteins toward maintaining grain protein levels. *Front. Plant Sci.* 13. doi.org/10.3389/fpls.2022.935851



Abstract

Gluten proteins are responsible for the unique viscoelastic properties of wheat dough, but they also trigger the immune response in celiac disease patients. RNA interference (RNAi) wheat lines with strongly silenced gliadins were obtained to reduce the immunogenic response of wheat. The E82 line presents the highest reduction of gluten, but other grain proteins increased, maintaining a total nitrogen content comparable to that of the wild type. To better understand the regulatory mechanisms in response to gliadin silencing, we carried out a transcriptomic analysis of grain and leaf tissues of the E82 line during grain filling. A network of candidate transcription factors (TFs) that regulates the synthesis of the seed storage proteins (SSPs), α -amylase/trypsin inhibitors, lipid transfer proteins, serpins, and starch in the grain was obtained. Moreover, there were a high number of differentially expressed genes in the leaf of E82, where processes such as nutrient availability and transport were enriched. The source-sink communication between leaf and grain showed that many down-regulated genes were related to protease activity, amino acid and sugar metabolism, and their transport. In the leaf, specific proline transporters and lysine-histidine transporters were down- and up-regulated, respectively. Overall, the silencing of gliadins in the RNAi line is compensated mainly with lysine-rich globulins, which are not related to the proposed candidate network of TFs, suggesting that these proteins are regulated independently of the other SSPs. Results reported here can explain the protein compensation mechanisms and contribute to decipher the complex TF network operating during grain filling.

Keywords: RNA-seq, prolamin regulation, starch, protein compensation, TF network, RNAi silencing, source-sink communication



Resumen

Las proteínas del gluten son responsables de las propiedades viscoelásticas únicas de la masa de trigo, pero también desencadenan la respuesta inmune en pacientes celíacos. Líneas de trigo de ARN de interferencia (ARNi) con un fuerte silenciamiento de gliadinas se obtuvieron para reducir la respuesta inmunogénica del trigo. La línea E82 presenta la mayor reducción de gluten, pero otras proteínas del grano incrementan, manteniendo el contenido de nitrógeno total comparable al de la línea control. Para entender mejor los mecanismos de regulación en respuesta al silenciamiento de gliadinas, llevamos a cabo un análisis transcriptómico del tejido de grano y hoja de la línea E82 durante el llenado del grano. Se obtuvo una red de factores de transcripción (FTs) candidatos que regula la síntesis de las proteínas de reserva de la semilla (SSPs), los inhibidores de α -amilasa/tripsina, las proteínas de transferencia de lípidos, serpinas, y almidón en el grano. Además, hubo un gran número de genes diferencialmente expresados en la hoja de E82, donde los procesos como la disponibilidad de nutrientes y su transporte fueron enriquecidos. La comunicación fuente-sumidero entre la hoja y el grano mostró que muchos genes infraexpresados estaban relacionados con la actividad de las proteasas, el metabolismo de los amino ácidos y de los azúcares, y su transporte. En la hoja, transportadores específicos de prolina y lisina-histidina estaban infra y sobreexpresados, respectivamente. En general, el silenciamiento de gliadinas en la línea ARNi está compensado principalmente por globulinas ricas en lisina, que no están relacionadas con la red de FTs candidatos propuesta, sugiriendo que estas proteínas son reguladas independientemente de las otras SSPs. Los resultados obtenidos en este trabajo pueden explicar los mecanismos de compensación de proteína y contribuir a descifrar la compleja red de FTs que opera durante el llenado del grano.

Palabras clave: RNA-seq, regulación de prolaminas, almidón, compensación de proteínas, red de FT, silenciamiento por ARNi, comunicación fuente-sumidero



1. Introduction

Wheat is one of the most important staple foods worldwide. It has unique viscoelastic properties for breadmaking due to its grain protein composition, allowing the production of many foods like bread and pasta (Shewry, 2009). However, wheat is also related to adverse reactions in humans, particularly proteins harboring epitopes that can trigger the immune response in celiac disease (CD), a chronic enteropathy mediated by human leukocyte antigen (HLA) HLA-DQ2 and HLA-DQ8 (Sollid, 2002). In addition to CD, there are other pathologies related to wheat consumption, such as Non-Celiac Wheat Sensitivity (NCWS), and wheat allergy, which includes baker's asthma and Wheat Dependent Exercise Induced Anaphylaxis (WDEIA) (Aziz et al., 2016).

Gluten proteins are composed of monomeric and polymeric proteins, gliadins and glutenins, respectively. Gliadins are soluble in alcoholic solutions and are divided into three structural groups; ω -, α/β - and γ -gliadins based on their mobility in A-PAGE electrophoresis gels. Glutenins are the polymeric fraction, insoluble in alcoholic solutions and classified into High Molecular Weight (HMW) and Low Molecular Weight (LMW) glutenin subunits, based on their mobility in SDS-PAGE electrophoresis gels (Shewry and Halford, 2002).

The development of wheat varieties lacking the proteins responsible for triggering adverse reactions to wheat is an appealing target for Plant Biotechnology, and both RNA interference (RNAi) and CRISPR/Cas9 were applied, providing wheat lines with the CD-related gliadins strongly downregulated or knock-out, respectively (Barro et al., 2016; SánchezLeón et al., 2018). In particular, RNAi technology was highly effective for the silencing of all three ω -, α/β -, and γ -gliadin fractions in wheat grain. However, the depletion of gliadin proteins affects the bread-making quality depending on the protein fraction being silenced (Gil-Humanes et al., 2012, 2014b). One wheat line denoted as E82 highlights among all RNAi lines produced, with a reduction of 98% of gluten content (Gil-Humanes et al., 2010), and the bread made using flour from this line was tolerated by NCWS patients (Haro et al., 2018), and did not elicit an immunogenic response in a blind randomized crossover gluten challenge with CD patients (Guzmán-López et al., 2021).

As a consequence of the down-regulation of gliadins by RNAi, this E82 line exhibits a marked readjustment in the grain proteins distribution in comparison to that of the wild type (WT); the gliadins were strongly decreased, the HMW glutenin subunits were not affected while the LMW glutenin subunits decrease, and the non-gluten proteins (NGPs), as globulins and triticins, increased, ensuring that the total nitrogen (N) content of the grain remains comparable to that of the WT (GilHumanes et al., 2010; Pistón et al., 2013; Ozuna et al., 2015; Barro et al., 2016). These protein compensation processes are particularly important to understand the molecular mechanisms operating in the grain, as they would allow these pathways to be targeted by using New Plant Breeding Techniques (NPBTs), such as CRISPR/Cas, for the development of new wheat varieties with specific nutritional profiles and unable to elicit an immunogenic response.



The accumulation of wheat prolamins during grain filling is mainly regulated at the transcriptional level by the interaction between the cis-regulatory elements (CREs), present in prolamins promoters, and the transcription factors (TFs) which recognize them (Kawakatsu and Takaiwa, 2010). Several CREs have been identified in the promoter sequences of seed storage protein (SSP) genes like gliadins and glutenins, being the endosperm box (E box), which is comprised of the prolamins box (P box) and the GCN4 motif, one of the most important (Kawakatsu and Takaiwa, 2010). The Storage Protein Activator (TaSPA), a wheat ortholog of basic leucine zipper (bZIP) opaque2 (O2) from maize, recognizes the GCN4 motif, while the DNA binding with one finger (DOF) Prolamins Binding Factor (TaPBF) recognizes the P box (Kawakatsu and Takaiwa, 2010). There are other well-conserved CREs and TFs involved in transcriptional regulation of SSPs genes, like the TaGAMYB TF that binds to the 50 -C/TAACAAA/C-30 -like motif, and the TaFUSCA3 that binds to the RY motif (Kawakatsu and Takaiwa, 2010; Guo et al., 2015; Sun et al., 2017).

Apart from proteins, wheat grain is also an important source of carbohydrates, with starch accounting for ~ 80% of its composition. Starch synthesis is catalyzed by enzymes like starch synthase (SS), and starch branching enzyme (SBE), among others (Shevkani et al., 2017). They are coded by starch synthesis-related genes (SSRGs) and they are regulated by TFs as bZIPs and NACs (Wang et al., 2013; Zhang et al., 2019). In addition, there are TFs actively involved in the regulation of both SSPs and SSRGs, like TaSPA and TaPBF in wheat, or O2 and PBF in maize (Zhang et al., 2016; Guo et al., 2020; Orman-Ligeza et al., 2020).

During grain filling, active source-sink communication is a key point for the accumulation of SSPs and starch, and for therefore, a major determinant for wheat quality and grain yield. This source-sink communication controls the production and transport of photoassimilates from leaves to grains, and the capacity of grain to store them (Aguirrezábal et al., 2015; Smith et al., 2018). Moreover, around 70% of the N in wheat grain comes from senescing leaves, mainly in amino acid form, and it is used for the synthesis of SSPs during the development of the embryo (Peeters and Van Laere, 1994; Masclaux-Daubresse et al., 2008).

For the E82 line, with strongly down-regulated gliadins, a strong protein compensatory mechanism operates to maintain the grain N level. This line is an excellent material to understand the mechanisms of protein regulation in the grain, and how this compensatory phenomenon is driven by or affects other tissues such as leaves that act as the main suppliers of compounds that ultimately determine the nutritional profile of the grain. In the present work, we aimed to progress in deciphering the genetic and molecular mechanisms of protein compensation which operate in the grain of the wheat RNAi lines, to move forward in the development of new wheat varieties, safe for patients with gluten/wheat intolerance and possess exceptional nutritional properties.



2. Materials and Methods

2.1. Plant material and experimental design

In this work, we used the E82 RNAi wheat line, derived from BW208, used as the WT (Gil-Humanes et al., 2010). The E82 line is the result of eight generations of self-pollination. This RNAi line contains the plasmids pGhp- ω/α and pghp8.1, harboring inverted repeat sequences of the most conserved regions from ω , α , and γ -gliadin genes, driven by endosperm-specific promoters (Pistón et al., 2008, 2009).

The two lines were grown in a greenhouse with a day/night regime of 12/12 h at 24/16°C. The seeds and flag leaf from three different plants as biological replicates were collected at 20 days after anthesis (DAA) for RNA extraction for each line and immediately frozen with liquid nitrogen and stored at -80°C . The seeds of three additional plants were also collected at harvest time for prolamin content determination.

2.2. RNA extraction and sequencing

The grains and flag leaves of the WT and E82 at 20 DAA were milled, and the total RNA was extracted using the protocol of Direct-zol RNA Kits (zymoresearch.com). For grain tissue, we also performed the first step described in the protocol of Meng and Feldman (2010).

Total RNA was sent to Novogene (United Kingdom) Company Limited (25 Cambridge Science Park Milton Road, CB4 0FW, United Kingdom) for sequencing. The quality control of total RNA was performed by NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, United States) for preliminary quantitation and by 2% agarose gel electrophoresis for RNA integrity. The RNA integrity and quantification were also carried out by Novogene (Agilent 2100). The rRNA was removed using the Ribo-Zero kit. cDNA synthesis, library construction of 2×150 paired-end reads, and sequencing with Novaseq 6000 system (Illumina) were performed by Novogene.

2.3. RNA-seq data analysis

Bioinformatic analysis was performed on a server with 64 cores and 128 GB of RAM with the operative system GNU/Linux Ubuntu version 18.04 (www.ubuntu.com, United Kingdom).

2.3.1. Trimming

The trimming of raw reads and the adapters removing were performed with Trimmomatic-0.36 (params: ILLUMINACLIP:TruSeq2-PE.fa:2:30:10 AVGQUAL:20 LEADING:3 TRAILING:3 MINLEN:36) (Bolger et al., 2014). The $\sim 2\%$ of reads per sample were discarded (Supplementary Table 4.1).



2.3.2. Mapping

Reads were pseudoaligned by kallisto v0.44.0 (Bray et al., 2016) with default parameters (--pseudobam included) to the International Wheat Genome Sequencing Consortium (IWGSC) RefSeqv1.1 transcriptome annotation (Appels et al., 2018). The control quality of pseudoalignment was performed with samtools v1.3.1 (Li et al., 2009) (Supplementary Table 4.2). The summarized Transcript Per Million (TPM) matrix at gene level was generated with tximport v.1.14.2 (Soneson et al., 2016).

2.3.3. Differential gene expression (DGE) analysis

DGE analysis was performed with edgeR v.3.22.5 (Robinson et al., 2009) in R v.3.6.3 (R Core Development Team, 2020). We established that differentially expressed (DE) genes had a False Discovery Rate (FDR) < 0.05. Two pair-wise comparisons were performed: WT grain tissue vs. E82 grain tissue, and WT leaf tissue vs. E82 leaf tissue. The genes with a sum of Counts Per Million (CPM) in all samples higher than 1 were considered in the DGE analysis. The complete expression data for all the genes and DGE analysis results are deposited in Gene Expression Omnibus (GEO) database with the identification number GSE199525.

2.3.4. Gene Ontology (GO) and pathway enrichment analyzes

The GO and the Kyoto Encyclopedia of Genes and Genomes (KEGG) terms were obtained from RefSeq v1.1 annotation. We only use high confidence genes for the GO enrichment analysis and this was carried out with g:Profiler2 (Raudvere et al., 2019) for up- and down-regulated genes separately in each comparison. For pathway enrichment analysis we used also high-confidence genes.

2.4. Seed-borne allergens and immune-responsive proteins

The annotation of genes coding for allergens and immunogenic proteins implicated in human diseases was obtained from Juhász et al. (2018). Prolamins genes from the work of Altenbach et al. (2020) and Huo et al. (2018a,b) were searched in the wheat genome to include them in the analysis.

2.5. Gene Network Inference with Ensemble of trees (GENIE3) and Weighted Gene Co-expression Network Analysis (WGCNA) networks in bread wheat

Annotation for TF genes was obtained from Ramírez González et al. (2018). The Gene Network Inference with Ensemble of trees (GENIE3) network (Ramírez-González et al., 2018) was used for predicting the downstream target genes of the DE TFs in our DGE analysis for WT vs. E82 comparison in the grain tissue. The top 1 million interactions of the GENIE3 network were used for the analysis, according to previous researches (Huang et al., 2018; Ramírez-González et al., 2018; Harrington et al., 2020). We calculated the shared Ratios of each DE TF gene (shared Ratio: number of DE target genes/number of target genes, for each TF gene) as described in Harrington et al. (2020). The DE TF genes with a shared Ratio > 0 were considered candidate TFs. The relationship between the candidate TF genes obtained from the GENIE3 network in grain tissue was



visualized with igraph 1.2.4.2 (Csardi and Nepusz, 2006) and nodes (TF genes) were grouped into communities based on greedy optimization of modularity (Csardi and Nepusz, 2006; Murata, 2010). For those TF genes without annotation, the gene names of orthologs in *Arabidopsis thaliana* were included along with TFs gene ID.

To study the candidate TFs functions and their relationship with SSPs and SSRGs, we calculated the percentage of DE genes in the grain of E82 in each module of the Weighted Gene Coexpression Network Analysis (WGCNA) grain network, a coexpression network of 850 RNA-seq of wheat in different tissues and conditions (Ramírez-González et al., 2018). We repeated this process for the leaf tissue.

2.6. Gene expression by quantitative real-time PCR (qPCR)

We selected DE genes to validate the results by qPCR. cDNA was synthesized using 1 µg of total RNA with qScript cDNA SuperMix (Quantabio, United States). qPCR was performed with SYBR Green Supermix (Bio-Rad, United States) on CFX Connect Real-time PCR Detection System (Bio-Rad). The BioRad CFX Manager 3.1 was used for data analysis. For target efficiency, serial dilutions 1:2 were prepared for a bulk of grain and leaf samples separately. For data normalization, CDC, RLI and ADP-RF genes were used as reference genes (Giménez et al., 2011) using the primers listed in Supplementary Table 4.3.

2.7. Prolamin quantification by Reverse Phase-High Performance Liquid Chromatography (RP-HPLC)

For prolamin extraction, 100 mg of flour from each line at 10, 20, 30, and 40 DAA were used per each biological replicate. The gliadin and glutenin were sequentially extracted following the protocol of Pistón et al. (2011), and protein quantification was performed by RP-HPLC (1200 Series Quaternary LC System liquid chromatography from Agilent Technologies) with a DAD UV-V detector at 210 nm. A 25 cm long column LiChrospher R 100 RP8 (5 µm) (Merck) was used.

2.8. Total protein, thousand kernel weight, and starch content

For total protein, grains of the WT and the E82 RNAi lines were collected at harvest and milled. Total protein content was calculated with Kjeldahl (%N × 5.7) method [standard ICC no. 105/2] for each biological replicate. The weight of thousand kernels was measured in triplicate. The starch content was determined with the standard ICC method no. 123/1 for each biological replicate. All these measures were expressed by the dry weight (DW) of flour.

2.9. Proteomic data analysis

Grains of the WT and the E82 RNAi lines at harvest were collected and milled for proteomic data analysis. Protein digestion, liquid chromatographic and mass spectrometric analysis were performed as described in Vaquero et al. (2018) followed by custom Python scripts (github.com/MiriamMarinS/ProtEp) for counting the number of unique peptides assigned to each wheat protein fraction.



3. Results

We carried out a transcriptomic (RNA-seq) study for the E82 RNAi line and its WT (BW208) using grain and leaf samples at 20 DAA. Trimmed reads were pseudoaligned with Kallisto to the RefSeq v1.1 transcriptome annotation, obtaining an average of 71.54% of clean reads mapped (Supplementary Table 4.4). Two pair-wise comparisons were performed for the DGE analysis: WT vs. E82 for grain and leaf tissues separately. The WT vs. E82 grain pair-wise comparison provided 109 genes up-regulated in E82 compared to the WT, while 859 genes were down-regulated. In contrast, the WT vs. E82 leaf pair-wise comparison provided 4,230 and 1,593 genes up-and downregulated, respectively. For grain, the most enriched GO terms for the up-regulated genes in E82 were related to the Cellular component domain, some of them specifically in the chloroplast location, and the Biological Process domain, with genes related to response to abiotic stimulus (Figure 4.1A). While GO terms related to nutrient reservoir activity and starch metabolism were enriched in GO enrichment analysis for down-regulated genes in the E82 line (Figure 4.1B).

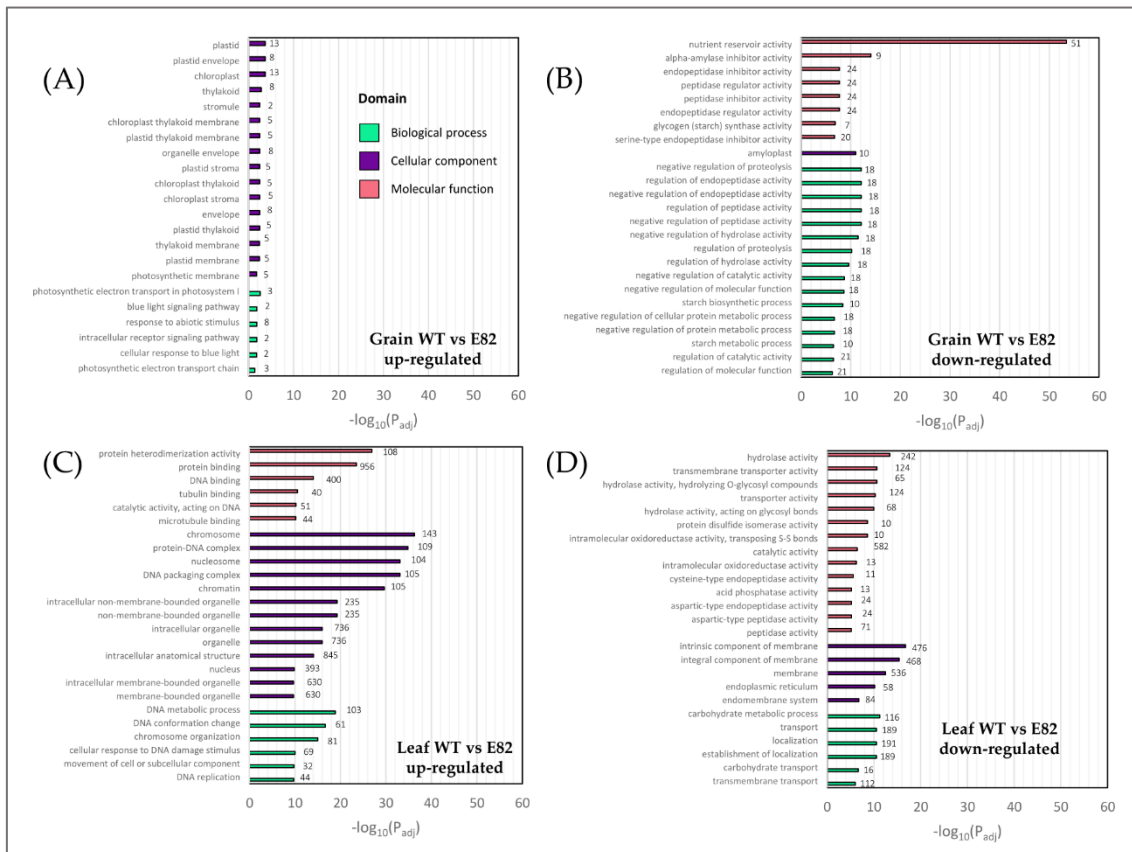


Figure 4.1. Gene Ontology (GO) enrichment analysis of up- and down-regulated genes in the wild type (WT) vs. E82 comparison in grain (A,B) and leaf (C,D) tissues. The top GOs are represented. Values on the bars represent the number of differentially expressed (DE) genes associated with each GO term. *P*-value was adjusted with the Benjamini-Hochberg method.



3.1. Grain proteins are strongly affected by the RNAi

The expression of gliadin genes, targeted by the RNAi silencing fragment, was strongly decreased in the E82 line at 20 DAA, but glutenin genes and other non-targeted grain proteins like avenin-like proteins (ALPs), α -amylase/trypsin inhibitors (ATIs), puroindolines, purinins, and purothionins, had lower expression in E82, which correspond with a high number of DE genes in WT vs. E82 grain comparison (Figure 4.2). Interestingly, globulins and triticins genes were not DE in the E82 RNAi line (Figure 4.2), except for one gene encoding 19 kDa globulin.

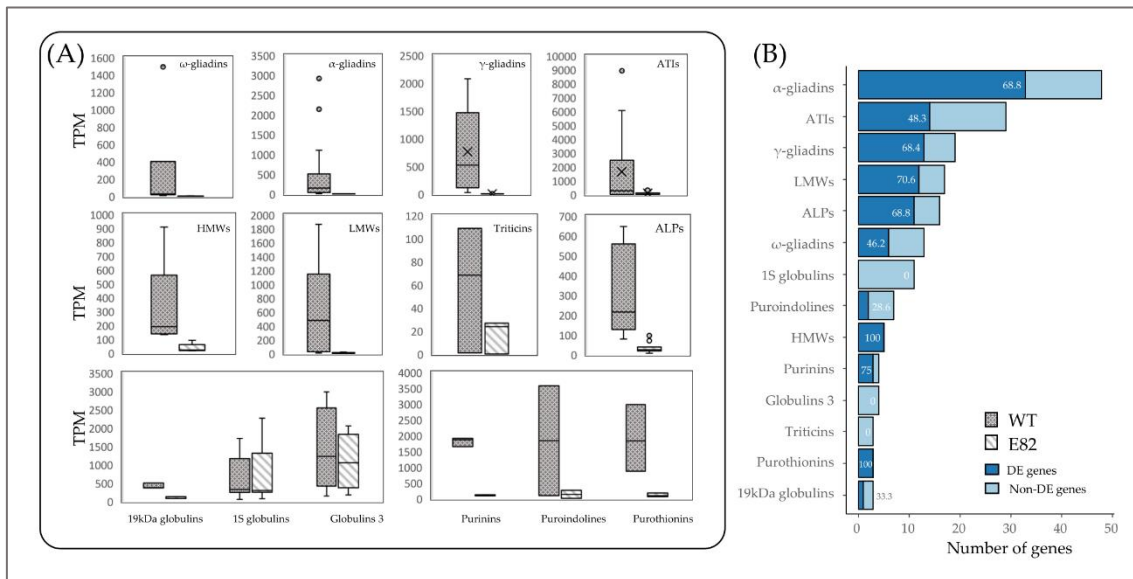


Figure 4.2. (A) Overview of the expression of differentially expressed (DE) genes coding for grain proteins between the wild type (WT) and E82. For globulins and triticins, we represented all genes. The line in the box represents the median value. (B) Number of DE and non-DE genes for each protein type. The number in bars indicates the percentage of DE genes. Gene annotations were extracted from the RefSeq v1.1, Juhász et al. (2018) and Altenbach et al. (2020) and Huo et al. (2018a,b), and for triticins, we performed a blastp of three protein sequences of triticins from Butte 86 summarized in Altenbach et al. (2020) against the wheat genome.

The synthesis and deposition of gliadins and glutenins were also monitored by RP-HPLC throughout grain filling. As shown in Supplementary Figure 4.1, the E82 line had lower content than the WT for the α - and γ -gliadins from the first stages of grain filling, including 20 DAA, while the ω -gliadins content was lower in E82 only at 40 DAA (Supplementary Figure 4.1A). In the case of glutenin proteins, the HMW content was higher in E82 than the WT in the first stages (Supplementary Figure 4.1A) while the LMW content was lower in E82 at 20 DAA (Supplementary Figure 4.1A).

The total protein content for the different gliadin fractions obtained by RP-HPLC (Supplementary Figure 4.1B) and the abundance of unique peptides by LC-MS/MS (Supplementary Table 4.5) at harvesting confirmed that E82 has less content of gliadins and a lower abundance of unique peptides for all the gliadin fractions than the WT



(Supplementary Figure 4.1B and Supplementary Table 4.5). The content of glutenin protein fractions and their abundance of unique peptides varied in the same way, decreasing only for the LMW in the E82 RNAi line (Supplementary Figure 4.1B and Supplementary Table 4.5). Moreover, the NGPs that were increased in E82 at harvesting (Supplementary Figure 4.1C), particularly globulins and triticins, lipid transfer proteins (LTPs), and serpins, also increased the abundance of unique peptides (Supplementary Table 4.5). All these protein rearrangements do not affect the total grain N content, which was not statistically different between E82 and the WT (Supplementary Figure 4.1C).

3.2. Starch synthesis-related genes are also affected in the RNAi line

RNAi silencing in the E82 line also affects the starch content and the genes involved in its regulation. For grain tissue, 14 SSRGs, coding for key enzymes in starch synthesis, were down-regulated compared to the WT at 20 DAA (Supplementary Table 4.6 and Figure 4.3), which agrees with the lower content of starch and the lower kernel weight detected in the E82 line at harvesting (Supplementary Figure 4.1D,E), and with the lower normalized gene expression of SSIIIa/SSIII-2 genes in the grain of E82 obtained by qPCR (Supplementary Figure 4.2). Moreover, for the flag leaf, the expression of five and four SSRGs were up- and down-regulated, respectively, in the E82 line (Supplementary Table 4.6 and Figure 4.3).

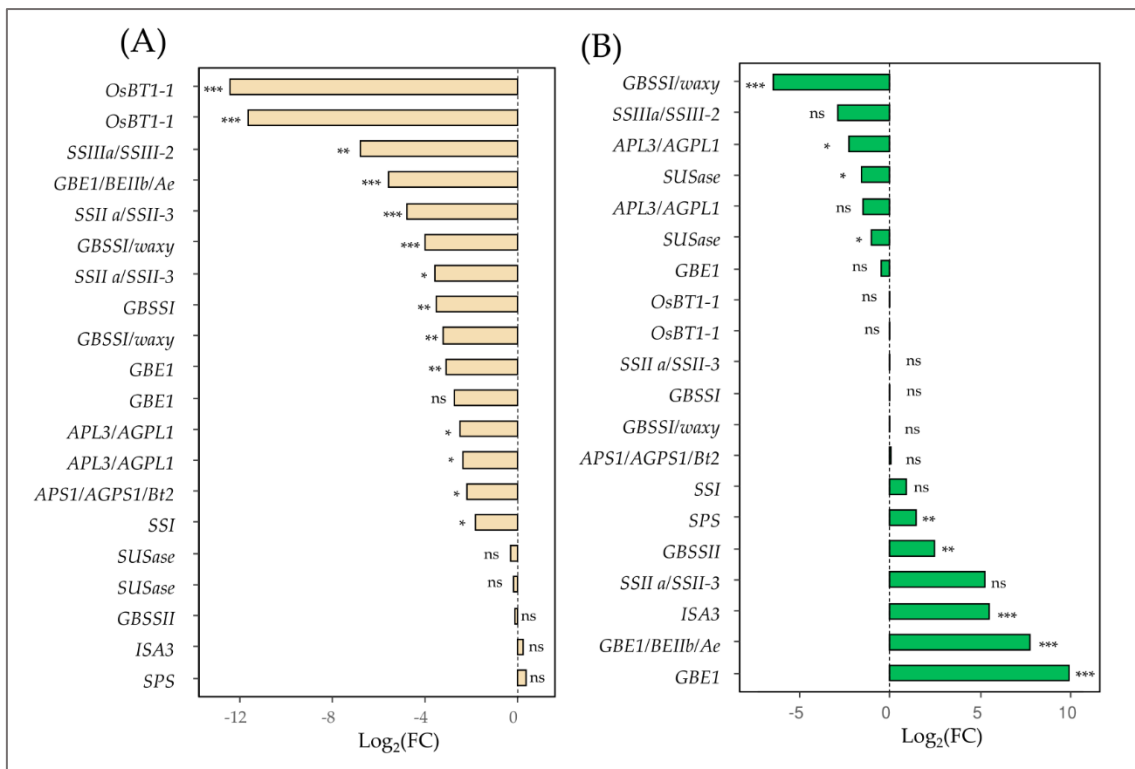


Figure 4.3. Starch synthesis-related genes (SSRGs) differentially expressed (DE) in (A) grain and/or (B) leaf tissue for the pair-wise comparison wild type (WT) vs. E82. *, FDR < 0.05; **, FDR < 0.01; ***, FDR < 0.001; ns, non-significant; FDR, False Discovery Rate.



3.3. Candidate TFs for the grain protein and starch synthesis regulation in the E82 RNAi line

Many of the genes proposed in previous research for the regulation of SSPs and starch synthesis are listed in Supplementary Table 4.7, including the *TabZIPs* which play a role in the amylose biosynthesis in wheat (Kumar et al., 2018). Among these genes, transcriptomic data showed that the A-subgenome homeolog of the *TaPBF* was significantly downregulated in the E82 line, while the other two homeologs present a low, not significant expression (Supplementary Table 4.7). Moreover, in the qPCR analysis, the pooled three homeologs for *TaPBF* were down-regulated in the E82 line (Supplementary Figure 4.2). However, the rest of the TFs listed in Supplementary Table 4.7 were not DE in the grain of E82, and therefore they may not be relevant for explaining the rearrangements in the regulation of the SSP and starch synthesis genes in the E82 line. The expression of homeologs of the *TaSPA*, which has a key role in the biosynthesis of prolamins and starch (Guo et al., 2020), was not over the threshold to be included in the DGE analysis.

Because of the above results, the GENIE3 network was used to find other TFs related to the SSPs and starch synthesis. This network provides biologically-relevant information based on TF's targets prediction, and it was demonstrated for bread wheat TFs related to senescence (Harrington et al., 2020). We searched the DE TFs in the grain of E82 in the wheat GENIE3 network, selecting those which had shared Ratio values higher than zero (at least one of their target genes was DE in the grain of E82). Only 30 out of 3,000 TFs present in the GENIE3 network met both criteria (Figure 4.4), being clustered in five communities (Figures 4.4B,C). The TF *TaPBF-A* belongs to community 3, which together with community 5, has the highest shared Ratio values (Figure 4.4C). Therefore, these TFs form a complex regulatory cascade and will be considered candidate TFs throughout this work.

To establish the putative function of each TF, we summarized (Supplementary Table 4.8) the top three GO terms most enriched for each TF from previously GO enrichment analysis for the GENIE3 network. Many of the TFs were in communities 3, 4, and 5, and associated with carbohydrate-mediated signaling, detection of nutrients, glycogen metabolism, lipid storage, regulation of endopeptidase activity, regulation of catalytic activity, energy reserve metabolic process, and seed oil body biogenesis. TFs in communities 2 and 1 were related to photosynthesis and defense response processes, respectively (Supplementary Table 4.8).

Each candidate TF was then assigned to a module according to the grain co-expression network constructed with WGCNA (Figure 4.4C), and the percentage of DE genes in each module for the E82 grain was represented (Figure 4.5A). The *TaPBF-A* was addressed to module 13 (Figure 4.4C), together with a high number of other down-regulated genes in E82 (Figure 4.5A). Next, we extracted all the genes from this module



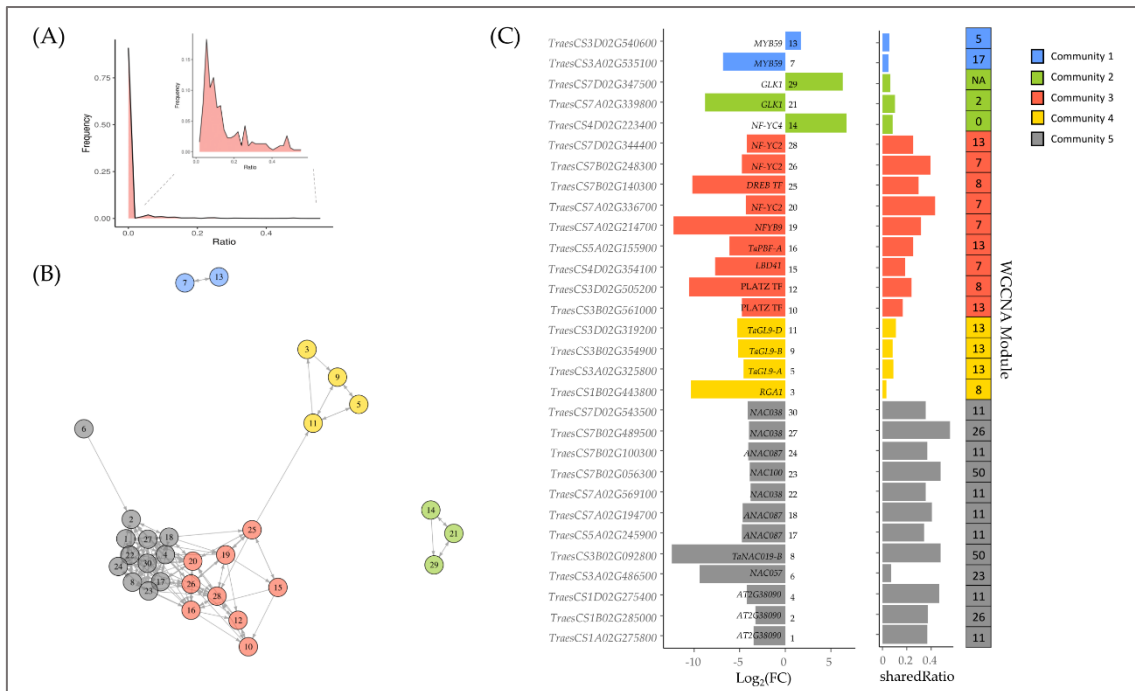


Figure 4.4. (A) Frequency of shared genes between Gene Network Inference with Ensemble of trees (GENIE3) transcription factors (TFs) target genes and differentially expressed (DE) genes in the grain of E82 compared to the wild type (WT). For improving the representation of frequencies, we remove those TFs with zero shared genes in the insert. (B) Network of candidate TFs implicated in the protein and starch changes in the grain of E82. All these genes are DE TF with a shared Ratio > 0. The different colors represent the community of nodes detected based on greedy optimization of modularity. (C) Fold-change (FC) values and shared Ratio of each gene in the network. Gene name (A. thaliana ortholog information) and network ID are represented in the plot. On the right, there is the module of the Weighted Gene Co-expression Network Analysis (WGCNA) network for grain tissue to which each TF belongs.

and performed GO enrichment analysis; the GO terms related to carbohydrate metabolism were enriched (Figures 4.5B,C), and GOs related to amino acid metabolism stand out (Supplementary Table 4.9). In addition, module 11 was also overrepresented among the candidate TFs and has a high percentage of DE genes in the grain of E82 (Figures 4.4C, 5A). As for module 13, we performed GO enrichment analysis of genes included in module 11, and we found nutrient reservoir activity and α -amylase inhibitor activity terms enriched (Figures 4.5B,C).

The next step was to establish the match between the modules where the candidate TF genes were and the modules in which the grain protein genes and SSRGs were included. Other grain proteins such as LTPs and serpins, in the list of allergens and immune-responsive proteins implicated in human diseases listed in Juhász et al. (2018),



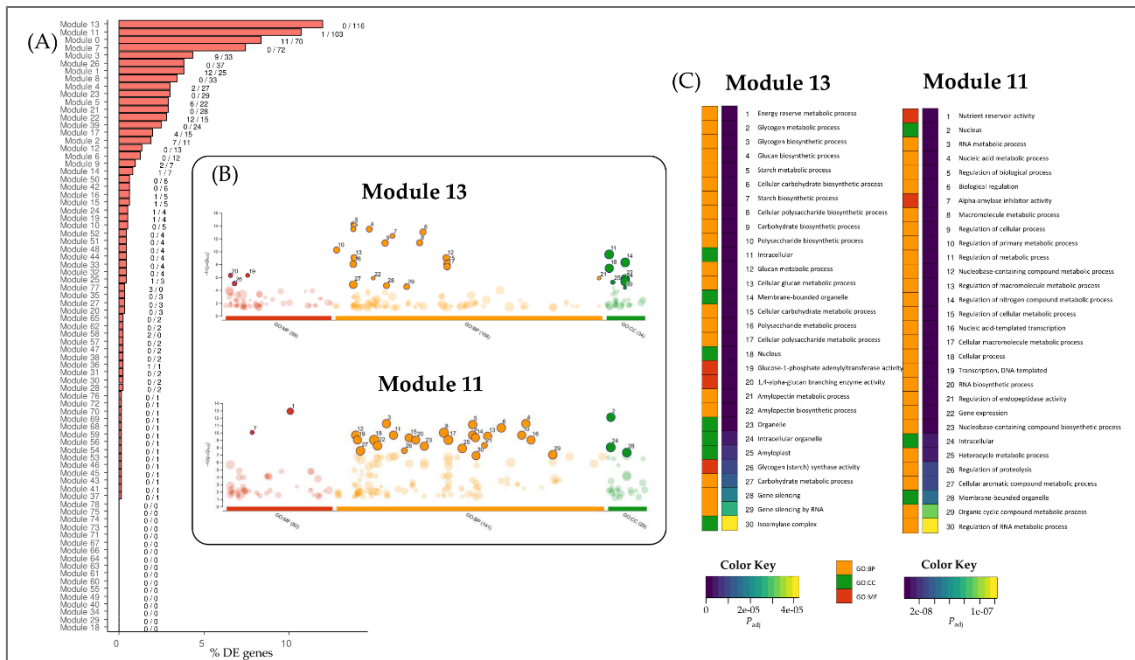


Figure 4.5. (A) Percentage of differentially expressed (DE) genes that belong to each module in the Weighted Gene Co-expression Network Analysis (WGCNA) network for grain tissue. The label close to the bars indicates the number of up/down-regulated genes in each module. (B) Manhattan plot of the Gene Ontology (GO) enrichment analysis of genes in modules 13 and 11 of the WGCNA network from grain tissue. The number of enriched GOs per domain is indicated between parentheses. The enriched GOs are represented, and the top 30 GO enriched are highlighted. (C) The adjusted P-value with the Benjamini-Hochberg method of the top 30 enriched GOs is represented per domain in modules 13 and 11. GO:MF, Molecular Function domain; GO:BP, Biological Process domain; GO:CC, Cellular Component domain.

were added to the study. We also searched these genes in the WGCNA grain co-expression network (Figure 4.6). Briefly, many gliadin and glutenin DE genes in the grain of E82 shared the same modules as the candidate TFs (Figure 4.6). The same occurs with other grain protein genes such as ATIs, ALPs, purothionins, puroindolines, purinins, serpins, and LTPs (Figure 4.6). Some grain protein genes such as globulins, including triticins genes, were not DE in our RNA-seq data analysis, although some of them showed a slight tendency to decrease in E82, being one 19 kDa globulin gene down-regulated (Figures 4.2A,B). While the genes for globulin 1S and globulin 3 were included in module 4 (not containing any of the candidate TFs), the 19 kDa globulin and triticin genes were included in module 26, and modules 11 and 50 respectively, with candidate TFs of community 5 (Figure 4.6). On the other hand, all the SSRGs down-regulated in E82 grain were also included in the same modules as the candidate TFs, except for SSIIa/SSII-3 genes (Figure 4.6).



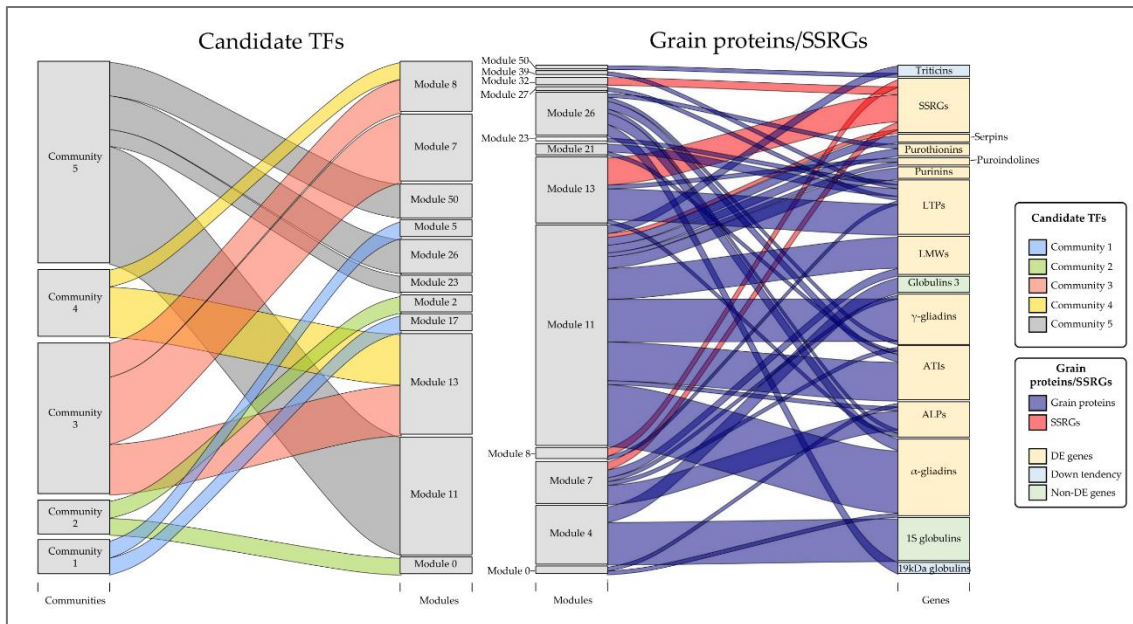


Figure 4.6. Relationships between communities of candidate transcription factors (TFs) and modules of Weighted Gene Co-expression Network Analysis (WGCNA) grain co-expression network (left), and between these modules and grain proteins/starch synthesis-related genes (SSRGs) (right). The size of the gray boxes indicates the number of genes contained in each category. All the grain protein genes and SSRGs used for representation were differentially expressed (DE) in our RNA-seq analysis. In the case of grain protein genes, the globulin ones non-DE were included for discussion purposes. Down tendency: genes with non-significant lower expression in E82, except for one DE gene encoding 19 kDa globulin.

3.4. The source-sink communication during grain filling is greatly affected by RNAi silencing

We also carried out an RNA-Seq analysis in leaves to determine if the transcriptomic changes found in the grain of the E82 line during grain filling could modulate the transcriptome of the leaf. As shown in Figure 4.7, there were many transcriptomic changes in the leaf of E82 against the WT. In fact, there were more DE genes in the E82 leaf than in the grain, and a higher number were up- than down-regulated. Many of the up-regulated genes were related to the DNA metabolism and chromosome organization (Figure 4.1C), while the downregulated ones were related to peptidase activity, carbohydrate metabolism, and transport (Figure 4.1D, Supplementary Table 4.6 and Figure 4.3B). Moreover, starch and sucrose metabolism and amino acids metabolism pathways were enriched in the downregulated genes of E82 (Figure 4.7B). All these processes are linked to source-sink interactions during grain-filling (Yu et al., 2015). As for grain tissue, the WGCNA co-expression modules ordered by their percentage of DE genes were represented (Figure 4.7C). The highest percentages of DE genes in the leaf were in modules 0, 1, 2, 3, 6, and 7 (Figure 4.7C). Then, a GO enrichment analysis was performed for the up-and downregulated genes in the modules with a higher percentage of DE genes: protein heterodimerization, hydrolase activity, and glutamate and



glutamine biosynthesis processes, among others, were enriched; GOs as transmembrane transport and carbohydrate metabolism were only enriched in the downregulated genes (Figure 4.7C).

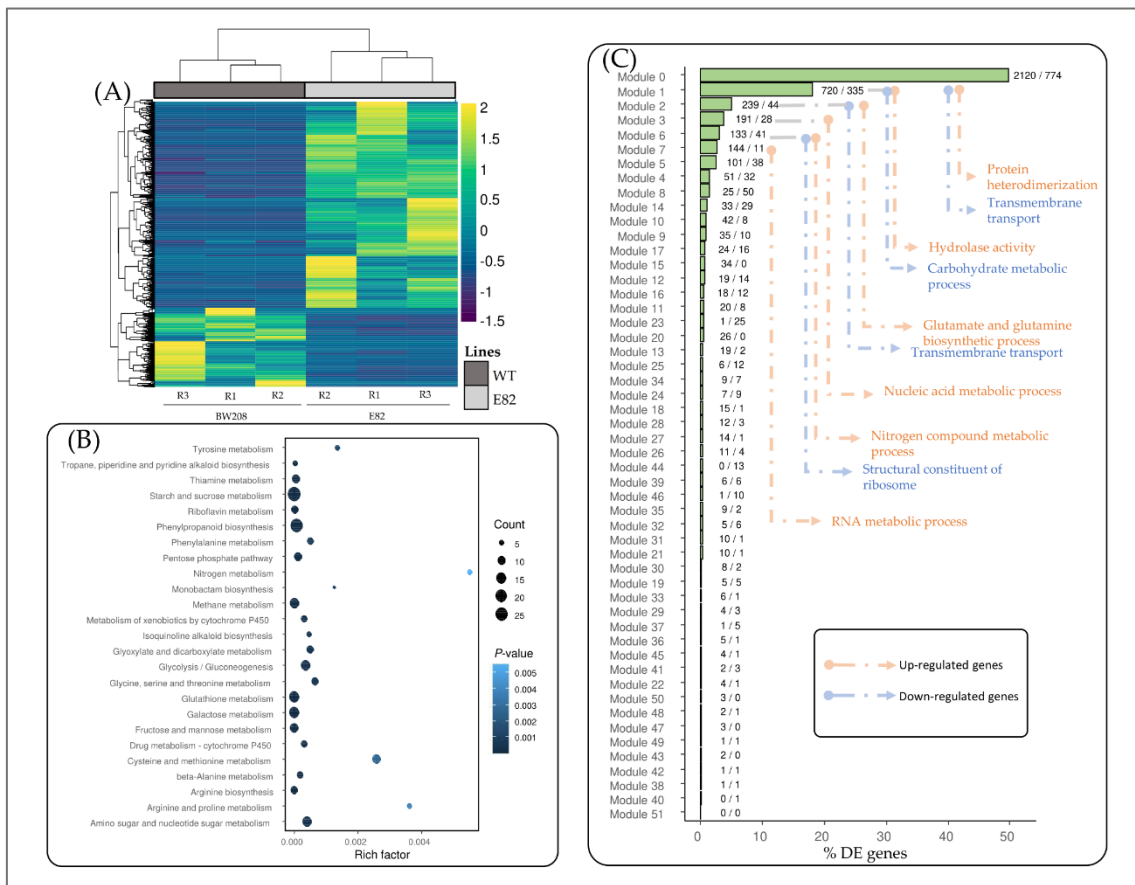


Figure 4.7. (A) Heatmaps of Transcript Per Million (TPM) for differentially expressed (DE) genes for the wild type (WT) vs. E82 comparison in leaf tissue. Values are scaled by row (–1.5 to 2). Dendrograms of genes and lines were represented. (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis for down-regulated genes in E82 leaf tissue against the WT. The rich factor is the division between the number of DE genes annotated and the total number of genes annotated in each pathway. The count value is the number of DE genes annotated in this Kegg Orthology (KO). (C) Percentage of DE genes which belong to each module in the Weighted Gene Co-expression Network Analysis (WGCNA) network for leaf tissue. The label close to the bars indicates the number of up-/down-regulated genes in each module. The most enriched GOs for the set of up- and down-regulated genes in each module were represented following the arrows. R: replicate.

Then, we decided to study more in detail the genes related to peptidase activity and nutrients transport in the leaf of E82 because of their implication in the source-sink communication and, in the end, in the SSP and starch content in grain.

3.4.1. Proteases and protease inhibitors in the leaf of E82

The Papain-like cysteine proteases (CysProt), belonging to C1A according to the MEROPS database (Rawlings et al., 2016), play a main role in the proteolytic activity and



nutrient remobilization during leaf senescence. Then, genes annotated in RefSeq v1.1 with the PFAM “Peptidase_C1” (PF00112) were searched on WheatMine website (<https://urgi.versailles.inra.fr/WheatMine>). Among them, one ortholog of *HvPap1* of *Hordeum vulgare* (NCBI Acc. N.: BN000093), was downregulated in E82 (Supplementary Table 4.10). In addition, four and five (*HvPap4* orthologs, RD19D, *Psy1-D1* ortholog, and *HvPap14* ortholog) CysProt genes were down- and upregulated in E82, respectively (Supplementary Table 4.10). In the case of cystatins, specific modulators of C1A peptidase activities, we looked for the cystatin domain (IPR000010) in WheatMine. There were five DE cystatin genes in E82. One of them was up-regulated in E82 (ortholog of *Icy2* of *H. vulgare*) (Supplementary Table 4.10). The serine proteases, together with CysProt, are also the most abundant enzymes associated with leaf senescence (Roberts et al., 2012). In the case of these peptidases, we filter genes with serine-type peptidase and endopeptidase activity in RefSeq v1.1. Briefly, 20 of these genes were up-regulated and 15 were down-regulated (Supplementary Table 4.11).

3.4.2. Sugar and amino acid transporters in the leaf and grain of E82

Both, carbohydrate and transmembrane transport GOs were enriched among the down-regulated genes in the leaf of E82 (Figure 4.1D). The carbohydrate transport genes were identified accordingly to the work of Gautam et al. (2019), and the DE ones are listed in Supplementary Table 4.12. For carbohydrate transport, many of the *TaSWEET* and *SUT* genes were downregulated in the leaf of E82 (Figure 4.8 and Supplementary Table 4.12). In addition, the expression of *TaSWEET13* homeolog genes (*TraesCS6B02G421800* and *TraesCS6D02G367400*) was also studied by qPCR, confirming the down-regulation in the leaf of E82 against the WT (Supplementary Figure 4.2). While in the grain, only three *TaSWEET* genes were down-regulated in E82 (Figure 4.8).

Concerning amino acid transporters, they have been identified accordingly to the work of Tian et al. (2020), and the DE genes were listed in Supplementary Table 4.12. Many of the amino acids permease (AAP) genes, which transport a wide range of amino acids (Kohl et al., 2012), were downregulated in the leaf of E82, except for three of them which were up-regulated: *TaAAP22-7A*, *TaAAP6-2A*, and *TaAAP8-3A* (Figure 4.8 and Supplementary Table 4.12). Four genes encoding proline transporters (ProTs) 2 and 3, proline-specific transport, were down-regulated in the E82 leaf and one encoding ProT1 was slightly up-regulated, while two of them were upregulated in the E82 grain (Figure 4.8 and Supplementary Table 4.12). Interestingly, three genes encoding lysine-histidine-like transporters (LHTs), associated with specific lysine-histidine transport, were highly up-regulated in the E82 leaf (Figure 4.8 and Supplementary Table 4.12).



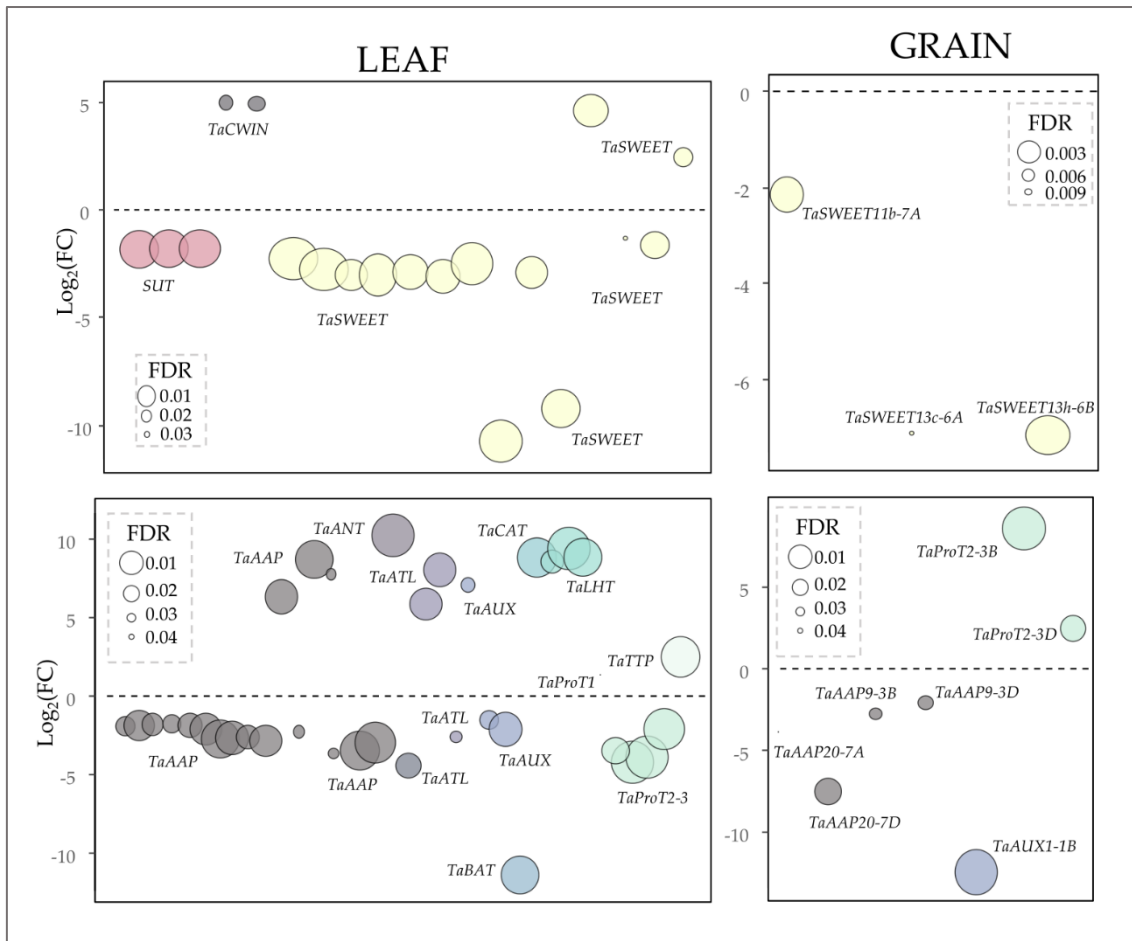


Figure 4.8. Sugar transporters (up) and amino acid transporters family (ATF) (down) genes differentially expressed (DE) in the leaf and grain of E82. The Log₂(FC) is represented, those genes with Log₂(FC) > 0 are up-regulated in E82, and those with Log₂(FC) are down-regulated. The size of the dots corresponds to the FDR of each gene. FC, fold-change; FDR, False Discovery Rate.

4. Discussion

The down-regulation of wheat gliadins by RNAi provided lines with a reduction of up to 98% of gluten in the grain (Gil-Humanes et al., 2010). The most outstanding line, named E82, showed the lowest amount of CD immunogenic epitopes (Sánchez-León et al., 2019), and elicited no immunogenic response in celiac disease patients after the oral consumption of E82 bread (Guzmán-López et al., 2021). Moreover, this line presents sensory properties comparable to that of the WT flour, suggesting that gliadins could not play a relevant role in texture, aroma, flavor and appearance as much as other grain components (Cho and Peterson, 2010; Gil-Humanes et al., 2014a). A common feature of this and other RNAi lines with the SSPs silenced was the rearrangement of the grain proteins to provide total N contents comparable to the WT (Gil-Humanes et al., 2012; Altenbach et al., 2019; García-Molina et al., 2019), which indicates that a compensatory mechanism is operating. However, depending on the SSP being targeted, there are different profiles of protein compensation. For example, line D793, with all the gliadin fractions down-regulated, shows a higher content of HMW glutenin subunits and lower



LMW content than the WT (Gil-Humanes et al., 2011; Marín-Sanz et al., 2022). In contrast, when only the γ -gliadins are targeted, RNAi lines present higher content of both ω - and α -gliadins (Pistón et al., 2011). In RNAi lines for silencing the ω -5 gliadins, an increase in NGPs, HMW and LMW glutenin subunits, and α -gliadins were reported (Altenbach et al., 2014). However, if ω -1,2 gliadins were silenced, all gliadins and LMW glutenin subunits were also suppressed, and major compensation with HMW glutenin subunits and NGPs occurred (Altenbach et al., 2019). The protein compensation was also reported in maize RNAi lines, in which the reduction of the 22-kDa α -zeins was compensated by the 19-kDa α -zeins, and vice versa (Huang et al., 2004), and in rice mutants, in which the low glutelin content was compensated by other SSPs such as prolamins (Iida et al., 1993). In the case of the E82 line, with all three ω -, α - and γ -gliadin fractions down-regulated, an increase in the NGPs and a reduction in LMW were reported, leading to the decrease in the content of total glutenins (Pistón et al., 2013; Barro et al., 2016). The reduction of LMW could be due to an off-target silencing because of the high homology between the LMW mRNA sequence and the small interfering RNA (siRNA) formed from the hairpin RNA (hpRNA) (Gil-Humanes et al., 2014b). Altogether, it is clear that different compensatory mechanisms are at work in response to the depletion of the different protein fractions by RNAi. Moreover, CRISPR/Cas targeted mutations seem to provide a similar response in the wheat grain (Sánchez-León et al., 2018). Understanding the molecular mechanisms underlying this process of protein compensation in the grain will allow us to make fine adjustments, using CRISPR/Cas techniques, toward developing wheat lines with improved nutritional properties and suitable for people who suffer from any wheat-related pathology.

In this work, the transcriptome of the line E82 in both grain and leaf during grain filling provides new insights into both protein rearrangements and grain-to-leaf communication pathways when all gliadins are strongly silenced. A very high number of genes were DE in the leaf of E82, comprising processes such as nutrient availability and nutrient transport, while in the grain, many genes related to the protein synthesis, particularly genes encoding gluten proteins, along with ALPs, ATIs and LTPs, and genes related to the starch synthesis, were down-regulated, in concordance with the grain protein composition and the slight reduction of the starch content found in the grain of E82. Among the well-known TFs involved in prolamins and SSRGs regulation, the DOF-type *TaPBF-A* which binds to the P-box motif of prolamins genes (Mena et al., 1998; Zhang et al., 2016; Guo et al., 2020), was down-regulated in E82. Despite the relation between *TaPBF* and the SSRGs has to be further studied, some genes implicated in the starch synthesis also present in their promoters the *TaPBF* binding motifs (P box and Pyrimidine box), and they can be targets of this TF, as predicted by the GENIE3 network (Orman-Ligeza et al., 2020), and in line with our transcriptomic results. Surprisingly, the rest of the well-known TFs related to the regulation of the synthesis of prolamins and starch were not DE in the grain of E82, which prompted us to expand the study to all the TFs to elucidate their role in the grain composition of the E82 line.



Among all the TFs, we selected those that were DE in E82, obtaining a putative TF regulatory cascade for the proteins and starch synthesis in the grain. This network comprised five communities, being three of them closely linked to each other, and the *TaPBF-A* gene included in one of them (community 3). The relationship between *TaPBF-A* and the regulation of prolamins, and some of the starch synthesis enzymes, suggests that not only this TF could be responsible for the shifts in the protein fractions observed in the grain of E82, but also those TFs closely related in the network would be potential candidates to regulate the grain protein genes and the SSRGs expression. Specifically, *TaNAC019* binds to the promoters of *TaGlu-1* loci on the homeologs chromosomes 1A, 1B, and 1D, which encode the HMW glutenin subunits. The *TaNAC019-B* homeolog was also present in our candidate TFs network in the community containing NACs, and preceding the *TaPBF-A* one. Recently, the role of this gene in the synthesis of SSPs and starch has been determined by using mutants for *TaNAC019* in wheat (Gao et al., 2021). The searching of the candidate TFs, protein genes and SSRGs in the co-expression modules of the WGCNA network, provided that many of the DE prolamins genes were classified in the same modules as the candidate TFs in the two largest communities in our network, which include *TaPBF-A* and *TaNAC019-B*. In addition, the DE genes that encode for ATIs, ALPs, purinins, purothionins, puroindolines, serpins, and LTPs were also classified in the same co-expression modules as the candidate TFs, suggesting that all these grain proteins may be subjected to the same regulatory network. However, the genes annotated as ω -gliadins and HMW were not considered as they are not present in the published WGCNA network for wheat, probably because the network was proposed before these genes were included in the wheat genome annotation. Except for one gene encoding 19 kDa globulin, globulin genes did not change their expression significantly in the grain of E82, but participated in the protein compensation of this and other RNAi lines (Gil-Humanes et al., 2011; Barro et al., 2016), suggesting that globulins would be regulated separately to prolamins and other grain proteins. In fact, globulins 3 and 1S globulins were not included in the same modules as the candidate TFs. Interestingly, an increase in the lysine-rich proteins (globulins), in addition to a high content of free lysine, was obtained in a triple null wheat line for the *TaPBF* gene homeologs (Moehs et al., 2019), suggesting that no relationship between the *TaPBF* and the increasing of globulins content exists. Moreover, the above triple null mutant presented lower content of gliadins and LMW, which agrees with the E82 profile. In addition, this mutant also showed a lower content of free proline and glutamine in the grain compared to its WT, as expected because of the role of this TF in the prolamins regulation (Moehs et al., 2019). Similarly, the *lys3a*, a previous mutant of barley with a missense allele in a domain of the PBF, showed a reduction in hordein content, while free and protein-bound lysine increased (Shewry et al., 1980; Moehs et al., 2019). Overall, it seems that most of the globulin genes are not the targets of the *TaPBF*, and their increase in the E82 line could be a consequence of a different regulation pathway and/or they are favored by the free amino acids profile lead after the down-regulation of gliadins.

Recently, a list of candidate TFs to regulate the SSPs was proposed for *Triticum urartu* by using a gene co-expression analysis (Luo et al., 2021). Our results agree with



those, as eight orthologs TFs of bread wheat were in community 5 (NAC community) and five in community 3 (*TaPBF-A* community), so they are potential candidates for the regulation of the SSPs in bread wheat. Interestingly, in our NAC community, one of the TF (*TraesCS3A02G486500*) which encodes the NAC057 ortholog of *A. thaliana* and has been identified as NAM-like protein (Guerin et al., 2019), was not listed as a putative regulator for SSPs by Luo et al. (2021). This gene is not the target of any DE TFs in the grain of E82 and its function in wheat is still unknown. However, considering its position in the candidate TFs regulatory network, it could have a prominent role in the regulation of the SSPs. In addition to the communities containing *TaPBF-A* and *TaNAC019-B*, there is a third community (community 4) connected to them by *TraesCS7B02G140300*, which encodes a TF member of the DREB subfamily A-5 of ERF/AP2 family (Ramírez-González et al., 2018). In this community, the sequence of three of the four TFs found (*TraesCS3A02G325800*, *TraesCS3B02G354900* and *TraesCS3D02G319200*) presents a high similarity with the TF TaGL9 (GLABRA2-like clone 9) from *Triticum aestivum*. This TaGL9 contains a putative lipid-binding domain and it could be involved in the regulation of lipid biosynthesis and their transport (Kovalchuk et al., 2012). In fact, the *TaGL9* genes have the same expression pattern as genes encoding LTPs, suggesting that either the LTP genes are regulated by TaGL9 or both genes are regulated by other TFs (Boutrot et al., 2007; Kovalchuk et al., 2012; Chew et al., 2013). Our results agree with both, as some of the LTP genes were down-regulated in the grain of E82, as well as all three *TaGL9* homeologs, being most of them located in the same coexpression module (module 13). The connection of *TaGL9* with *TaPBF-A* and *TaNAC019-B* described above suggests a common regulation of SSPs, starch, and LTPs.

We also focused on the leaf transcriptome of E82 during the grain filling as the source-sink communication is important for plant productivity, enhanced by source activity and sink strength. Source activity comprises the photosynthesis or nutrient remobilization rates, and sink strength is the capacity to import photoassimilates or biosynthetic rate of starch and proteins among others (Yu et al., 2015). However, not much information on the regulation of such source-sink communication during grain filling is available (Yu et al., 2015). Concerning the leaf transcriptome of E82, we have obtained many DE genes compared to the WT, many of them related to amino acids transport, carbohydrate transport, amino acids and sugar metabolisms, and peptidase activity. In this respect, carbohydrates are the main photoassimilate produced in the leaves. The leaf of E82 is also affected in the production of carbohydrates as many genes related to the carbohydrate metabolism process were down-regulated, and this may lead to lower availability of this nutrient to be transported to filling grain. This may explain that carbohydrate transporters were mainly down-regulated in the leaf of E82, contributing to the lower starch content found in the grain of E82 at harvest in comparison to that of the WT.

Protein degradation allows the recycling of N and other nutrients during leaf senescence (Bhalerao et al., 2003; Guo et al., 2004). In the line E82, we found many DE genes among the C1A CysProt and the serine-type peptidases, two of the most abundant



protease families in plants and important during leaf senescence for the remobilization of nutrients from source tissues to the storage ones (Roberts et al., 2012; Rawlings and Salvesen, 2013; Díaz-Mendoza et al., 2014; Liu et al., 2018). For example, the orthologs in wheat of the HvPap14, HvPap4, and RD19D genes in the leaf of E82 were up-regulated, suggesting that those proteases may actively provide substrates for nutrient remobilization to the grain of E82. Regarding amino acids, the synthesis of some of them was enriched among the down-regulated genes in the RNAi line, which may indicate a lower availability. As we hypothesized in previous works, the protein compensation studied in RNAi lines is a selective process that tries to prioritize the compensation of proteins with similar amino acid profiles to the silenced proteins over the least similar ones, as the protein compensation is closely linked to the availability of amino acids (Pistón et al., 2013). As mentioned before, we observed that the γ -gliadins reduced lines presented an increase of the rest of the prolamin proteins with similar amino acid profiles (Pistón et al., 2011). However, the lines with a reduction of all the gliadins and the LMW presented an increase in HMW and NGPs, including albumins and globulins, which present different amino acid compositions (Pistón et al., 2013). Among the wheat protein fractions, globulins present the highest lysine content and the lowest proline content, while prolamins are rich in proline and glutamine, and have the lowest content of lysine and other essential amino acids such as histidine (Shewry et al., 2009). Accordingly, genes involved in the metabolism of proline were down-regulated in the leaf of E82, in line with the fact that the grain would require less content of this amino acid. Moreover, regarding ProTs, which have a high affinity for proline and play a key role in the rapid distribution of this amino acid (Ueda et al., 2008), *TaProTs* 2-3 were down-regulated in the leaf of E82, which could be a consequence of the downregulation of proline metabolism in the leaf. Regarding lysine-rich proteins, three genes encoding LHTs were strongly upregulated in the leaf of the RNAi line, which could contribute to the high content of total lysine and lysine-rich proteins in the grain of E82 as reported previously (Gil-Humanes et al., 2014a; Barro et al., 2016). In addition to ProTs and LHTs, a large number of genes encoding AAPs were DE in the leaf of E82, being the expression of AAPs regulated in the sink-source communication (Kohl et al., 2012). The above results suggest that the amino acids transporters could play an important role in the E82 grain protein compensation, however, further analyzes are needed to study their implication in this process.

Taking together, the results reported here show that gliadin silencing in wheat has a great transcriptomic effect on the leaves than on the grain. However, the changes observed in the leaf transcriptome are probably a consequence of the changes observed in the grain during grain filling. Therefore, no changes in the leaves are expected before or after this physiological process. These results provide new insights to decipher the complex regulation network which operates during grain filling in response to the down-regulation of prolamins by RNAi: as a summary, the network of TFs in the grain could regulate the starch and SSPs synthesis, but some globulins would be exempt from this regulation, having the principal role in protein compensation (Figure 4.9). This transcriptomic information also highlights the grain-leaf communication (Figure 4.9)



and could be used to redesign the wheat grain protein composition and control the nutrient supply chain during grain filling. All these objectives could be now achievable using CRISPR/Cas technology, enabling the development of new healthier wheat varieties, totally safe for those suffering adverse reactions to wheat.

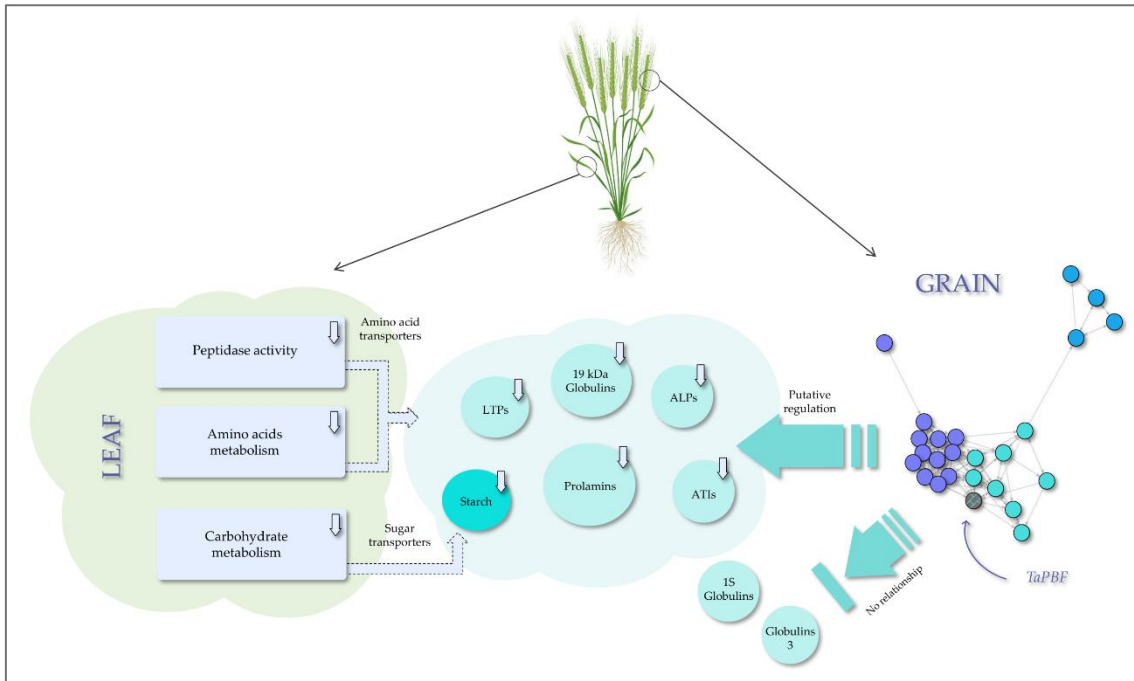


Figure 4.9. Model of the transcription factors (TFs) network role in the regulation of the seed storage proteins (SSPs) and starch synthesis, and the effect on the grain-leaf communication-related processes in a RNAi line with gliadins down-regulated.



5. References

- Aguirrezábal, L., Martre, P., Pereyra-Irujo, G., Echarte, M. M., and Izquierdo, N. (2015). Improving grain quality: ecophysiological and modeling tools to develop management and breeding strategies. *Crop Physiol.*, 423–465.
- Altenbach, S. B., Chang, H. C., Simon-Buss, A., Mohr, T., Huo, N., and Gu, Y. Q. (2020). Exploiting the reference genome sequence of hexaploid wheat: a proteomic study of flour proteins from the cultivar Chinese Spring. *Funct. Integr. Genomics* 20, 1–16. doi:10.1007/s10142-019-00694-z.
- Altenbach, S. B., Chang, H. C., Yu, X. B., Seabourn, B. W., Green, P. H., and Alaedini, A. (2019). Elimination of omega-1,2 gliadins from bread wheat (*Triticum aestivum*) flour: Effects on immunogenic potential and end-use quality. *Front. Plant Sci.* 10, 580. doi:10.3389/fpls.2019.00580.
- Altenbach, S. B., Tanaka, C. K., and Allen, P. V. (2014). Quantitative proteomic analysis of wheat grain proteins reveals differential effects of silencing of omega-5 gliadin genes in transgenic lines. *J. Cereal Sci.* 59, 118–125. doi:10.1016/j.jcs.2013.11.008.
- Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* (80-.). 361, eaar7191.
- Aziz, I., Dwivedi, K., and Sanders, D. S. (2016). From coeliac disease to noncoeliac gluten sensitivity; Should everyone be gluten free? *Curr. Opin. Gastroenterol.* 32, 120–127. doi:10.1097/MOG.0000000000000248.
- Barro, F., Iehisa, J. C. M., Giménez, M. J., García-Molina, M. D., Ozuna, C. V., Comino, I., et al. (2016). Targeting of prolamins by RNAi in bread wheat: Effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotechnol. J.* 14, 986–996. doi:10.1111/pbi.12455.
- Bhalerao, R., Keskitalo, J., Sterky, F., Erlandsson, R., Björkbacka, H., Birve, S. J., et al. (2003). Gene expression in autumn leaves. *Plant Physiol.* 131, 430–442. doi:10.1104/pp.012732.
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.
- Boutrot, F., Meynard, D., Guiderdoni, E., Joudrier, P., and Gautier, M. F. (2007). The *Triticum aestivum* non-specific lipid transfer protein (TaLtp) gene family: Comparative promoter activity of six TaLtp genes in transgenic rice. *Planta* 225, 843–862. doi:10.1007/s00425-006-0397-7.
- Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34, 525–527.
- Chew, W., Hrmova, M., and Lopato, S. (2013). Role of homeodomain leucine zipper (HD-Zip) iv transcription factors in plant development and plant protection from deleterious environmental factors. *Int. J. Mol. Sci.* 14, 8122–8147. doi:10.3390/ijms14048122.
- Cho, I. H., and Peterson, D. G. (2010). Chemistry of bread aroma: A review. *Food Sci. Biotechnol.* 19, 575–582.
- Core Development Team, R. (2020). A Language and Environment for Statistical Computing. R Found. Stat. Comput. 2, <https://www.R-project.org>. Available at: <http://www.r-project.org>.
- Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Syst. Complex Sy.* 1695. Available at: <http://igraph.sf.net>.
- Díaz-Mendoza, M., Velasco-Arroyo, B., González-Melendi, P., Martínez, M., and Díaz, I. (2014). C1A cysteine protease-cystatin interactions in leaf senescence. *J. Exp. Bot.* 65, 3825–3833. doi:10.1093/jxb/eru043.



- Gao, Y., An, K., Guo, W., Chen, Y., Zhang, R., Zhang, X., et al. (2021). The endosperm-specific transcription factor TaNAC019 regulates glutenin and starch accumulation and its elite allele improves wheat grain quality. *Plant Cell* 33, 603–622. doi:10.1093/plcell/koaa040.
- García-Molina, M. D., Giménez, M. J., Sánchez-León, S., and Barro, F. (2019). Gluten free wheat: Are we there? *Nutrients* 11, 487. doi:10.3390/nu11030487.
- Gautam, T., Saripalli, G., Gahlaut, V., Kumar, A., Sharma, P. K., Balyan, H. S., et al. (2019). Further studies on sugar transporter (SWEET) genes in wheat (*Triticum aestivum* L.). *Mol. Biol. Rep.* 46, 2327–2353. doi:10.1007/s11033-019-04691-0.
- Gil-Humanes, J., Pistón, F., Altamirano-Fortoul, R., Real, A., Comino, I., Sousa, C., et al. (2014a). Reduced-gliadin wheat bread: An alternative to the gluten-free diet for consumers suffering gluten-related pathologies. *PLoS One* 9, e90898. doi:10.1371/journal.pone.0090898.
- Gil-Humanes, J., Pistón, F., Barro, F., and Rosell, C. M. (2014b). The shutdown of celiac disease-related gliadin epitopes in bread wheat by RNAi provides flours with increased stability and better tolerance to over-mixing. *PLoS One* 9, e91931. doi:10.1371/journal.pone.0091931.
- Gil-Humanes, J., Pistón, F., Rosell, C. M., and Barro, F. (2012). Significant down-regulation of γ -gliadins has minor effect on gluten and starch properties of bread wheat. *J. Cereal Sci.* 56, 161–170. doi:10.1016/j.jcs.2012.02.009.
- Gil-Humanes, J., Pistón, F., Shewry, P. R., Tosi, P., and Barro, F. (2011). Suppression of gliadins results in altered protein body morphology in wheat. *J. Exp. Bot.* 62, 4203–4213. doi:10.1093/jxb/err119.
- Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L. M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proc. Natl. Acad. Sci. U. S. A.* 107, 17023–17028. doi:10.1073/pnas.1007773107.
- Giménez, M. J., Pistón, F., and Atienza, S. G. (2011). Identification of suitable reference genes for normalization of qPCR data in comparative transcriptomics analyses in the Triticeae. *Planta* 233, 163–173. doi:10.1007/s00425-010-1290-y.
- Guerin, C., Roche, J., Allard, V., Ravel, C., Mouzeyar, S., and Bouzidi, M. F. (2019). Genome-wide analysis, expansion and expression of the NAC family under drought and heat stresses in bread wheat (*T. aestivum* L.). *PLoS One* 14, e0213390.
- Guo, D., Hou, Q., Zhang, R., Lou, H., Li, Y., Zhang, Y., et al. (2020). Over-expressing *taspa-b* reduces prolamin and starch accumulation in wheat (*Triticum aestivum* L.) grains. *Int. J. Mol. Sci.* 21, 3257. doi:10.3390/ijms21093257.
- Guo, W., Yang, H., Liu, Y., Gao, Y., Ni, Z., Peng, H., et al. (2015). The wheat transcription factor TaGAMYB recruits histone acetyltransferase and activates the expression of a high-molecular-weight glutenin subunit gene. *Plant J.* 84, 347–359. doi:10.1111/tpj.13003.
- Guo, Y., Cai, Z., and Gan, S. (2004). Transcriptome of Arabidopsis leaf senescence. *Plant, Cell Environ.* 27, 521–549. doi:10.1111/j.1365-3040.2003.01158.x.
- Guzmán-López, M. H., Sánchez-León, S., Marín-Sanz, M., Comino, I., Segura, V., Vaquero, L., et al. (2021). Oral Consumption of Bread from an RNAi Wheat Line with Strongly Silenced Gliadins Elicits No Immunogenic Response in a Pilot Study with Celiac Disease Patients. *Nutrients* 13, 4548.
- Haro, C., Villatoro, M., Vaquero, L., Pastor, J., Giménez, M. J., Ozuna, C. V., et al. (2018). The dietary intervention of transgenic low-gliadin wheat bread in patients with non-celiac gluten sensitivity (NCGS) showed no differences with gluten free diet (GFD) but provides better gut microbiota profile. *Nutrients* 10, 1964. doi:10.3390/nu10121964.



- Harrington, S. A., Backhaus, A. E., Singh, A., Hassani-Pak, K., and Uauy, C. (2020). The wheat GENIE3 network provides biologically-Relevant information in polyploid wheat. *G3 Genes, Genomes, Genet.* 10, 3675–3686. doi:10.1534/g3.120.401436.
- Huang, J., Zheng, J., Yuan, H., and McGinnis, K. (2018). Distinct tissue-specific transcriptional regulation revealed by gene regulatory networks in maize. *BMC Plant Biol.* 18, 1–14. doi:10.1186/s12870-018-1329-y.
- Huang, S., Adams, W. R., Zhou, Q., Malloy, K. P., Voyles, D. A., Anthony, J., et al. (2004). Improving nutritional quality of maize proteins by expressing sense and antisense zein genes. *J. Agric. Food Chem.* 52, 1958–1964.
- Huo, N., Zhang, S., Zhu, T., Dong, L., Wang, Y., Mohr, T., et al. (2018a). Gene duplication and evolution dynamics in the homeologous regions harboring multiple prolamin and resistance gene families in hexaploid wheat. *Front. Plant Sci.* 9, 673. doi:10.3389/fpls.2018.00673.
- Huo, N., Zhu, T., Altenbach, S., Dong, L., Wang, Y., Mohr, T., et al. (2018b). Dynamic Evolution of α -Gliadin Prolamin Gene Family in Homeologous Genomes of Hexaploid Wheat. *Sci. Rep.* 8, 1–13. doi:10.1038/s41598-018-23570-5.
- Iida, S., Amano, E., and Nishio, T. (1993). A rice (*Oryza sativa* L.) mutant having a low content of glutelin and a high content of prolamine. *Theor. Appl. Genet.* 87, 374–378.
- Juhász, A., Belova, T., Florides, C. G., Maulis, C., Fischer, I., Gell, G., et al. (2018). Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci. Adv.* 4, eaar8602. doi:10.1126/sciadv.aar8602.
- Kawakatsu, T., and Takaiwa, F. (2010). Cereal seed storage protein synthesis: Fundamental processes for recombinant protein production in cereal grains. *Plant Biotechnol. J.* 8, 939–953. doi:10.1111/j.1467-7652.2010.00559.x.
- Kohl, S., Hollmann, J., Blattner, F. R., Radchuk, V., Andersch, F., Steuernagel, B., et al. (2012). A putative role for amino acid permeases in sink-source communication of barley tissues uncovered by RNA-seq. *BMC Plant Biol.* 12, 1–18. doi:10.1186/1471-2229-12-154.
- Kovalchuk, N., Wu, W., Eini, O., Bazanova, N., Pallotta, M., Shirley, N., et al. (2012). The scutellar vascular bundle-specific promoter of the wheat HD-Zip IV transcription factor shows similar spatial and temporal activity in transgenic wheat, barley and rice. *Plant Biotechnol. J.* 10, 43–53. doi:10.1111/j.1467-7652.2011.00633.x.
- Kumar, P., Mishra, A., Sharma, H., Sharma, D., Rahim, M. S., Sharma, M., et al. (2018). Pivotal role of bZIPs in amylose biosynthesis by genome survey and transcriptome analysis in wheat (*Triticum aestivum* L.) mutants. *Sci. Rep.* 8, 1–15. doi:10.1038/s41598-018-35366-8.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi:10.1093/bioinformatics/btp352.
- Liu, H., Hu, M., Wang, Q., Cheng, L., and Zhang, Z. (2018). Role of papain-like cysteine proteases in plant development. *Front. Plant Sci.* 871, 1717. doi:10.3389/fpls.2018.01717.
- Luo, G., Shen, L., Zhao, S., Li, R., Song, Y., Song, S., et al. (2021). Genome-wide identification of seed storage protein gene regulators in wheat through coexpression analysis. *Plant J.* 108, 1704–1720.
- Marín-Sanz, M., Iehisa, J. C. M., and Barro, F. (2022). New transcriptomic insights in two RNAi wheat lines with the gliadins strongly down-regulated by two endosperm specific promoters. *Crop J.* 10, 194–203. doi:10.1016/j.cj.2021.04.009.
- Masclaux-Daubresse, C., Reisdorf-Cren, M., and Orsel, M. (2008). Leaf nitrogen remobilisation for plant development and grain filling. *Plant Biol.* 10, 23–36. doi:10.1111/j.1438-8677.2008.00097.x.



- Mena, M., Vicente-Carbajosa, J., Schmidt, R. J., and Carbonero, P. (1998). An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. *Plant J.* 16, 53–62. doi:10.1046/j.1365-313X.1998.00275.x.
- Meng, L., and Feldman, L. (2010). A rapid TRIzol-based two-step method for DNA-free RNA extraction from *Arabidopsis* siliques and dry seeds. Wiley Online Library.
- Moehs, C. P., Austill, W. J., Holm, A., Large, T. A. G., Loeffler, D., Mullenberg, J., et al. (2019). Development of decreased-gluten wheat enabled by determination of the genetic basis of lys3a barley. *Plant Physiol.* 179, 1692–1703.
- Murata, T. (2010). “Detecting Communities in Social Networks,” in *Handbook of Social Network Technologies and Applications* (Springer), 269–280. doi:10.1007/978-1-4419-7142-5_12.
- Orman-Ligeza, B., Borrill, P., Chia, T., Chirico, M., Doležal, J., Drea, S., et al. (2020). LYS3 encodes a prolamin-box-binding transcription factor that controls embryo growth in barley and wheat. *J. Cereal Sci.* 93, 102965. doi:10.1016/j.jcs.2020.102965.
- Ozuna, C. V., Iehisa, J. C. M., Giménez, M. J., Alvarez, J. B., Sousa, C., and Barro, F. (2015). Diversification of the celiac disease α -gliadin complex in wheat: A 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *Plant J.* 82, 794–805. doi:10.1111/tpj.12851.
- PEETERS, K. M. U., and VAN LAERE, A. J. (1994). Amino acid metabolism associated with N-mobilization from the flag leaf of wheat (*Triticum aestivum* L.) during grain development. *Plant. Cell Environ.* 17, 131–141. doi:10.1111/j.1365-3040.1994.tb00276.x.
- Pistón, F., Gil-Humanes, J., and Barro, F. (2013). Integration of promoters, inverted repeat sequences and proteomic data into a model for high silencing efficiency of coeliac disease related gliadins in bread wheat. *BMC Plant Biol.* 13, 136. doi:10.1186/1471-2229-13-136.
- Pistón, F., Gil-Humanes, J., Rodríguez-Quijano, M., and Barro, F. (2011). Down-Regulating γ -Gliadins in bread wheat leads to Non-Specific increases in other gluten proteins and has no major effect on dough gluten strength. *PLoS One* 6, e24754. doi:10.1371/journal.pone.0024754.
- Pistón, F., León, E., Lazzeri, P. A., and Barro, F. (2008). Isolation of two storage protein promoters from *Hordeum chilense* and characterization of their expression patterns in transgenic wheat. *Euphytica* 162, 371–379. doi:10.1007/s10681-007-9530-3.
- Pistón, F., Marín, S., Hernando, A., and Barro, F. (2009). Analysis of the activity of a γ -gliadin promoter in transgenic wheat and characterization of gliadin synthesis in wheat by MALDI-TOF during grain development. *Mol. Breed.* 23, 655–667. doi:10.1007/s11032-009-9263-1.
- Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of polyploid wheat. *Science* (80-.). 361. doi:10.1126/science.aar6089.
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). G:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 47, W191–W198. doi:10.1093/nar/gkz369.
- Rawlings, N. D., Barrett, A. J., and Finn, R. (2016). Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 44, D343–D350. doi:10.1093/nar/gkv1118.
- Rawlings, N. D., and Salvesen, G. (2013). *Handbook of Proteolytic Enzymes*. Academic press doi:10.1016/C2009-1-60990-4.
- Roberts, I. N., Caputo, C., Criado, M. V., and Funk, C. (2012). Senescence-associated proteases in plants. *Physiol. Plant.* 145, 130–139. doi:10.1111/j.1399-3054.2012.01574.x.



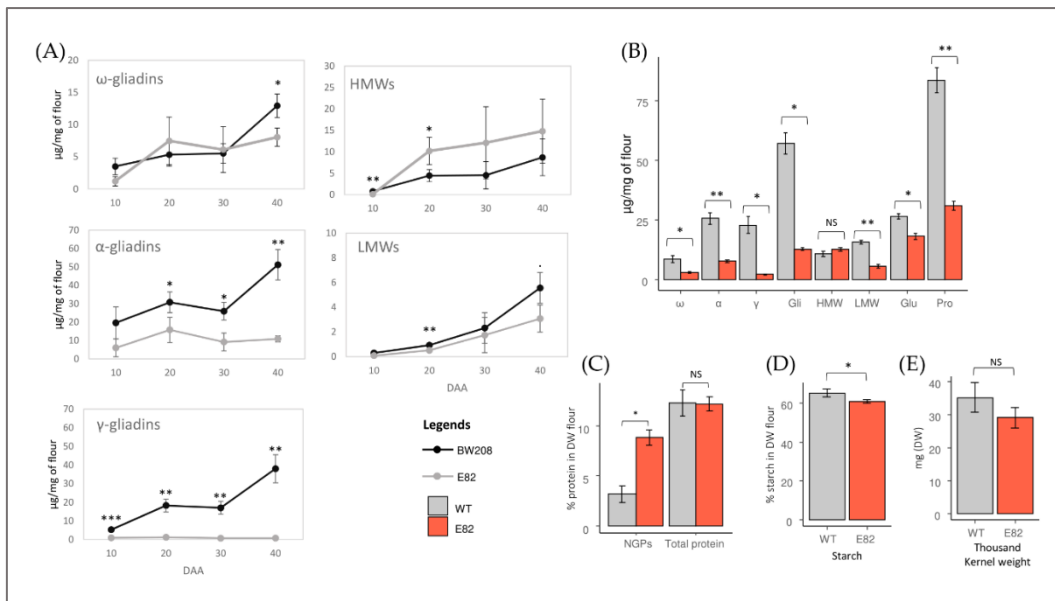
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi:10.1093/bioinformatics/btp616.
- Sánchez-León, S., Gil-Humanes, J., Ozuna, C. V., Giménez, M. J., Sousa, C., Voytas, D. F., et al. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant Biotechnol. J.* 16, 902–910. doi:10.1111/pbi.12837.
- Sánchez-León, S., Giménez, M. J., Comino, I., Sousa, C., López Casado, M. Á., Torres, M. I., et al. (2019). Stimulatory response of celiac disease peripheral blood mononuclear cells induced by RNAi wheat lines differing in grain protein composition. *Nutrients* 11, 2933.
- Shevkani, K., Singh, N., Bajaj, R., and Kaur, A. (2017). Wheat starch production, structure, functionality and applications—a review. *Int. J. Food Sci. Technol.* 52, 38–58. doi:10.1111/ijfs.13266.
- Shewry, P. R. (2009). Wheat. *J. Exp. Bot.* 60, 1537–1553.
- Shewry, P. R., D'Ovidio, R., Lafiandra, D., Jenkins, J. A., Mills, E. N. C., and Békés, F. (2009). Wheat grain proteins. *Wheat Chem. Technol.*, 223–298.
- Shewry, P. R., Faulks, A. J., and Miflin, B. J. (1980). Effect of high-lysine mutations on the protein fractions of barley grain. *Biochem. Genet.* 18, 133–151.
- Shewry, P. R., and Halford, N. G. (2002). Cereal seed storage proteins: Structures, properties and role in grain utilization. *J. Exp. Bot.* 53, 947–958. doi:10.1093/jexbot/53.370.947.
- Smith, M. R., Rao, I. M., and Merchant, A. (2018). Source-sink relationships in crop plants and their influence on yield development and nutritional quality. *Front. Plant Sci.* 871, 1889. doi:10.3389/fpls.2018.01889.
- Sollid, L. M. (2002). Coeliac disease: Dissecting a complex inflammatory disorder. *Nat. Rev. Immunol.* 2, 647–655. doi:10.1038/nri885.
- Soneson, C., Love, M. I., and Robinson, M. D. (2016). Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research* 4. doi:10.12688/F1000RESEARCH.7563.2.
- Sun, F., Liu, X., Wei, Q., Liu, J., Yang, T., Jia, L., et al. (2017). Functional characterization of TaFUSCA3, a B3-superfamily transcription factor gene in the wheat. *Front. Plant Sci.* 8, 1133. doi:10.3389/fpls.2017.01133.
- Tian, R., Yang, Y., and Chen, M. (2020). Genome-wide survey of the amino acid transporter gene family in wheat (*Triticum aestivum* L.): Identification, expression analysis and response to abiotic stress. *Int. J. Biol. Macromol.* 162, 1372–1387. doi:10.1016/j.ijbiomac.2020.07.302.
- Ueda, A., Shi, W., Shimada, T., Miyake, H., and Takabe, T. (2008). Altered expression of barley proline transporter causes different growth responses in *Arabidopsis*. *Planta* 227, 277–286. doi:10.1007/s00425-007-0615-y.
- Vaquero, L., Comino, I., Vivas, S., Rodríguez-Martín, L., Giménez, M. J., Pastor, J., et al. (2018). Tritordeum: a novel cereal for food processing with good acceptability and significant reduction in gluten immunogenic peptides in comparison with wheat. *J. Sci. Food Agric.* 98, 2201–2209. doi:10.1002/jsfa.8705.
- Wang, J. C., Xu, H., Zhu, Y., Liu, Q. Q., and Cai, X. L. (2013). OsbZIP58, a basic leucine zipper transcription factor, regulates starch biosynthesis in rice endosperm. *J. Exp. Bot.* 64, 3453–3466. doi:10.1093/jxb/ert187.
- Yu, S. M., Lo, S. F., and Ho, T. H. D. (2015). Source-Sink Communication: Regulated by Hormone, Nutrient, and Stress Cross-Signaling. *Trends Plant Sci.* 20, 844–857. doi:10.1016/j.tplants.2015.10.009.
- Zhang, Z., Dong, J., Ji, C., Wu, Y., and Messing, J. (2019). NAC-type transcription factors regulate accumulation of starch and protein in maize seeds. *Proc. Natl. Acad. Sci. U. S. A.* 166, 11223–11228. doi:10.1073/pnas.1904995116.



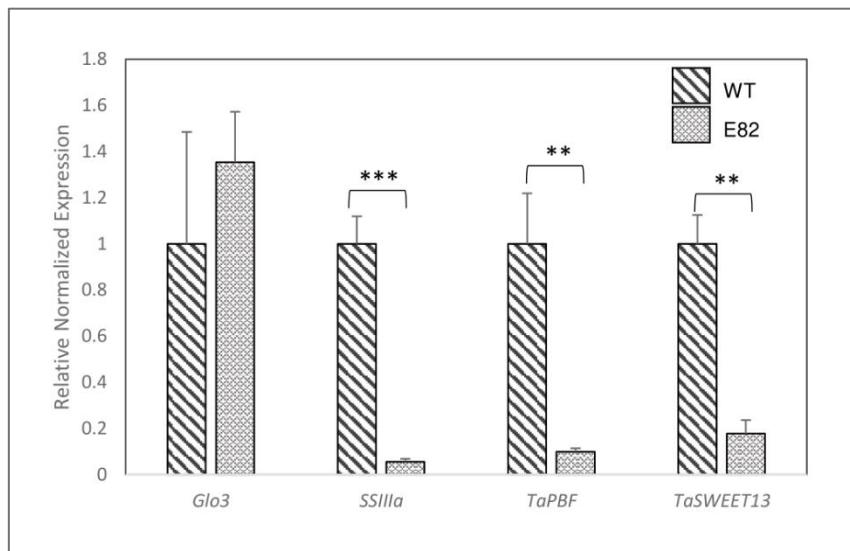
Zhang, Z., Zheng, X., Yang, J., Messing, J., and Wu, Y. (2016). Maize endosperm-specific transcription factors O2 and PBF network the regulation of protein and starch synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 113, 10842–10847. doi:10.1073/pnas.1613721113.



6. Supplementary Materials



Supplementary Figure 4.1. (A) Content of gliadins and glutenins of the wild type (WT) and E82 throughout grain development. Bars indicate standard deviation. ANOVA was performed for each DAA between BW208 and E82. (B) Comparative gliadin and glutenin grain protein content at harvest between the WT and E82, (C) non-gluten proteins (NGP) and total grain protein contents at harvest between the WT and E82, (D) Starch content in grain tissue at harvest for BW208 and E82. (E) Thousand kernel weight for the WT and E82. Bars indicate the standard error of three biological replicates. ·, $P \sim 0.05$; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; NS, non-significant. Gli: total gliadins; Glu: total glutenins; Pro: total prolamins; DW: dry weight; DAA: days after anthesis.



Supplementary Figure 4.2. Relative normalized expression of the wild type (WT) and E82 grains and leaves at 20 DAA for *TaPBF*, *Glo3*, *SSIIIa*, and *TaSWEET13* genes by qPCR. Bars indicate the standard error of three biological replicates. **, $P < 0.01$; ***, $P < 0.001$.



Supplementary Table 4.1. Trimming of raw reads for RNA-seq data analysis. PE reads: paired-end reads, Rep: replicate.

Sample	Number of PE reads	Number of clean PE reads	Effective rate (%)	Forward reads only surviving (%)	Reverse reads only surviving (%)	Dropped reads (%)
BW208 Leaf Rep 1	74,752,131	73,037,333	97.71	2.19	0.04	0.06
BW208 Leaf Rep 2	71,353,011	69,950,624	98.03	1.85	0.07	0.05
BW208 Leaf Rep 3	56,516,174	55,259,622	97.78	2.04	0.09	0.10
E82 Leaf Rep 1	84,242,149	83,170,383	98.73	1.24	0.03	0.00
E82 Leaf Rep 2	81,215,756	79,890,887	98.37	1.59	0.03	0.01
E82 Leaf Rep 3	78,480,143	77,444,904	98.68	1.28	0.03	0.01
BW208 Grain Rep 1	74,518,649	73,520,129	98.66	1.27	0.06	0.01
BW208 Grain Rep 2	69,791,185	68,721,043	98.47	1.48	0.04	0.01
BW208 Grain Rep 3	72,447,691	71,661,718	98.92	1.04	0.04	0.01
E82 Grain Rep 1	66,306,206	66,306,206	99.09	0.86	0.03	0.01
E82 Grain Rep 2	70,719,358	69,495,582	98.27	1.67	0.04	0.02
E82 Grain Rep 3	50,926,417	50,231,971	98.64	1.31	0.04	0.02

Supplementary Table 4.2. Results of quality control of aligned reads by samtools. Parameters: default parameters. Reads are considered as the sum of forward and reverse ones. QC: quality control. Rep: replicate.

Sample	QC passed reads	QC failed reads	Number of reads secondary	Number of mapped reads (forward and reverse separately)	Rate of mapped reads (%)	Number of reads paired in sequencing	Reads properly paired (%)	Singletons (%)
BW208 Leaf Rep 1	165,858,823	0	19,784,157	97,917,955	59.04	146,074,666	51.72	1.77



Sample	QC passed reads	QC failed reads	Number of reads secondary	Number of mapped reads (forward and reverse separately)	Rate of mapped reads (%)	Number of reads paired in sequencing	Reads properly paired (%)	Singletons (%)
BW208 Leaf Rep 2	159,944,144	0	20,042,896	105,211,830	65.78	139,901,248	58.05	2.83
BW208 Leaf Rep 3	122,664,967	0	12,145,723	69,129,593	56.36	110,519,244	49.68	1.89
E82 Leaf Rep 1	209,745,327	0	43,404,561	171,003,105	81.53	166,340,766	74.38	2.33
E82 Leaf Rep 2	196,404,544	0	36,622,770	155,214,808	79.03	159,781,774	71.56	2.66
E82 Leaf Rep 3	195,765,253	0	40,875,445	160,022,010	81.74	154,889,808	74.76	2.17
BW208 Grain Rep 1	172,826,311	0	25,786,053	133,146,643	77.04	147,040,258	70.77	2.24
BW208 Grain Rep 2	161,795,752	0	24,353,666	126,835,486	78.39	137,442,086	72.18	2.38
BW208 Grain Rep 3	166,343,532	0	23,020,096	129,988,550	78.14	143,323,436	72.74	1.89
E82 Grain Rep 1	152,749,397	0	21,338,639	118,389,468	77.51	131,410,758	71.81	2.04
E82 Grain Rep 2	161,221,617	0	22,230,453	121,834,633	75.57	138,991,164	75.57	1.57
E82 Grain Rep 3	116,361,061	0	15,897,119	87,716,282	75.38	100,463,942	69.80	1.69

Supplementary Table 4.3. List of primers for the qPCR assay for gene expression analysis.

Gene name	Gene ID	Primer sequence (5'-3')	Reference
<i>TaPBF^a</i>	<i>TraesCS5A02G155900</i>	AGAGCAGAAGGTGGAATGCC	This work.
	<i>TraesCS5B02G154100</i>	GTAGCGGGGCTGAGACATAC	
<i>Glo3^a</i>	<i>TraesCS5D02G161000</i>		This work.
	<i>TraesCS4A02G296100</i>	ATCGCCAAGATTCTCCACAC	
<i>SSIII^a</i>	<i>TraesCS4A02G296000</i>	GGACGATGGAGATGGACTTC	This work.
	<i>TraesCS1B02G119300</i>	TGGAATGGAGGCTTTTCAC	
	<i>TraesCS1D02G100100</i>	TGTAGGCCTCCTTGGGTATG	



Gene name	Gene ID	Primer sequence (5'-3')	Reference
<i>TraesCS1A02G091500</i>			
<i>TaSWEET13^b</i>	<i>TraesCS6B02G421800</i>	CGAGTCATCAAGACCAAGAGTG	This work.
	<i>TraesCS6D02G367400</i>	AGGCCGTAGAGGAACCAGAC	
<i>CDC^c</i>	EU267938	CAGCTGCTGACTGAGATGGA ATGTCTGGCCTGTTGGTAGC	Giménez et al., 2011
<i>ADP-RF^c</i>	AB050957	TTCATGGTTGGTCTCGATG GGATGGTGGTGACGATCTCT	Giménez et al., 2011
<i>RLI^c</i>	AK331207	TTGAGCAACTCATGGACCAG GCTTCCAAGGCACAAACAT	Giménez et al., 2011

^a Three homeologs are amplified.

^b Two homeologs are amplified.

^c The accession number from NCBI is indicated.

Supplementary Table 4.4. Results of the alignment of each sample by kallisto. Parameters: default parameters and --pseudobam. PE: paired-end reads. Rep: replicate.

Sample	Number of clean PE reads	Number of aligned PE reads	Rate of alignment (%)
BW208 Leaf Rep 1	73,037,333	40,356,415	55.25
BW208 Leaf Rep 2	69,950,624	44,562,509	63.71
BW208 Leaf Rep 3	55,259,622	29,533,590	53.45
E82 Leaf Rep 1	83,170,383	65,734,409	79.04
E82 Leaf Rep 2	79,890,887	61,420,333	76.88
E82 Leaf Rep 3	77,444,904	61,251,809	79.09
BW208 Grain Rep 1	73,520,129	55,329,460	75.26
BW208 Grain Rep 2	68,721,043	52,879,418	76.95
BW208 Grain Rep 3	71,661,718	54,840,952	76.53
E82 Grain Rep 1	65,705,379	49,868,499	75.90



Sample	Number of clean PE reads	Number of aligned PE reads	Rate of alignment (%)
E82 Grain Rep 2	69,495,582	50,894,143	73.23
E82 Grain Rep 3	50,231,971	36,756,483	73.17

Supplementary Table 4.5. Proteomic data for grain proteins at harvest in the wild type (WT) and E82. The average of two biological replicates and the results of the t-test for the WT vs E82 comparison are represented.

Protein	Enzyme	WT		E82		Significance of WT vs E82
		Number unique peptides ^a	SD ^b	Number unique peptides ^a	SD ^b	
ALPs	Trypsin	93	7.1	103.5	6.4	NS
ATIs	Trypsin	202	25.5	200.5	9.2	NS
Triticins	Trypsin	5.5	2.1	23.5	0.7	**
ω -gliadins	Chymotrypsin	27.5	2.1	4.5	0.7	**
α -gliadins	Chymotrypsin	97.5	16.3	22	0	*
γ -gliadins	Chymotrypsin	107	9.9	40.5	0.7	*
HMWs	Trypsin	215.5	19.1	258.5	20.5	NS
LMWs	Trypsin	150	11.3	103	7.1	*
Globulins	Trypsin	29.5	0.7	57.5	3.5	**
LTPs	Trypsin	24.5	0.7	29	0	*
Serpins	Trypsin	71.5	9.2	179	26.9	*

*, $P < 0.05$.

** , $P < 0.01$.

***, $P < 0.001$.

NS, Non-significant.

^a Number of unique peptides: peptides with e -value < 0.05 , non-redundant, included in proteins with more than 1 peptide assigned.

^b SD: standard deviation.



Supplementary Table 4.6. Differential Gene Expression (DGE) analysis of Starch Synthesis-Related Genes (SSRGs) for the pair-wise comparison wild type (WT) vs E82.

Gene name	Gene ID	Grain		Leaf	
		WT vs E82		WT vs E82	
		Log ₂ FC	FDR	Log ₂ FC	FDR
<i>SUSase</i>	<i>TraesCS7B02G063400</i>	-0.16	ns	-1.57	*
<i>SUSase</i>	<i>TraesCS7D02G159800</i>	-0.29	ns	-1.02	*
<i>APL3/AGPL1</i>	<i>TraesCS1B02G449700</i>	-2.36	*	-2.27	*
<i>APL3/AGPL1</i>	<i>TraesCS1D02G427400</i>	-2.47	*	-1.55	ns
<i>APS1/AGPS1/Bt2</i>	<i>TraesCS7A02G287400</i>	-2.17	*	0.01	ns
<i>ISA3</i>	<i>TraesCS5A02G248700</i>	0.26	ns	5.50	***
<i>OsBT1-1</i>	<i>TraesCS6B02G210000</i>	-12.45	***	NA	ns
<i>OsBT1-1</i>	<i>TraesCS6D02G168200</i>	-11.64	***	NA	ns
<i>GBE1/BEI1b/Ae</i>	<i>TraesCS2A02G310300</i>	-5.59	***	7.76	***
<i>GBE1</i>	<i>TraesCS7A02G549100</i>	-3.08	**	-0.51	ns
<i>GBE1</i>	<i>TraesCS7B02G472500</i>	-2.75	ns	9.92	***
<i>SSI</i>	<i>TraesCS7B02G018600</i>	-1.85	*	0.89	ns
<i>SSII a/SSII-3</i>	<i>TraesCS7B02G093800</i>	-4.80	***	5.24	ns
<i>SSII a/SSII-3</i>	<i>TraesCS7D02G190100</i>	-3.57	*	NA	ns
<i>GBSSII</i>	<i>TraesCS2B02G390700</i>	-0.11	ns	2.43	**
<i>GBSSI</i>	<i>TraesCS4A02G418200</i>	-3.53	**	NA	ns
<i>GBSSI/waxy</i>	<i>TraesCS7A02G070100</i>	-3.24	**	-6.54	***
<i>GBSSI/waxy</i>	<i>TraesCS7D02G064300</i>	-3.98	***	NA	ns
<i>SSIIIa/SSIII-2</i>	<i>TraesCS1B02G119300</i>	-6.78	**	-2.92	ns
<i>SPS</i>	<i>TraesCS4B02G091100</i>	0.34	ns	1.43	**

FC: Fold-Change.

FDR: False Discovery Rate.

ns: non-significant; *, FDR < 0.05; **, FDR < 0.01; ***, FDR < 0.001.

NA: Non-applicable. Genes not expressed in either the WT or the E82.



Supplementary Table 4.7. Differential Gene Expression (DGE) analysis of well-known Transcription Factors (TFs) genes that regulate Seed Storage Proteins (SSPs) genes and Starch Synthesis-Related Genes (SSRGs) expression. Significant results (FDR < 0.05) are indicated.

TF name	Gene ID	Grain WT vs E82	
		Log ₂ FC	FDR
<i>TaFUSCA3</i>	<i>TraesCS3B02G278000</i>	-0.07	ns
<i>TaFUSCA3</i>	<i>TraesCS3D02G249100</i>	-0.27	ns
<i>TaFUSCA3</i>	<i>TraesCS3A02G249100</i>	0.18	ns
<i>TaPBF</i>	<i>TraesCS5A02G155900</i>	-6.10	**
<i>TaPBF</i>	<i>TraesCS5B02G154100</i>	-4.98	ns
<i>TaPBF</i>	<i>TraesCS5D02G161000</i>	-5.24	ns
<i>TaSHP</i>	<i>TraesCS5A02G440400</i>	-0.21	ns
<i>TaSHP</i>	<i>TraesCS5D02G447500</i>	0.06	ns
<i>TaSHP</i>	<i>TraesCS5B02G444100</i>	-0.36	ns
<i>GAMYB</i>	<i>TraesCS3A02G336500</i>	-2.66	ns
<i>GAMYB</i>	<i>TraesCS3B02G367500</i>	-0.22	ns
<i>GAMYB</i>	<i>TraesCS3D02G329400</i>	-0.46	ns
<i>MCB1</i>	<i>TraesCS1D02G220500</i>	-0.26	ns
<i>MCB1</i>	<i>TraesCS1B02G232300</i>	-0.37	ns
<i>MCB1</i>	<i>TraesCS1A02G218800</i>	-0.15	ns
<i>MYBS3</i>	<i>TraesCS1A02G219400</i>	-0.34	ns
<i>MYBS3</i>	<i>TraesCS1D02G221000</i>	0.01	ns
<i>MYBS3</i>	<i>TraesCS1B02G232800</i>	0.02	ns
<i>TBP1</i>	<i>TraesCS1A02G139600</i>	-0.38	ns
<i>TBP1</i>	<i>TraesCS1B02G151700</i>	-0.61	ns
<i>TBP2</i>	<i>TraesCS4D02G014000</i>	0.37	ns
<i>TBP2</i>	<i>TraesCS4B02G015700</i>	0.07	ns
<i>TBP2</i>	<i>TraesCS4A02G298000</i>	0.25	ns
<i>TaRSR1</i>	<i>TraesCS1B02G076300</i>	-0.50	ns
<i>TaRSR1</i>	<i>TraesCS1A02G058400</i>	-0.24	ns
<i>TaRSR1</i>	<i>TraesCS1D02G059200</i>	0.10	ns
<i>TaGBF1</i>	<i>TraesCS1A02G409800</i>	-0.44	ns
<i>TaGBF1</i>	<i>TraesCS1B02G439800</i>	-0.65	ns



TF name	Gene ID	Grain WT vs E82	
		Log ₂ FC	FDR
<i>TaGBF1</i>	<i>TraesCS1D02G417100</i>	0.32	ns
<i>TabZIP229.1</i>	<i>TraesCS7B02G114300</i>	1.24	ns
<i>TabZIP206</i>	<i>TraesCS7B02G391800</i>	0.50	ns
<i>TabZIP217.1</i>	<i>TraesCS7D02G475100</i>	-1.15	ns
<i>TabZIP110</i>	<i>TraesCS4B02G175800</i>	-0.14	ns
<i>TabZIP236</i>	<i>TraesCS1D02G306000</i>	-0.40	ns
<i>TabZIP151</i>	<i>TraesCS5D02G447400</i>	0.36	ns
<i>TabZIP194.3</i>	<i>TraesCS7A02G488600</i>	0.94	ns
<i>TabZIP117.1</i>	<i>TraesCS4B02G178600</i>	0.06	ns
<i>TabZIP229.1</i>	<i>TraesCS7B02G114300</i>	1.24	ns
<i>TabZIP238.1</i>	<i>TraesCS7A02G207100</i>	1.50	ns
<i>TabZIP167.2</i>	<i>TraesCS5D02G178800</i>	-3.15	ns
<i>TabZIP184.2</i>	<i>TraesCS6B02G193200</i>	-0.12	ns
<i>TabZIP59.2</i>	<i>TraesCS3B02G404800</i>	0.40	ns
<i>TabZIP77.1</i>	<i>TraesCS3D02G365200</i>	-0.13	ns
<i>TabZIP145.3</i>	<i>TraesCS5B02G142200</i>	-0.14	ns
<i>TabZIP101.1</i>	<i>TraesCS4D02G230200</i>	0.15	ns
<i>TabZIP137</i>	<i>TraesCS5A02G299400</i>	-0.32	ns
<i>TabZIP157.1</i>	<i>TraesCS5D02G308600</i>	0.64	ns
<i>TabZIP111</i>	<i>TraesCS4B02G113400</i>	-0.37	ns
<i>TabZIP121</i>	<i>TraesCS4D02G115200</i>	0.27	ns
<i>TabZIP54.1</i>	<i>TraesCS2B02G269600</i>	0.42	ns

FC: Fold-Change.

FDR: False Discovery Rate.

ns: non-significant; *, FDR < 0.05; **, FDR < 0.01; ***, FDR < 0.001.



Supplementary Table 4.8. Top 3 Gene Ontology (GO) terms enriched for each candidate transcription factor (TF) summarized from previously published GO enrichment analysis for GENIE3 network genes in Ramírez-González et al., 2018. The black boxes indicate that this is an enriched GO for this TF gene.

Community	ID in candidate network	Candidate TF	Annotation	Carbohydrate mediated signaling	Seed oilbody biogenesis	Negative regulation of endopeptidase activity	Glycogen biosynthetic process	Negative regulation of molecular function	Glycogen metabolic process	Negative regulation of catalytic activity	Photosynthesis	Plastid organization	Inorganic anion transport	Defense response	Sugar mediated signaling pathway	Lipid storage	Regulation of endopeptidase activity	Regulation of cellular metabolic process	Regulation of macromolecule metabolic	Energy reserve metabolic process	Innate immune response	Cellular response to carbohydrate stimulus	Siderophore metabolic process	Gibberellin biosynthetic process	Negative regulation of peptidase activity	Regulation of catalytic activity	Light reaction	Immune system process	Glyceraldehyde-3-phosphate metabolic	Detection of nutrient
Community 5	1	<i>TraesCS1A02G275800</i>	<i>AT2G38090</i>	■											■															
	2	<i>TraesCS1B02G285000</i>	<i>AT2G38090</i>	■	■										■	■														
	4	<i>TraesCS1D02G275400</i>	<i>AT2G38090</i>	■											■								■							
	6	<i>TraesCS3A02G486500</i>	<i>NAC057</i>		■																			■						
	8	<i>TraesCS3B02G092800</i>	<i>TaNAC019-B</i>			■												■	■											
	17	<i>TraesCS5A02G245900</i>	<i>ANAC087</i>				■											■	■	■										
	18	<i>TraesCS7A02G194700</i>	<i>ANAC087</i>				■											■	■	■										
	22	<i>TraesCS7A02G569100</i>	<i>NAC038</i>			■												■	■	■				■						
	23	<i>TraesCS7B02G056300</i>	<i>NAC100</i>			■												■	■	■										■
	24	<i>TraesCS7B02G100300</i>	<i>ANAC087</i>			■												■	■	■										
	27	<i>TraesCS7B02G489500</i>	<i>NAC038</i>			■												■	■	■						■				
	30	<i>TraesCS7D02G543500</i>	<i>NAC038</i>			■					■							■	■	■						■				
Community 4	3	<i>TraesCS1B02G443800</i>	<i>RGA1</i>					■																						
	5	<i>TraesCS3A02G325800</i>	<i>TaGL9-A</i>					■		■																				
	9	<i>TraesCS3B02G354900</i>	<i>TaGL9-B</i>					■		■																■				



									700, TRAESCS4D02G007200, TRAESCS4D02G022900, TRAESCS4D02G057500, TRAESCS4D02G086000, TRAESCS4D02G127500, TRAESCS4D02G152200, TRAESCS4D02G162800, TRAESCS4D02G163000, TRAESCS4D02G167100, TRAESCS4D02G171500, TRAESCS4D02G180500, TRAESCS4D02G185200, TRAESCS4D02G188500, TRAESCS4D02G199800, TRAESCS4D02G229400, TRAESCS4D02G232500, TRAESCS4D02G243900, TRAESCS4D02G246600, TRAESCS4D02G254200, TRAESCS4D02G297900, TRAESCS4D02G304500, TRAESCS4D02G318400, TRAESCS4D02G327300, TRAESCS4D02G356700, TRAESCS5A02G07300, TRAESCS5A02G094100, TRAESCS5A02G058500, TRAESCS5A02G059300, TRAESCS5A02G060400, TRAESCS5A02G069800, TRAESCS5A02G085700, TRAESCS5A02G130700, TRAESCS5A02G132300, TRAESCS5A02G146800, TRAESCS5A02G155900, TRAESCS5A02G174200, TRAESCS5A02G177400, TRAESCS5A02G200800, TRAESCS5A02G246000, TRAESCS5A02G266500, TRAESCS5A02G270700, TRAESCS5A02G274700, TRAESCS5A02G275000, TRAESCS5A02G282800, TRAESCS5A02G283100, TRAESCS5A02G283100, TRAESCS5A02G326000, TRAESCS5A02G326400, TRAESCS5A02G395200, TRAESCS5A02G434900, TRAESCS5A02G466100, TRAESCS5A02G501900, TRAESCS5A02G554000, TRAESCS5B02G054200, TRAESCS5B02G058400, TRAESCS5B02G062900, TRAESCS5B02G071300, TRAESCS5B02G071600, TRAESCS5B02G102400, TRAESCS5B02G112100, TRAESCS5B02G130600, TRAESCS5B02G154100, TRAESCS5B02G158100, TRAESCS5B02G168600, TRAESCS5B02G177800, TRAESCS5B02G179500, TRAESCS5B02G179600, TRAESCS5B02G182200, TRAESCS5B02G190100, TRAESCS5B02G199400, TRAESCS5B02G213600, TRAESCS5B02G238900, TRAESCS5B02G260500, TRAESCS5B02G271200, TRAESCS5B02G274400, TRAESCS5B02G282100, TRAESCS5B02G297300, TRAESCS5B02G379700, TRAESCS5B02G384600, TRAESCS5B02G386500, TRAESCS5B02G400000, TRAESCS5B02G404000, TRAESCS5B02G405700, TRAESCS5B02G426000, TRAESCS5B02G432900, TRAESCS5B02G465000, TRAESCS5B02G475100, TRAESCS5B02G483400, TRAESCS5B02G484300, TRAESCS5B02G527300, TRAESCS5D02G027300, TRAESCS5D02G048400, TRAESCS5D02G051200, TRAESCS5D02G051700, TRAESCS5D02G070200, TRAESCS5D02G070400, TRAESCS5D02G071900, TRAESCS5D02G072300, TRAESCS5D02G081900, TRAESCS5D02G109200, TRAESCS5D02G138200, TRAESCS5D02G161000, TRAESCS5D02G178800, TRAESCS5D02G184900, TRAESCS5D02G186000, TRAESCS5D02G199600, TRAESCS5D02G278500, TRAESCS5D02G282300, TRAESCS5D02G282600, TRAESCS5D02G289800, TRAESCS5D02G290600, TRAESCS5D02G292700, TRAESCS5D02G327300, TRAESCS5D02G328300, TRAESCS5D02G332500, TRAESCS5D02G364900, TRAESCS5D02G388600, TRAESCS5D02G396600, TRAESCS5D02G404500, TRAESCS5D02G411100, TRAESCS5D02G421000, TRAESCS5D02G453200, TRAESCS5D02G468200, TRAESCS5D02G510300, TRAESCS5D02G521200, TRAESCS5D02G550700, TRAESCS5D02G556600, TRAESCS6A02G017500, TRAESCS6A02G017600, TRAESCS6A02G017700, TRAESCS6A02G023900, TRAESCS6A02G029700, TRAESCS6A02G099600, TRAESCS6A02G105000, TRAESCS6A02G117100, TRAESCS6A02G119000, TRAESCS6A02G159400, TRAESCS6A02G160600, TRAESCS6A02G166900, TRAESCS6A02G203000, TRAESCS6A02G207600, TRAESCS6A02G219000, TRAESCS6A02G226600, TRAESCS6A02G227900, TRAESCS6A02G234000, TRAESCS6A02G235300, TRAESCS6A02G250800, TRAESCS6A02G258600, TRAESCS6A02G277800, TRAESCS6A02G278300, TRAESCS6A02G283600, TRAESCS6A02G420800, TRAESCS6B02G033700, TRAESCS6B02G036100, TRAESCS6B02G065000, TRAESCS6B02G087500, TRAESCS6B02G094300, TRAESCS6B02G096400, TRAESCS6B02G099700, TRAESCS6B02G100100, TRAESCS6B02G112600, TRAESCS6B02G118800, TRAESCS6B02G132500, TRAESCS6B02G163300, TRAESCS6B02G172900, TRAESCS6B02G191900, TRAESCS6B02G205400, TRAESCS6B02G243200, TRAESCS6B02G253900, TRAESCS6B02G256000, TRAESCS6B02G262800, TRAESCS6B02G274600, TRAESCS6B02G279300, TRAESCS6B02G292900, TRAESCS6B02G296200, TRAESCS6B02G299700, TRAESCS6B02G302500, TRAESCS6B02G305400, TRAESCS6B02G351900, TRAESCS6B02G365300, TRAESCS6B02G408800, TRAESCS6B02G422400, TRAESCS6B02G429800, TRAESCS6B02G429900, TRAESCS6B02G442400, TRAESCS6B02G442500, TRAESCS6B02G457000, TRAESCS6B02G453300, TRAESCS6B02G467500, TRAESCS6D02G006100, TRAESCS6D02G019800, TRAESCS6D02G023500, TRAESCS6D02G027200, TRAESCS6D02G029000, TRAESCS6D02G029200, TRAESCS6D02G036200, TRAESCS6D02G040500, TRAESCS6D02G040600, TRAESCS6D02G047300, TRAESCS6D02G052300, TRAESCS6D02G056900, TRAESCS6D02G057300, TRAESCS6D02G062900, TRAESCS6D02G107300, TRAESCS6D02G133700, TRAESCS6D02G138200, TRAESCS6D02G153100, TRAESCS6D02G155400, TRAESCS6D02G159400, TRAESCS6D02G164900, TRAESCS6D02G168400, TRAESCS6D02G189600, TRAESCS6D02G210400, TRAESCS6D02G236400, TRAESCS6D02G244300, TRAESCS6D02G258000, TRAESCS6D02G263200, TRAESCS6D02G268800, TRAESCS6D02G364500, TRAESCS6D02G374100, TRAESCS6D02G377200, TRAESCS6D02G405000, TRAESCS7A02G036500, TRAESCS7A02G076500, TRAESCS7A02G105100, TRAESCS7A02G134300, TRAESCS7A02G135000, TRAESCS7A02G135100, TRAESCS7A02G135200, TRAESCS7A02G149000, TRAESCS7A02G146400, TRAESCS7A02G163000, TRAESCS7A02G175900, TRAESCS7A02G193100, TRAESCS7A02G236900, TRAESCS7A02G249900, TRAESCS7A02G250800, TRAESCS7A02G259900, TRAESCS7A02G260200, TRAESCS7A02G282600, TRAESCS7A02G287000, TRAESCS7A02G287400, TRAESCS7A02G299300, TRAESCS7A02G304700, TRAESCS7A02G3257500, TRAESCS7A02G425300, TRAESCS7A02G481100, TRAESCS7A02G504800, TRAESCS7A02G511900, TRAESCS7A02G516000, TRAESCS7A02G548500, TRAESCS7A02G548900, TRAESCS7B02G028500, TRAESCS7B02G068900, TRAESCS7B02G071400, TRAESCS7B02G098400, TRAESCS7B02G183300, TRAESCS7B02G197600, TRAESCS7B02G205000, TRAESCS7B02G211200, TRAESCS7B02G214800, TRAESCS7B02G219000, TRAESCS7B02G226600, TRAESCS7B02G226800, TRAESCS7B02G304900, TRAESCS7B02G308200, TRAESCS7B02G363200, TRAESCS7B02G395100, TRAESCS7B02G403200, TRAESCS7B02G432000, TRAESCS7B02G434000, TRAESCS7B02G434000, TRAESCS7B02G435200, TRAESCS7B02G455500, TRAESCS7B02G461400, TRAESCS7D02G032300, TRAESCS7D02G072200, TRAESCS7D02G093000, TRAESCS7D02G093800, TRAESCS7D02G127400, TRAESCS7D02G135000, TRAESCS7D02G167400, TRAESCS7D02G194400, TRAESCS7D02G199200, TRAESCS7D02G212800, TRAESCS7D02G232200, TRAESCS7D02G248600, TRAESCS7D02G282200, TRAESCS7D02G284900, TRAESCS7D02G287900, TRAESCS7D02G290200, TRAESCS7D02G307300, TRAESCS7D02G310800, TRAESCS7D02G313500, TRAESCS7D02G314300, TRAESCS7D02G344400, TRAESCS7D02G365200, TRAESCS7D02G399400, TRAESCS7D02G473600, TRAESCS7D02G487900, TRAESCS7D02G493200, TRAESCS7D02G500300, TRAESCS7D02G535200, TRAESCS7D02G506300, TRAESCS7D02G508900, TRAESCS7D02G510200, TRAESCS7D02G524700, TRAESCS7D02G525900, TRAESCS7D02G528800, TRAESCS7D02G535200, TRAESCSU02G000800, TRAESCSU02G028900, TRAESCSU02G032400, TRAESCSU02G068500, TRAESCSU02G103200, TRAESCSU02G110700, TRAESCSU02G133100, TRAESCSU02G172000
GO:MF	glycogen debranching enzyme activity	GO:0004133	4.87E-04	3.31	12	818	4	63217	TRAESCS2A02G159300, TRAESCS7A02G251400, TRAESCS7B02G139700, TRAESCS7D02G249500
GO:MF	ligase activity, forming carbon-oxygen bonds	GO:0016875	4.87E-04	3.31	234	818	13	63217	TRAESCS1B02G169100, TRAESCS1D02G198800, TRAESCS3A02G111500, TRAESCS3A02G163300, TRAESCS3A02G186200, TRAESCS3D02G169600, TRAESCS3D02G190200, TRAESCS4D02G243900, TRAESCS7A02G304700, TRAESCS7A02G516000, TRAESCS7B02G205000, TRAESCS7B02G432400, TRAESCS7D02G506300
GO:MF	aminoacyl-tRNA ligase activity	GO:0004812	4.87E-04	3.31	234	818	13	63217	TRAESCS1B02G169100, TRAESCS1D02G198800, TRAESCS3A02G111500, TRAESCS3A02G163300, TRAESCS3A02G186200, TRAESCS3D02G169600, TRAESCS3D02G190200, TRAESCS4D02G243900, TRAESCS7A02G304700, TRAESCS7A02G516000, TRAESCS7B02G205000, TRAESCS7B02G432400, TRAESCS7D02G506300
GO:MF	isoamylase activity	GO:0019156	1.29E-03	2.89	6	818	3	63217	TRAESCS7A02G251400, TRAESCS7B02G139700, TRAESCS7D02G249500
GO:MF	aminoacyl-tRNA editing activity	GO:0002161	1.29E-03	2.89	49	818	6	63217	TRAESCS3A02G163300, TRAESCS3D02G169600, TRAESCS6A02G330600, TRAESCS7A02G516000, TRAESCS7B02G432400, TRAESCS7D02G506300
GO:MF	linear malto-oligosaccharide phosphorylase activity	GO:0102250	1.98E-03	2.70	7	818	3	63217	TRAESCS5A02G395200, TRAESCS5B02G400000, TRAESCS5D02G404500
GO:MF	SHG alpha-glucan phosphorylase activity	GO:0102499	1.98E-03	2.70	7	818	3	63217	TRAESCS5A02G395200, TRAESCS5B02G400000, TRAESCS5D02G404500
GO:MF	ribonuclease III activity	GO:0004525	2.72E-03	2.57	37	818	5	63217	TRAESCS1A02G169000, TRAESCS1D02G158000, TRAESCS7B02G248100, TRAESCS7D02G344100, TRAESCS7D02G344200
GO:MF	double-stranded RNA-specific ribonuclease activity	GO:0032296	2.72E-03	2.57	37	818	5	63217	TRAESCS1A02G169000, TRAESCS1D02G158000, TRAESCS7B02G248100, TRAESCS7D02G344100, TRAESCS7D02G344200
GO:MF	isoleucine-tRNA ligase activity	GO:0004822	3.65E-03	2.44	9	818	3	63217	TRAESCS7A02G516000, TRAESCS7B02G432400, TRAESCS7D02G506300
GO:MF	limit dextrinase activity	GO:0010303	3.65E-03	2.44	2	818	2	63217	TRAESCS7A02G133500, TRAESCS7D02G133100
GO:MF	glycogen phosphorylase activity	GO:0008184	3.65E-03	2.44	9	818	3	63217	TRAESCS5A02G395200, TRAESCS5B02G400000, TRAESCS5D02G404500
GO:MF	alpha-mannosidase activity	GO:0004559	4.14E-03	2.38	23	818	4	63217	TRAESCS1A02G087500, TRAESCS1B02G106400, TRAESCS7A02G236900, TRAESCS7D02G199200



GO:MF	catalytic activity, acting on a tRNA	GO:0140101	8.15E-03	2.09	372	818	14	63217	TRAESC51B02G169100,TRAESC51D02G198800,TRAESC53A02G111500,TRAESC53A02G163300,TRAESC53A02G186200,TRAESC53D02G169600,TRAESC53D02G190200,TRAESC54D02G243900,TRAESC56A02G330600,TRAESC57A02G304700,TRAESC57A02G516000,TRAESC57B02G205000,TRAESC57B02G432400,TRAESC57D02G506300
GO:MF	mannosidase activity	GO:0015923	8.40E-03	2.08	28	818	4	63217	TRAESC51A02G087500,TRAESC51B02G106400,TRAESC57A02G236900,TRAESC57D02G199200
GO:MF	ATP phosphoribosyltransferase activity	GO:0003879	8.44E-03	2.07	3	818	2	63217	TRAESC54D02G327300,TRAESC55A02G501900
GO:MF	glucosyltransferase activity	GO:0046527	8.44E-03	2.07	379	818	14	63217	TRAESC51A02G091500,TRAESC51A02G339300,TRAESC51B02G119300,TRAESC51B02G351600,TRAESC51D02G100100,TRAESC51D02G341100,TRAESC53A02G292700,TRAESC53B02G327500,TRAESC56A02G154400,TRAESC57A02G070100,TRAESC57A02G120300,TRAESC57B02G018600,TRAESC57B02G048700,TRAESC57D02G117800
GO:MF	2-isopropylmalate synthase activity	GO:0003852	8.44E-03	2.07	3	818	2	63217	TRAESC55A02G131300,TRAESC55D02G139800
GO:MF	starch binding	GO:2001070	1.09E-02	1.96	31	818	4	63217	TRAESC51A02G091500,TRAESC51B02G119300,TRAESC51D02G100100,TRAESC54A02G149300
GO:MF	histone methyltransferase activity	GO:0042054	1.13E-02	1.95	114	818	7	63217	TRAESC51D02G366800,TRAESC52A02G382000,TRAESC52A02G389600,TRAESC55B02G177800,TRAESC55D02G051200,TRAESC56A02G226600,TRAESC57D02G127400
GO:MF	histone-lysine N-methyltransferase activity	GO:0018024	1.13E-02	1.95	114	818	7	63217	TRAESC51D02G366800,TRAESC52A02G382000,TRAESC52A02G389600,TRAESC55B02G177800,TRAESC55D02G051200,TRAESC56A02G226600,TRAESC57D02G127400
GO:MF	UDP-glycosyltransferase activity	GO:0008194	1.18E-02	1.93	491	818	16	63217	TRAESC51A02G091500,TRAESC51A02G339300,TRAESC51B02G119300,TRAESC51B02G351600,TRAESC51D02G100100,TRAESC51D02G341100,TRAESC53A02G292700,TRAESC53B02G327500,TRAESC56A02G154400,TRAESC57A02G070100,TRAESC57A02G120300,TRAESC57B02G018600,TRAESC57B02G048700,TRAESC57D02G117800
GO:MF	tyrosine-tRNA ligase activity	GO:0004831	1.26E-02	1.90	15	818	3	63217	TRAESC51B02G169100,TRAESC57A02G304700,TRAESC57B02G205000
GO:MF	SUMO activating enzyme activity	GO:0019948	1.38E-02	1.86	4	818	2	63217	TRAESC52A02G192200,TRAESC52B02G214100
GO:MF	jasmonate-amino synthetase activity	GO:0080123	2.16E-02	1.67	5	818	2	63217	TRAESC51A02G425100,TRAESC51D02G434100
GO:MF	starch synthase activity	GO:0009011	2.16E-02	1.67	5	818	2	63217	TRAESC57A02G120300,TRAESC57D02G117800
GO:MF	1,3-beta-D-glucan synthase activity	GO:0003843	2.22E-02	1.65	40	818	4	63217	TRAESC5A02G292700,TRAESC53B02G327500,TRAESC56A02G154400,TRAESC57B02G048700
GO:MF	protein-lysine N-methyltransferase activity	GO:0016279	2.22E-02	1.65	133	818	7	63217	TRAESC51D02G366800,TRAESC52A02G382000,TRAESC52A02G389600,TRAESC55B02G177800,TRAESC55D02G051200,TRAESC56A02G226600,TRAESC57D02G127400
GO:MF	lysine N-methyltransferase activity	GO:0016278	2.22E-02	1.65	133	818	7	63217	TRAESC51D02G366800,TRAESC52A02G382000,TRAESC52A02G389600,TRAESC55B02G177800,TRAESC55D02G051200,TRAESC56A02G226600,TRAESC57D02G127400
GO:MF	7S RNA binding	GO:0008312	2.31E-02	1.64	41	818	4	63217	TRAESC51D02G329300,TRAESC53A02G402600,TRAESC53B02G436000,TRAESC53D02G397600
GO:MF	tRNA binding	GO:0000049	2.33E-02	1.63	70	818	5	63217	TRAESC53B02G376000,TRAESC54D02G243900,TRAESC57A02G516000,TRAESC57B02G432400,TRAESC57D02G506300
GO:MF	3-dehydroquinate dehydratase activity	GO:0003855	2.33E-02	1.63	6	818	2	63217	TRAESC55B02G102400,TRAESC55D02G109200
GO:MF	guanylate kinase activity	GO:0004385	2.33E-02	1.63	6	818	2	63217	TRAESC55A02G100300,TRAESC55D02G112800
GO:MF	1,4-alpha-oligoglucan phosphorylase activity	GO:0004645	2.33E-02	1.63	21	818	3	63217	TRAESC55A02G395200,TRAESC55B02G400000,TRAESC55D02G404500
GO:MF	ubiquitin-protein transferase regulator activity	GO:0055106	2.33E-02	1.63	21	818	3	63217	TRAESC53A02G527600,TRAESC53B02G595600,TRAESC56B02G296200
GO:MF	shikimate 3-dehydrogenase (NADP+) activity	GO:0004764	2.33E-02	1.63	6	818	2	63217	TRAESC55B02G102400,TRAESC55D02G109200
GO:MF	carboxylic ester hydrolase activity	GO:0052689	2.33E-02	1.63	305	818	11	63217	TRAESC52A02G135400,TRAESC52A02G135500,TRAESC52B02G159300,TRAESC52B02G159400,TRAESC53A02G163300,TRAESC53D02G169600,TRAESC56A02G330600,TRAESC56B02G375900,TRAESC57A02G516000,TRAESC57B02G432400,TRAESC57D02G506300
GO:MF	thiol oxidase activity	GO:0016972	2.33E-02	1.63	6	818	2	63217	TRAESC51B02G399000,TRAESC51D02G385200
GO:MF	flavin-linked sulfhydryl oxidase activity	GO:0016971	2.33E-02	1.63	6	818	2	63217	TRAESC51B02G399000,TRAESC51D02G385200
GO:MF	ubiquitin-protein transferase activator activity	GO:0097027	2.33E-02	1.63	21	818	3	63217	TRAESC53A02G527600,TRAESC53B02G595600,TRAESC56B02G296200
GO:MF	MAP kinase activity	GO:0004707	2.79E-02	1.55	46	818	4	63217	TRAESC51A02G402400,TRAESC51B02G431400,TRAESC51D02G410100,TRAESC56A02G099600



										50, TRAESCS7A02G251400, TRAESCS7A02G287400, TRAESCS7A02G549200, TRAESCS7A02G549300, TRAESCS7B02G048700, TRAESCS7B02G183300, TRAESCS7B02G472400, TRAESCS7D02G17800, TRAESCS7D02G133100, TRAESCS7D02G284900, TRAESCS7D02G535500, TRAESCS7D02G535600
GO:BP	amylopectin metabolic process	GO:2000896	1.32E-06	5.88	8	547	5	42199	TRAESCS7A02G120300, TRAESCS7A02G251400, TRAESCS7B02G139700, TRAESCS7D02G117800, TRAESCS7D02G249500	
GO:BP	amylopectin biosynthetic process	GO:0010021	1.32E-06	5.88	8	547	5	42199	TRAESCS7A02G120300, TRAESCS7A02G251400, TRAESCS7B02G139700, TRAESCS7D02G117800, TRAESCS7D02G249500 TRAESCS1A02G087500, TRAESCS1A02G339300, TRAESCS1A02G419600, TRAESCS1B02G094700, TRAESCS1B02G106400, TRAESCS1B02G343500, TRAESCS1B02G351600, TRAESCS1B02G449700, TRAESCS1D02G041200, TRAESCS1D02G332200, TRAESCS1D02G341100, TRAESCS1D02G427400, TRAESCS2A02G135400, TRAESCS2A02G159300, TRAESCS2A02G161000, TRAESCS2A02G163100, TRAESCS2A02G160900, TRAESCS2B02G159400, TRAESCS2B02G187000, TRAESCS2B02G327300, TRAESCS2D02G386400, TRAESCS2D02G402200, TRAESCS2D02G495900, TRAESCS3A02G292700, TRAESCS3A02G365800, TRAESCS3A02G489900, TRAESCS3B02G148500, TRAESCS3B02G227500, TRAESCS3B02G253400, TRAESCS3B02G259500, TRAESCS3D02G141900, TRAESCS3D02G216700, TRAESCS3D02G388900, TRAESCS3D02G475600, TRAESCS4A02G149300, TRAESCS4A02G263900, TRAESCS4A02G485600, TRAESCS4A02G101000, TRAESCS4A02G39300, TRAESCS4B02G202200, TRAESCS4B02G400000, TRAESCS4B02G404500, TRAESCS4A02G154400, TRAESCS4A02G2064100, TRAESCS4A02G070100, TRAESCS4A02G120300, TRAESCS4A02G133500, TRAESCS4A02G236900, TRAESCS4A02G251400, TRAESCS4A02G287400, TRAESCS4A02G549200, TRAESCS4A02G549300, TRAESCS4B02G034600, TRAESCS4B02G039700, TRAESCS4B02G183300, TRAESCS4B02G192200, TRAESCS4B02G472400, TRAESCS4D02G054800, TRAESCS4D02G117800, TRAESCS4D02G133100, TRAESCS4D02G199200, TRAESCS4D02G249500, TRAESCS4D02G284900, TRAESCS4D02G535500, TRAESCS4D02G535600	
GO:BP	carbohydrate metabolic process	GO:0005975	1.34E-05	4.87	2619	547	66	42199	TRAESCS1A02G169900, TRAESCS1A02G333000, TRAESCS1A02G445500, TRAESCS2A02G252400, TRAESCS3D02G231600, TRAESCS3D02G366200, TRAESCS4A02G197000, TRAESCS4A02G281500, TRAESCS4A02G495800, TRAESCS4B02G031500, TRAESCS4D02G029200, TRAESCS6A02G223100, TRAESCS6D02G12100, TRAESCS7A02G002800, TRAESCS7D02G003000	
GO:BP	gene silencing	GO:0016458	1.85E-05	4.73	225	547	15	42199	TRAESCS1A02G169900, TRAESCS1A02G333000, TRAESCS1A02G445500, TRAESCS2A02G252400, TRAESCS3D02G231600, TRAESCS3D02G366200, TRAESCS4A02G197000, TRAESCS4A02G281500, TRAESCS4A02G495800, TRAESCS4B02G031500, TRAESCS4D02G029200, TRAESCS6A02G223100, TRAESCS6D02G12100, TRAESCS7A02G002800, TRAESCS7D02G003000	
GO:BP	gene silencing by RNA	GO:0031047	2.78E-05	4.56	173	547	13	42199	TRAESCS1A02G169900, TRAESCS1A02G333000, TRAESCS1A02G445500, TRAESCS2A02G252400, TRAESCS3D02G231600, TRAESCS3D02G366200, TRAESCS4A02G197000, TRAESCS4A02G281500, TRAESCS4A02G495800, TRAESCS4B02G031500, TRAESCS4D02G029200, TRAESCS6A02G223100, TRAESCS6D02G12100, TRAESCS7A02G002800, TRAESCS7D02G003000 TRAESCS1A02G169900, TRAESCS1A02G333000, TRAESCS1A02G445500, TRAESCS2A02G252400, TRAESCS3D02G231600, TRAESCS3D02G366200, TRAESCS4A02G197000, TRAESCS4A02G281500, TRAESCS4A02G495800, TRAESCS4B02G031500, TRAESCS4D02G029200, TRAESCS6A02G223100, TRAESCS6D02G12100, TRAESCS7A02G002800, TRAESCS7D02G003000 TRAESCS1B02G085500, TRAESCS1B02G118200, TRAESCS1B02G123200, TRAESCS1B02G129400, TRAESCS1B02G173200, TRAESCS1B02G2271800, TRAESCS1B02G343500, TRAESCS1B02G399000, TRAESCS1B02G433100, TRAESCS1B02G445400, TRAESCS1B02G464000, TRAESCS1D02G044200, TRAESCS1D02G047800, TRAESCS1D02G142900, TRAESCS1D02G303900, TRAESCS1D02G32200, TRAESCS1D02G388200, TRAESCS1D02G410100, TRAESCS1D02G411400, TRAESCS1D02G434100, TRAESCS2A02G044800, TRAESCS2A02G104500, TRAESCS2A02G189700, TRAESCS2A02G252400, TRAESCS2A02G292700, TRAESCS2A02G320700, TRAESCS2A02G320700, TRAESCS2A02G32100, TRAESCS2B02G039100, TRAESCS2B02G035400, TRAESCS2B02G035800, TRAESCS2D02G074300, TRAESCS2D02G090300, TRAESCS2D02G135200, TRAESCS2D02G198100, TRAESCS2D02G295300, TRAESCS2D02G426700, TRAESCS2D02G452400, TRAESCS2D02G503900, TRAESCS3A02G008000, TRAESCS3A02G080700, TRAESCS3A02G154400, TRAESCS3A02G325800, TRAESCS3A02G429100, TRAESCS3A02G489900, TRAESCS3A02G527600, TRAESCS3B02G021200, TRAESCS3B02G141100, TRAESCS3B02G181800, TRAESCS3B02G204200, TRAESCS3B02G278000, TRAESCS3B02G330800, TRAESCS3B02G354900, TRAESCS3B02G412300, TRAESCS3B02G413800, TRAESCS3B02G415800, TRAESCS3B02G466800, TRAESCS3B02G510500, TRAESCS3B02G525400, TRAESCS3B02G594200, TRAESCS3B02G595600, TRAESCS3D02G048000, TRAESCS3D02G123900, TRAESCS3D02G144600, TRAESCS3D02G169600, TRAESCS3D02G179800, TRAESCS3D02G231600, TRAESCS3D02G366200, TRAESCS3D02G424600, TRAESCS3D02G475600, TRAESCS3D02G532000, TRAESCS4A02G197000, TRAESCS4A02G281500, TRAESCS4A02G495800, TRAESCS4B02G006500, TRAESCS4B02G031500, TRAESCS4B02G178900, TRAESCS4B02G329700, TRAESCS4B02G268800, TRAESCS4D02G011300, TRAESCS4D02G029200, TRAESCS4D02G246100, TRAESCS4D02G316900, TRAESCS4D02G0294100, TRAESCS5A02G155900, TRAESCS5A02G174200, TRAESCS5A02G295700, TRAESCS5B02G054200, TRAESCS5B02G154100, TRAESCS5B02G185200, TRAESCS5B02G196800, TRAESCS5B02G384600, TRAESCS5B02G405700, TRAESCS5D02G161000, TRAESCS5D02G178800, TRAESCS5D02G192100, TRAESCS5D02G411100, TRAESCS5D02G495500, TRAESCS5D02G66900, TRAESCS6A02G017100, TRAESCS6A02G017600, TRAESCS6A02G017700, TRAESCS6A02G117100, TRAESCS6A02G223100, TRAESCS6A02G227900, TRAESCS6A02G23300, TRAESCS6A02G418900, TRAESCS6B02G02253900, TRAESCS6B02G256000, TRAESCS6B02G296200, TRAESCS6B02G332100, TRAESCS6B02G408800, TRAESCS6D02G019800, TRAESCS6D02G021700, TRAESCS6D02G023500, TRAESCS6D02G036200, TRAESCS6D02G218400, TRAESCS6D02G221200, TRAESCS6D02G258000, TRAESCS7A02G002800, TRAESCS7A02G066600, TRAESCS7A02G184400, TRAESCS7A02G258000, TRAESCS7A02G357500, TRAESCS7A02G497800, TRAESCS7B02G308200, TRAESCS7B02G308200, TRAESCS7B02G363200, TRAESCS7B02G402000, TRAESCS7B02G461400, TRAESCS7D02G003000, TRAESCS7D02G012700, TRAESCS7D02G232200, TRAESCS7D02G248600, TRAESCS7D02G265600, TRAESCS7D02G282200, TRAESCS7D02G365200, TRAESCS7D02G525900, TRAESCS7D02G241200	
GO:BP	biological regulation	GO:0065007	9.82E-05	4.01	8183	547	151	42199	U02G241200	
GO:BP	branched-chain amino acid biosynthetic process	GO:0009082	1.24E-04	3.91	47	547	7	42199	TRAESCS1B02G222500, TRAESCS1D02G211700, TRAESCS1D02G419800, TRAESCS2B02G069800, TRAESCS5A02G131300, TRAESCS5B02G403400, TRAESCS5D02G139800	
GO:BP	energy derivation by oxidation of organic compounds	GO:0015980	1.98E-04	3.70	312	547	16	42199	TRAESCS1A02G419600, TRAESCS1B02G449700, TRAESCS1D02G427400, TRAESCS2A02G310300, TRAESCS2B02G327300, TRAESCS2D02G295200, TRAESCS3D02G420100, TRAESCS4A02G251400, TRAESCS7A02G287400, TRAESCS7A02G549200, TRAESCS7A02G549300, TRAESCS7B02G183300, TRAESCS7B02G472400, TRAESCS7D02G284900, TRAESCS7D02G535500, TRAESCS7D02G535600	
GO:BP	branched-chain amino acid metabolic process	GO:0009081	3.36E-04	3.47	55	547	7	42199	TRAESCS1B02G222500, TRAESCS1D02G211700, TRAESCS1D02G419800, TRAESCS2B02G069800, TRAESCS5A02G131300, TRAESCS5B02G403400, TRAESCS5D02G139800	
GO:BP	tRNA aminoacylation for protein translation	GO:0006418	5.97E-04	3.22	233	547	13	42199	TRAESCS1B02G169100, TRAESCS1D02G198800, TRAESCS3A02G111500, TRAESCS3A02G163300, TRAESCS3A02G186200, TRAESCS3D02G169600, TRAESCS3D02G190200, TRAESCS4D02G243900, TRAESCS4D02G304700, TRAESCS7A02G2516000, TRAESCS7B02G205000, TRAESCS7B02G432400, TRAESCS7D02G506300 TRAESCS1A02G043600, TRAESCS1A02G169900, TRAESCS1A02G190100, TRAESCS1A02G304400, TRAESCS1A02G333000, TRAESCS1A02G403500, TRAESCS1A02G425100, TRAESCS1A02G445500, TRAESCS1B02G085500, TRAESCS1B02G118200, TRAESCS1B02G123200, TRAESCS1B02G129400, TRAESCS1B02G173200, TRAESCS1B02G2271800, TRAESCS1B02G343500, TRAESCS1B02G399000, TRAESCS1B02G433100, TRAESCS1B02G445400, TRAESCS1B02G464000, TRAESCS1D02G044200, TRAESCS1D02G047800, TRAESCS1D02G142900, TRAESCS1D02G303900, TRAESCS1D02G32200, TRAESCS1D02G388200, TRAESCS1D02G410100, TRAESCS1D02G411400, TRAESCS1D02G434100, TRAESCS2A02G044800, TRAESCS2A02G104500, TRAESCS2A02G189700, TRAESCS2A02G252400, TRAESCS2A02G292700, TRAESCS2A02G320700, TRAESCS2A02G320700, TRAESCS2A02G32100, TRAESCS2B02G039100, TRAESCS2B02G035400, TRAESCS2B02G035800, TRAESCS2D02G074300, TRAESCS2D02G090300, TRAESCS2D02G135200, TRAESCS2D02G198100, TRAESCS2D02G295300, TRAESCS2D02G426700, TRAESCS2D02G452400, TRAESCS3A02G008000, TRAESCS3A02G080700, TRAESCS3A02G154400, TRAESCS3A02G325800, TRAESCS3A02G429100, TRAESCS3A02G527600, TRAESCS3B02G141100, TRAESCS3B02G181800, TRAESCS3B02G204200, TRAESCS3B02G278000, TRAESCS3B02G330800, TRAESCS3B02G354900, TRAESCS3B02G412300, TRAESCS3B02G413800, TRAESCS3B02G415800, TRAESCS3B02G466800, TRAESCS3B02G510500, TRAESCS3B02G525400, TRAESCS3B02G594200, TRAESCS3B02G595600, TRAESCS3D02G123900, TRAESCS3D02G144600, TRAESCS3D02G169600, TRAESCS3D02G179800, TRAESCS3D02G231600, TRAESCS3D02G366200, TRAESCS3D02G424600, TRAESCS3D02G475600, TRAESCS4A02G197000, TRAESCS4A02G281500, TRAESCS4A02G495800, TRAESCS4B02G006500, TRAESCS4B02G031500, TRAESCS4B02G178900, TRAESCS4B02G329700, TRAESCS4B02G268800, TRAESCS4D02G011300, TRAESCS4D02G029200, TRAESCS4D02G246100, TRAESCS4D02G316900, TRAESCS4D02G304700, TRAESCS5A02G155900, TRAESCS5A02G174200, TRAESCS5A02G295700, TRAESCS5B02G054200, TRAESCS5B02G154100, TRAESCS5B02G185200, TRAESCS5B02G196800, TRAESCS5B02G384600, TRAESCS5D02G161000, TRAESCS5D02G178800, TRAESCS5D02G192100, TRAESCS5D02G411100, TRAESCS5D02G495500, TRAESCS5D02G66900, TRAESCS6A02G017100, TRAESCS6A02G017600, TRAESCS6A02G017700, TRAESCS6A02G223100, TRAESCS6A02G227900, TRAESCS6A02G23300, TRAESCS6A02G418900, TRAESCS6B02G02253900, TRAESCS6B02G253900, TRAESCS6B02G295200, TRAESCS6B02G352100, TRAESCS6B02G408800, TRAESCS6D02G019800, TRAESCS6D02G021700, TRAESCS6D02G023500, TRAESCS6D02G036200, TRAESCS6D02G212100, TRAESCS6D02G258000, TRAESCS7A02G002800, TRAESCS7A02G066600, TRAESCS7A02G184400, TRAESCS7A02G258000, TRAESCS7B02G308200, TRAESCS7B02G308200, TRAESCS7B02G363200, TRAESCS7D02G003000, TRAESCS7D02G012700, TRAESCS7D02G232200, TRAESCS7D02G248600, TRAESCS7D02G265600, TRAESCS7D02G282200, TRAESCS7D02G365200, TRAESCS7D02G525900, TRAESCS7D02G241200	
GO:BP	regulation of biological process	GO:0050789	6.46E-04	3.19	7606	547	138	42199	TRAESCS1B02G169100, TRAESCS1D02G198800, TRAESCS3A02G111500, TRAESCS3A02G163300, TRAESCS3A02G186200, TRAESCS3D02G169600, TRAESCS3D02G190200, TRAESCS4D02G243900, TRAESCS4D02G304700, TRAESCS7A02G2516000, TRAESCS7B02G205000, TRAESCS7B02G432400, TRAESCS7D02G506300	
GO:BP	amino acid activation	GO:0043038	8.67E-04	3.06	244	547	13	42199	TRAESCS1B02G169100, TRAESCS1D02G198800, TRAESCS3A02G111500, TRAESCS3A02G163300, TRAESCS3A02G186200, TRAESCS3D02G169600, TRAESCS3D02G190200, TRAESCS4D02G243900, TRAESCS4D02G304700, TRAESCS7A02G2516000, TRAESCS7B02G205000, TRAESCS7B02G432400, TRAESCS7D02G506300	
GO:BP	tRNA aminoacylation	GO:0043039	8.67E-04	3.06	244	547	13	42199	TRAESCS1B02G169100, TRAESCS1D02G198800, TRAESCS3A02G111500, TRAESCS3A02G163300, TRAESCS3A02G186200, TRAESCS3D02G169600, TRAESCS3D02G190200, TRAESCS4D02G243900, TRAESCS4D02G304700, TRAESCS7A02G2516000, TRAESCS7B02G205000, TRAESCS7B02G432400, TRAESCS7D02G506300	



GO:BP	carbohydrate homeostasis	GO:0033500	1.28E-02	1.89	30	547	4	42199	TRAESCS2D02G155200,TRAESCS3A02G480900,TRAESCS3B02G252400,TRAESCS3D02G475600
GO:BP	positive regulation of transferase activity	GO:0051347	1.28E-02	1.89	52	547	5	42199	TRAESCS3A02G527600,TRAESCS3B02G595600,TRAESCS6B02G296200,TRAESCS7A02G066300,TRAESCS7D02G282200
GO:BP	glucose homeostasis	GO:0042593	1.28E-02	1.89	30	547	4	42199	TRAESCS2D02G155200,TRAESCS3A02G480900,TRAESCS3B02G252400,TRAESCS3D02G475600
GO:BP	regulation of protein modification by small protein conjugation or removal	GO:1903320	1.37E-02	1.86	31	547	4	42199	TRAESCS3A02G527600,TRAESCS3B02G595600,TRAESCS6B02G296200,TRAESCS7D02G282200
GO:BP	regulation of protein ubiquitination	GO:0031396	1.37E-02	1.86	31	547	4	42199	TRAESCS3A02G527600,TRAESCS3B02G595600,TRAESCS6B02G296200,TRAESCS7D02G282200
GO:BP	negative regulation of metabolic process	GO:0009892	1.37E-02	1.86	885	547	24	42199	TRAESCS1A02G160900,TRAESCS1A02G333000,TRAESCS1A02G445500,TRAESCS2A02G252400,TRAESCS2B02G358800,TRAESCS2D02G090300,TRAESCS2D02G155200,TRAESCS2D02G295300,TRAESCS3A02G154400,TRAESCS3B02G180800,TRAESCS3D02G231600,TRAESCS3D02G366200,TRAESCS4A02G2281500,TRAESCS4A02G495800,TRAESCS4B02G031500,TRAESCS4D02G029200,TRAESCS6A02G223100,TRAESCS6A02G227900,TRAESCS6B02G253900,TRAESCS6D02G212100,TRAESCS6D02G258000,TRAESCS7A02G002800,TRAESCS7D02G003300
GO:BP	positive regulation of protein modification process	GO:0031401	1.55E-02	1.81	56	547	5	42199	TRAESCS3A02G527600,TRAESCS3B02G595600,TRAESCS6B02G296200,TRAESCS7A02G066300,TRAESCS7D02G282200
GO:BP	macromolecule metabolic process	GO:0043170	1.55E-02	1.81	20037	547	297	42199	TRAESCS1A02G043600,TRAESCS1A02G114900,TRAESCS1A02G137000,TRAESCS1A02G160900,TRAESCS1A02G190100,TRAESCS1A02G232300,TRAESCS1A02G304400,TRAESCS1A02G333000,TRAESCS1A02G402400,TRAESCS1A02G403300,TRAESCS1A02G419600,TRAESCS1A02G445500,TRAESCS1B02G085500,TRAESCS1B02G129400,TRAESCS1B02G134800,TRAESCS1B02G15400,TRAESCS1B02G155600,TRAESCS1B02G169100,TRAESCS1B02G200200,TRAESCS1B02G256600,TRAESCS1B02G294900,TRAESCS1B02G298100,TRAESCS1B02G343500,TRAESCS1B02G431400,TRAESCS1B02G433100,TRAESCS1B02G449700,TRAESCS1B02G464000,TRAESCS1D02G044200,TRAESCS1D02G132500,TRAESCS1D02G136400,TRAESCS1D02G142900,TRAESCS1D02G158000,TRAESCS1D02G198800,TRAESCS1D02G235500,TRAESCS1D02G303900,TRAESCS1D02G332200,TRAESCS1D02G366800,TRAESCS1D02G383400,TRAESCS1D02G401100,TRAESCS1D02G411400,TRAESCS1D02G427400,TRAESCS2A02G043100,TRAESCS2A02G044800,TRAESCS2A02G044900,TRAESCS2A02G044950,TRAESCS2A02G135400,TRAESCS2A02G159300,TRAESCS2A02G189700,TRAESCS2A02G192200,TRAESCS2A02G248200,TRAESCS2A02G252400,TRAESCS2A02G266700,TRAESCS2A02G297900,TRAESCS2A02G310300,TRAESCS2A02G338000,TRAESCS2A02G352100,TRAESCS2A02G382000,TRAESCS2A02G386700,TRAESCS2A02G389600,TRAESCS2A02G400900,TRAESCS2A02G507300,TRAESCS2B02G159400,TRAESCS2B02G175800,TRAESCS2B02G213800,TRAESCS2B02G214100,TRAESCS2B02G249900,TRAESCS2B02G269200,TRAESCS2B02G309100,TRAESCS2B02G327300,TRAESCS2B02G335400,TRAESCS2B02G358800,TRAESCS2D02G012000,TRAESCS2D02G022900,TRAESCS2D02G074300,TRAESCS2D02G090300,TRAESCS2D02G155200,TRAESCS2D02G198100,TRAESCS2D02G238800,TRAESCS2D02G295300,TRAESCS2D02G386400,TRAESCS2D02G426700,TRAESCS2D02G503900,TRAESCS3A02G080700,TRAESCS3A02G111500,TRAESCS3A02G154400,TRAESCS3A02G159900,TRAESCS3A02G163300,TRAESCS3A02G179700,TRAESCS3A02G186200,TRAESCS3A02G216100,TRAESCS3A02G222300,TRAESCS3A02G292700,TRAESCS3A02G325800,TRAESCS3A02G331500,TRAESCS3A02G402600,TRAESCS3A02G414700,TRAESCS3A02G429100,TRAESCS3A02G527600,TRAESCS3B02G018100,TRAESCS3B02G141100,TRAESCS3B02G180800,TRAESCS3B02G204200,TRAESCS3B02G252500,TRAESCS3B02G278000,TRAESCS3B02G327500,TRAESCS3B02G330800,TRAESCS3B02G354900,TRAESCS3B02G412300,TRAESCS3B02G413800,TRAESCS3B02G415800,TRAESCS3B02G4166800,TRAESCS3B02G510500,TRAESCS3B02G544700,TRAESCS3B02G595600,TRAESCS3D02G018000,TRAESCS3D02G123900,TRAESCS3D02G141900,TRAESCS3D02G164700,TRAESCS3D02G167100,TRAESCS3D02G169600,TRAESCS3D02G179800,TRAESCS3D02G190200,TRAESCS3D02G230100,TRAESCS3D02G231600,TRAESCS3D02G249100,TRAESCS3D02G305500,TRAESCS3D02G319200,TRAESCS3D02G334500,TRAESCS3D02G366200,TRAESCS3D02G376500,TRAESCS3D02G44600,TRAESCS3D02G466500,TRAESCS4A02G13700,TRAESCS4A02G149500,TRAESCS4A02G197000,TRAESCS4A02G256800,TRAESCS4A02G281500,TRAESCS4A02G322300,TRAESCS4A02G323800,TRAESCS4A02G329100,TRAESCS4B02G031500,TRAESCS4B02G090000,TRAESCS4B02G137100,TRAESCS4B02G138800,TRAESCS4B02G151800,TRAESCS4B02G157300,TRAESCS4B02G239700,TRAESCS4B02G268800,TRAESCS4B02G281900,TRAESCS4D02G011100,TRAESCS4D02G029200,TRAESCS4D02G044700,TRAESCS4D02G074800,TRAESCS4D02G082500,TRAESCS4D02G086000,TRAESCS4D02G171500,TRAESCS4D02G185200,TRAESCS4D02G199800,TRAESCS4D02G243900,TRAESCS4D02G289100,TRAESCS4D02G316900,TRAESCS4D02G318400,TRAESCS4D02G357800,TRAESCS5A02G049100,TRAESCS5A02G060300,TRAESCS5A02G155900,TRAESCS5A02G174200,TRAESCS5A02G200800,TRAESCS5A02G236300,TRAESCS5A02G246000,TRAESCS5A02G295700,TRAESCS5A02G321500,TRAESCS5A02G347900,TRAESCS5B02G040200,TRAESCS5B02G062900,TRAESCS5B02G154100,TRAESCS5B02G196800,TRAESCS5B02G199400,TRAESCS5B02G297300,TRAESCS5B02G299900,TRAESCS5B02G384600,TRAESCS5B02G405700,TRAESCS5B02G485300,TRAESCS5D02G071900,TRAESCS5D02G161000,TRAESCS5D02G178800,TRAESCS5D02G238500,TRAESCS5D02G282000,TRAESCS5D02G305200,TRAESCS5D02G311100,TRAESCS5D02G495500,TRAESCS6A02G017600,TRAESCS6A02G017700,TRAESCS6A02G099600,TRAESCS6A02G117100,TRAESCS6A02G119000,TRAESCS6A02G154400,TRAESCS6A02G200300,TRAESCS6A02G223100,TRAESCS6A02G226600,TRAESCS6A02G227900,TRAESCS6A02G235300,TRAESCS6A02G254900,TRAESCS6A02G26418900,TRAESCS6A02G26996400,TRAESCS6B02G02201100,TRAESCS6B02G253900,TRAESCS6B02G256600,TRAESCS6B02G270600,TRAESCS6B02G279300,TRAESCS6B02G291800,TRAESCS6B02G296200,TRAESCS6B02G305400,TRAESCS6B02G31900,TRAESCS6B02G410400,TRAESCS6B02G472000,TRAESCS6D02G019800,TRAESCS6D02G123500,TRAESCS6D02G136300,TRAESCS6D02G182900,TRAESCS6D02G164900,TRAESCS6D02G168400,TRAESCS6D02G189600,TRAESCS6D02G210400,TRAESCS6D02G212100,TRAESCS6D02G236200,TRAESCS6D02G236400,TRAESCS6D02G244300,TRAESCS6D02G258000,TRAESCS6D02G288200,TRAESCS6D02G362000,TRAESCS7A02G002800,TRAESCS7A02G065700,TRAESCS7A02G066300,TRAESCS7A02G070100,TRAESCS7A02G076500,TRAESCS7A02G105100,TRAESCS7A02G120300,TRAESCS7A02G133500,TRAESCS7A02G250800,TRAESCS7A02G250900,TRAESCS7A02G251400,TRAESCS7A02G287400,TRAESCS7A02G299300,TRAESCS7A02G304700,TRAESCS7A02G357500,TRAESCS7A02G502400,TRAESCS7A02G504800,TRAESCS7A02G511900,TRAESCS7A02G516000,TRAESCS7A02G549200,TRAESCS7A02G549300,TRAESCS7B02G0222900,TRAESCS7B02G048700,TRAESCS7B02G094000,TRAESCS7B02G131700,TRAESCS7B02G139700,TRAESCS7B02G183300,TRAESCS7B02G197600,TRAESCS7B02G205000,TRAESCS7B02G214800,TRAESCS7B02G248100,TRAESCS7B02G260600,TRAESCS7B02G308200,TRAESCS7B02G320200,TRAESCS7B02G363200,TRAESCS7B02G403200,TRAESCS7B02G423200,TRAESCS7B02G432400,TRAESCS7B02G455500,TRAESCS7B02G472400,TRAESCS7D02G003300,TRAESCS7D02G072200,TRAESCS7D02G117800,TRAESCS7D02G133100,TRAESCS7D02G134100,TRAESCS7D02G23200,TRAESCS7D02G48600,TRAESCS7D02G49500,TRAESCS7D02G265600,TRAESCS7D02G282200,TRAESCS7D02G284900,TRAESCS7D02G310800,TRAESCS7D02G324100,TRAESCS7D02G34100,TRAESCS7D02G344200,TRAESCS7D02G365200,TRAESCS7D02G365200,TRAESCS7D02G50900,TRAESCS7D02G50900,TRAESCS7D02G50900,TRAESCS7D02G50900,TRAESCS7D02G524700,TRAESCS7D02G535500,TRAESCS7D02G535600,TRAESCS8U02G068900,TRAESCSU02G110700,TRAESCSU02G138100,TRAESCSU02G207000
GO:BP	trehalose biosynthetic process	GO:0005992	1.55E-02	1.81	56	547	5	42199	TRAESCS1A02G339300,TRAESCS1B02G351600,TRAESCS1D02G341100,TRAESCS2A02G161000,TRAESCS2B02G187000
GO:BP	RNA metabolic process	GO:0016070	1.66E-02	1.78	6571	547	113	42199	TRAESCS1A02G043600,TRAESCS1A02G137000,TRAESCS1A02G160900,TRAESCS1A02G190100,TRAESCS1A02G304400,TRAESCS1A02G333000,TRAESCS1A02G403500,TRAESCS1B02G085500,TRAESCS1B02G129400,TRAESCS1B02G155600,TRAESCS1B02G169100,TRAESCS1B02G343500,TRAESCS1B02G433100,TRAESCS1B02G464000,TRAESCS1D02G044200,TRAESCS1D02G132500,TRAESCS1D02G142900,TRAESCS1D02G158000,TRAESCS1D02G198800,TRAESCS1D02G303900,TRAESCS1D02G332200,TRAESCS1D02G383400,TRAESCS1D02G411400,TRAESCS2A02G044800,TRAESCS2A02G189700,TRAESCS2A02G292700,TRAESCS2A02G297400,TRAESCS2A02G352100,TRAESCS2B02G309100,TRAESCS2B02G335400,TRAESCS2D02G074000,TRAESCS2D02G198100,TRAESCS2D02G295300,TRAESCS2D02G503900,TRAESCS3A02G080700,TRAESCS3A02G111500,TRAESCS3A02G154400,TRAESCS3A02G159900,TRAESCS3A02G163300,TRAESCS3A02G179700,TRAESCS3A02G186200,TRAESCS3A02G216100,TRAESCS3A02G222300,TRAESCS3A02G292700,TRAESCS3A02G325800,TRAESCS3A02G331500,TRAESCS3A02G402600,TRAESCS3A02G414700,TRAESCS3A02G429100,TRAESCS3A02G527600,TRAESCS3B02G141100,TRAESCS3B02G180800,TRAESCS3B02G204200,TRAESCS3B02G252500,TRAESCS3B02G278000,TRAESCS3B02G327500,TRAESCS3B02G330800,TRAESCS3B02G354900,TRAESCS3B02G412300,TRAESCS3B02G413800,TRAESCS3B02G415800,TRAESCS3B02G4166800,TRAESCS3B02G510500,TRAESCS3B02G544700,TRAESCS3B02G595600,TRAESCS3D02G018000,TRAESCS3D02G123900,TRAESCS3D02G141900,TRAESCS3D02G164700,TRAESCS3D02G167100,TRAESCS3D02G169600,TRAESCS3D02G179800,TRAESCS3D02G190200,TRAESCS3D02G230100,TRAESCS3D02G231600,TRAESCS3D02G249100,TRAESCS3D02G305500,TRAESCS3D02G319200,TRAESCS3D02G334500,TRAESCS3D02G366200,TRAESCS3D02G376500,TRAESCS3D02G44600,TRAESCS3D02G466500,TRAESCS4A02G13700,TRAESCS4A02G149500,TRAESCS4A02G197000,TRAESCS4A02G256800,TRAESCS4A02G281500,TRAESCS4A02G322300,TRAESCS4A02G323800,TRAESCS4A02G329100,TRAESCS4B02G031500,TRAESCS4B02G090000,TRAESCS4B02G137100,TRAESCS4B02G138800,TRAESCS4B02G151800,TRAESCS4B02G157300,TRAESCS4B02G239700,TRAESCS4B02G268800,TRAESCS4B02G281900,TRAESCS4D02G011100,TRAESCS4D02G029200,TRAESCS4D02G044700,TRAESCS4D02G074800,TRAESCS4D02G082500,TRAESCS4D02G086000,TRAESCS4D02G171500,TRAESCS4D02G185200,TRAESCS4D02G199800,TRAESCS4D02G243900,TRAESCS4D02G289100,TRAESCS4D02G316900,TRAESCS4D02G318400,TRAESCS4D02G357800,TRAESCS5A02G049100,TRAESCS5A02G060300,TRAESCS5A02G155900,TRAESCS5A02G174200,TRAESCS5A02G200800,TRAESCS5A02G236300,TRAESCS5A02G246000,TRAESCS5A02G295700,TRAESCS5A02G321500,TRAESCS5A02G347900,TRAESCS5B02G040200,TRAESCS5B02G062900,TRAESCS5B02G154100,TRAESCS5B02G196800,TRAESCS5B02G199400,TRAESCS5B02G297300,TRAESCS5B02G299900,TRAESCS5B02G384600,TRAESCS5B02G405700,TRAESCS5B02G485300,TRAESCS5D02G071900,TRAESCS5D02G161000,TRAESCS5D02G178800,TRAESCS5D02G238500,TRAESCS5D02G282000,TRAESCS5D02G305200,TRAESCS5D02G311100,TRAESCS5D02G495500,TRAESCS6A02G017600,TRAESCS6A02G017700,TRAESCS6A02G099600,TRAESCS6A02G117100,TRAESCS6A02G119000,TRAESCS6A02G154400,TRAESCS6A02G200300,TRAESCS6A02G223100,TRAESCS6A02G226600,TRAESCS6A02G227900,TRAESCS6A02G235300,TRAESCS6A02G254900,TRAESCS6A02G26418900,TRAESCS6A02G26996400,TRAESCS6B02G02201100,TRAESCS6B02G253900,TRAESCS6B02G256600,TRAESCS6B02G270600,TRAESCS6B02G279300,TRAESCS6B02G291800,TRAESCS6B02G296200,TRAESCS6B02G305400,TRAESCS6B02G31900,TRAESCS6B02G410400,TRAESCS6B02G472000,TRAESCS6D02G019800,TRAESCS6D02G123500,TRAESCS6D02G136300,TRAESCS6D02G182900,TRAESCS6D02G164900,TRAESCS6D02G168400,TRAESCS6D02G189600,TRAESCS6D02G210400,TRAESCS6D02G212100,TRAESCS6D02G236200,TRAESCS6D02G236400,TRAESCS6D02G244300,TRAESCS6D02G258000,TRAESCS6D02G288200,TRAESCS6D02G362000,TRAESCS7A02G002800,TRAESCS7A02G065700,TRAESCS7A02G066300,TRAESCS7A02G070100,TRAESCS7A02G076500,TRAESCS7A02G105100,TRAESCS7A02G120300,TRAESCS7A02G133500,TRAESCS7A02G250800,TRAESCS7A02G250900,TRAESCS7A02G251400,TRAESCS7A02G287400,TRAESCS7A02G299300,TRAESCS7A02G304700,TRAESCS7A02G357500,TRAESCS7A02G502400,TRAESCS7A02G504800,TRAESCS7A02G511900,TRAESCS7A02G516000,TRAESCS7A02G549200,TRAESCS7A02G549300,TRAESCS7B02G0222900,TRAESCS7B02G048700,TRAESCS7B02G094000,TRAESCS7B02G131700,TRAESCS7B02G139700,TRAESCS7B02G183300,TRAESCS7B02G197600,TRAESCS7B02G205000,TRAESCS7B02G214800,TRAESCS7B02G248100,TRAESCS7B02G260600,TRAESCS7B02G308200,TRAESCS7B02G320200,TRAESCS7B02G363200,TRAESCS7B02G403200,TRAESCS7B02G423200,TRAESCS7B02G432400,TRAESCS7B02G455500,TRAESCS7B02G472400,TRAESCS7D02G003300,TRAESCS7D02G072200,TRAESCS7D02G117800,TRAESCS7D02G133100,TRAESCS7D02G134100,TRAESCS7D02G23200,TRAESCS7D02G48600,TRAESCS7D02G49500,TRAESCS7D02G265600,TRAESCS7D02G282200,TRAESCS7D02G284900,TRAESCS7D02G310800,TRAESCS7D02G324100,TRAESCS7D02G34100,TRAESCS7D02G344200,TRAESCS7D02G365200,TRAESCS7D02G365200,TRAESCSU02G068900



GO:BP	macromolecule biosynthetic process	GO:0009059	1.68E-02	1.77	7756	547	130	42199	TRAESCS1A02G043600,TRAESCS1A02G190100,TRAESCS1A02G304400,TRAESCS1A02G403500,TRAESCS1A02G419600,TRAESCS1B02G085500,TRAESCS1B02G129400,TRAESCS1B02G169100,0,TRAESCS1B02G298100,TRAESCS1B02G343500,TRAESCS1B02G433100,TRAESCS1B02G449700,TRAESCS1B02G464000,TRAESCS1D02G044200,TRAESCS1D02G142900,TRAESCS1D02G198800,TRAESCS1D02G330900,TRAESCS1D02G332200,TRAESCS1D02G411400,TRAESCS1D02G427400,TRAESCS2A02G044800,TRAESCS2A02G189700,TRAESCS2A02G292700,TRAESCS2A02G297400,TRAESCS2A02G310300,TRAESCS2A02G352100,TRAESCS2B02G213800,TRAESCS2B02G309100,TRAESCS2B02G327300,TRAESCS2B02G335400,TRAESCS2B02G358800,TRAESCS2B02G374300,TRAESCS2D02G090300,TRAESCS2D02G198100,TRAESCS2D02G295300,TRAESCS2D02G386400,TRAESCS2D02G503900,TRAESCS3A02G080700,TRAESCS3A02G115000,TRAESCS3A02G163300,TRAESCS3A02G179700,TRAESCS3A02G186200,TRAESCS3A02G292700,TRAESCS3A02G325800,TRAESCS3A02G429100,TRAESCS3B02G141100,TRAESCS3B02G180800,TRAESCS3B02G204200,TRAESCS3B02G278000,TRAESCS3B02G327500,TRAESCS3B02G330800,TRAESCS3B02G354900,TRAESCS3B02G412300,TRAESCS3B02G466800,TRAESCS3D02G129900,TRAESCS3D02G169600,TRAESCS3D02G179800,TRAESCS3D02G190200,TRAESCS3D02G249100,TRAESCS3D02G319200,TRAESCS3D02G424600,TRAESCS3A02G013700,TRAESCS3A02G033000,TRAESCS3A02G038800,TRAESCS3A02G039700,TRAESCS3A02G068800,TRAESCS3A02G071100,TRAESCS3A02G086600,TRAESCS3A02G0924900,TRAESCS3A02G1289100,TRAESCS3A02G146900,TRAESCS3A02G1949100,TRAESCS3A02G215900,TRAESCS3A02G2174200,TRAESCS3A02G219500,TRAESCS3B02G354200,TRAESCS3B02G354400,TRAESCS3B02G354600,TRAESCS3D02G161000,TRAESCS3D02G178800,TRAESCS3D02G2003200,TRAESCS3D02G2495500,TRAESCS3A02G017600,TRAESCS6A02G117700,TRAESCS6A02G115400,TRAESCS6A02G154400,TRAESCS6A02G223100,TRAESCS6A02G235300,TRAESCS6A02G418900,TRAESCS6A02G019800,TRAESCS6A02G023500,TRAESCS6A02G036200,TRAESCS6A02G12100,TRAESCS6A02G36200,TRAESCS7A02G070100,TRAESCS7A02G076500,TRAESCS7A02G120300,TRAESCS7A02G133500,TRAESCS7A02G250800,TRAESCS7A02G250900,TRAESCS7A02G251400,TRAESCS7A02G287400,TRAESCS7A02G304700,TRAESCS7A02G357500,TRAESCS7A02G516000,TRAESCS7A02G549200,TRAESCS7A02G549300,TRAESCS7B02G048700,TRAESCS7B02G094000,TRAESCS7B02G139700,TRAESCS7B02G18300,TRAESCS7B02G205000,TRAESCS7B02G214800,TRAESCS7B02G308200,TRAESCS7B02G403200,TRAESCS7B02G432400,TRAESCS7B02G472400,TRAESCS7D02G072200,TRAESCS7D02G117800,TRAESCS7D02G133100,TRAESCS7D02G232200,TRAESCS7D02G248600,TRAESCS7D02G249500,TRAESCS7D02G265600,TRAESCS7D02G284900,TRAESCS7D02G310800,TRAESCS7D02G2636200,TRAESCS7D02G306300,TRAESCS7D02G335500,TRAESCS7D02G353600
GO:BP	chloroplast-nucleus signaling pathway	GO:0010019	1.84E-02	1.74	4	547	2	42199	TRAESCS5B02G405700,TRAESCS5D02G411100
GO:BP	trehalose metabolic process	GO:0005991	1.85E-02	1.73	59	547	5	42199	TRAESCS1A02G339300,TRAESCS1B02G351600,TRAESCS1D02G341100,TRAESCS2A02G61000,TRAESCS2B02G187000
GO:BP	leucine metabolic process	GO:0006551	1.92E-02	1.72	16	547	3	42199	TRAESCS2B02G609800,TRAESCS5A02G131300,TRAESCS5D02G139800
GO:BP	negative regulation of biological process	GO:0048519	1.98E-02	1.70	1142	547	28	42199	TRAESCS1A02G043600,TRAESCS1A02G137000,TRAESCS1A02G160900,TRAESCS1A02G190100,TRAESCS1A02G304400,TRAESCS1A02G333000,TRAESCS1A02G403500,TRAESCS1B02G085500,TRAESCS1B02G129400,TRAESCS1B02G155600,TRAESCS1B02G169100,TRAESCS1B02G298100,TRAESCS1B02G343500,TRAESCS1B02G433100,TRAESCS1B02G464000,TRAESCS1D02G044200,TRAESCS1D02G132500,TRAESCS1D02G142900,TRAESCS1D02G158000,TRAESCS1D02G198800,TRAESCS1D02G303900,TRAESCS1D02G332200,TRAESCS1D02G383400,TRAESCS1D02G411400,TRAESCS2A02G044800,TRAESCS2A02G189700,TRAESCS2A02G292700,TRAESCS2A02G297400,TRAESCS2A02G332100,TRAESCS2A02G507300,TRAESCS2B02G309100,TRAESCS2B02G355400,TRAESCS2D02G090300,TRAESCS2D02G198100,TRAESCS2D02G295300,TRAESCS2D02G386400,TRAESCS2D02G503900,TRAESCS3A02G080700,TRAESCS3A02G115000,TRAESCS3A02G163300,TRAESCS3A02G179700,TRAESCS3A02G186200,TRAESCS3A02G295300,TRAESCS3A02G414700,TRAESCS3A02G429100,TRAESCS3B02G141100,TRAESCS3B02G180800,TRAESCS3B02G204200,TRAESCS3B02G278000,TRAESCS3B02G330800,TRAESCS3B02G354200,TRAESCS3B02G354400,TRAESCS3B02G354600,TRAESCS3D02G161000,TRAESCS3D02G178800,TRAESCS3D02G198100,TRAESCS3D02G249500,TRAESCS3D02G295300,TRAESCS3D02G319200,TRAESCS3D02G424600,TRAESCS3A02G013700,TRAESCS3A02G033000,TRAESCS3A02G038800,TRAESCS3A02G039700,TRAESCS3A02G068800,TRAESCS3A02G071100,TRAESCS3A02G086600,TRAESCS3A02G0924900,TRAESCS3A02G1289100,TRAESCS3A02G146900,TRAESCS3A02G1949100,TRAESCS3A02G215900,TRAESCS3A02G2174200,TRAESCS3A02G219500,TRAESCS3B02G354200,TRAESCS3B02G354400,TRAESCS3B02G354600,TRAESCS3D02G161000,TRAESCS3D02G178800,TRAESCS3D02G198100,TRAESCS3D02G249500,TRAESCS3D02G295300,TRAESCS3A02G017600,TRAESCS6A02G117700,TRAESCS6A02G115400,TRAESCS6A02G154400,TRAESCS6A02G223100,TRAESCS6A02G235300,TRAESCS6A02G249500,TRAESCS6A02G418900,TRAESCS6B02G201100,TRAESCS6B02G207600,TRAESCS6B02G351900,TRAESCS6B02G0190800,TRAESCS6D02G023500,TRAESCS6D02G036200,TRAESCS6D02G164900,TRAESCS6D02G12100,TRAESCS6D02G236200,TRAESCS6D02G36200,TRAESCS7A02G070100,TRAESCS7A02G076500,TRAESCS7A02G120300,TRAESCS7A02G133500,TRAESCS7A02G250800,TRAESCS7A02G250900,TRAESCS7A02G251400,TRAESCS7B02G048700,TRAESCS7B02G094000,TRAESCS7B02G139700,TRAESCS7B02G18300,TRAESCS7B02G205000,TRAESCS7B02G214800,TRAESCS7B02G308200,TRAESCS7B02G403200,TRAESCS7B02G432400,TRAESCS7D02G072200,TRAESCS7D02G232200,TRAESCS7D02G248600,TRAESCS7D02G265600,TRAESCS7D02G284900,TRAESCS7D02G310800,TRAESCS7D02G335500,TRAESCS7D02G353600
GO:BP	nucleic acid metabolic process	GO:0090304	2.37E-02	1.62	7351	547	123	42199	
GO:BP	cell junction organization	GO:0034330	2.73E-02	1.56	5	547	2	42199	TRAESCS6A02G234000,TRAESCS6B02G262800
GO:BP	cell-cell junction organization	GO:0045216	2.73E-02	1.56	5	547	2	42199	TRAESCS6A02G234000,TRAESCS6B02G262800
GO:BP	regulation of starch metabolic process	GO:2000904	2.73E-02	1.56	5	547	2	42199	TRAESCS1B02G343500,TRAESCS1D02G332200
GO:BP	(1->3)-beta-D-glucan biosynthetic process	GO:0006075	2.82E-02	1.55	40	547	4	42199	TRAESCS3A02G292700,TRAESCS3B02G327500,TRAESCS6A02G154400,TRAESCS7B02G048700
GO:BP	(1->3)-beta-D-glucan metabolic process	GO:0006074	2.82E-02	1.55	40	547	4	42199	TRAESCS3A02G292700,TRAESCS3B02G327500,TRAESCS6A02G154400,TRAESCS7B02G048700
GO:BP	organelle fusion	GO:0048284	2.82E-02	1.55	40	547	4	42199	TRAESCS1B02G094700,TRAESCS5A02G094300,TRAESCS5B02G100400,TRAESCS5D02G106600
GO:BP	nucleobase-containing compound metabolic process	GO:0006139	2.89E-02	1.54	8333	547	136	42199	TRAESCS1A02G043600,TRAESCS1A02G190100,TRAESCS1A02G304400,TRAESCS1A02G403500,TRAESCS1A02G419600,TRAESCS1B02G085500,TRAESCS1B02G129400,TRAESCS1B02G169100,TRAESCS1B02G298100,TRAESCS1B02G343500,TRAESCS1B02G433100,TRAESCS1B02G464000,TRAESCS1D02G044200,TRAESCS1D02G132500,TRAESCS1D02G142900,TRAESCS1D02G158000,TRAESCS1D02G198800,TRAESCS1D02G303900,TRAESCS1D02G332200,TRAESCS1D02G383400,TRAESCS2A02G044800,TRAESCS2A02G189700,TRAESCS2A02G292700,TRAESCS2A02G297400,TRAESCS2A02G332100,TRAESCS2A02G507300,TRAESCS2B02G309100,TRAESCS2B02G355400,TRAESCS2D02G090300,TRAESCS2D02G198100,TRAESCS2D02G295300,TRAESCS2D02G386400,TRAESCS2D02G503900,TRAESCS3A02G080700,TRAESCS3A02G115000,TRAESCS3A02G163300,TRAESCS3A02G179700,TRAESCS3A02G186200,TRAESCS3A02G295300,TRAESCS3A02G414700,TRAESCS3A02G429100,TRAESCS3B02G141100,TRAESCS3B02G180800,TRAESCS3B02G204200,TRAESCS3B02G278000,TRAESCS3B02G330800,TRAESCS3B02G354200,TRAESCS3B02G354400,TRAESCS3B02G354600,TRAESCS3D02G161000,TRAESCS3D02G178800,TRAESCS3D02G198100,TRAESCS3D02G249500,TRAESCS3D02G295300,TRAESCS3A02G017600,TRAESCS6A02G117700,TRAESCS6A02G115400,TRAESCS6A02G154400,TRAESCS6A02G223100,TRAESCS6A02G235300,TRAESCS6A02G249500,TRAESCS6A02G418900,TRAESCS6B02G201100,TRAESCS6B02G207600,TRAESCS6B02G351900,TRAESCS6B02G0190800,TRAESCS6D02G023500,TRAESCS6D02G036200,TRAESCS6D02G164900,TRAESCS6D02G12100,TRAESCS6D02G236200,TRAESCS6D02G36200,TRAESCS7A02G070100,TRAESCS7A02G076500,TRAESCS7A02G120300,TRAESCS7A02G133500,TRAESCS7A02G250800,TRAESCS7A02G250900,TRAESCS7A02G251400,TRAESCS7B02G048700,TRAESCS7B02G094000,TRAESCS7B02G139700,TRAESCS7B02G18300,TRAESCS7B02G205000,TRAESCS7B02G214800,TRAESCS7B02G308200,TRAESCS7B02G403200,TRAESCS7B02G432400,TRAESCS7D02G072200,TRAESCS7D02G232200,TRAESCS7D02G248600,TRAESCS7D02G265600,TRAESCS7D02G284900,TRAESCS7D02G310800,TRAESCS7D02G335500,TRAESCS7D02G353600



									2G432400,TRAESC57D02G072200,TRAESC57D02G232200,TRAESC57D02G248600,TRAESC57D02G265600,TRAESC57D02G344100,TRAESC57D02G344200,TRAESC57D02G365200,TRAESC57D02G506300,TRAESC5U02G068900
GO:BP	phosphorelay signal transduction system	GO:0000160	2.89E-02	1.54	297	547	11	42199	TRAESC54B02G006500,TRAESC54B02G178900,TRAESC56A02G227900,TRAESC56B02G253900,TRAESC56B02G408800,TRAESC56D02G021700,TRAESC57A02G146400,TRAESC57B02G416400,TRAESC57D02G012700,TRAESC57D02G529900,TRAESC5U02G241200 TRAESC51A02G043600,TRAESC51A02G190100,TRAESC51A02G304400,TRAESC51A02G403500,TRAESC51A02G419600,TRAESC51B02G085500,TRAESC51B02G129400,TRAESC51B02G169100,TRAESC51B02G298100,TRAESC51B02G343500,TRAESC51B02G433100,TRAESC51B02G449700,TRAESC51B02G464000,TRAESC51D02G044200,TRAESC51D02G142900,TRAESC51D02G198800,TRAESC51D02G303900,TRAESC51D02G332200,TRAESC51D02G411400,TRAESC51D02G427400,TRAESC52A02G044800,TRAESC52A02G189700,TRAESC52A02G292700,TRAESC52A02G297400,TRAESC52A02G310300,TRAESC52A02G352100,TRAESC52B02G13800,TRAESC52B02G309100,TRAESC52B02G327300,TRAESC52B02G355400,TRAESC52B02G358800,TRAESC52D02G074300,TRAESC52D02G090300,TRAESC52D02G198100,TRAESC52D02G295300,TRAESC52D02G386400,TRAESC52D02G503900,TRAESC53A02G080700,TRAESC53A02G111500,TRAESC53A02G163300,TRAESC53A02G179700,TRAESC53A02G186200,TRAESC53A02G292700,TRAESC53A02G325800,TRAESC53A02G429100,TRAESC53B02G141100,TRAESC53B02G180800,TRAESC53B02G204200,TRAESC53B02G278000,TRAESC53B02G327500,TRAESC53B02G330800,TRAESC53B02G354900,TRAESC53B02G412300,TRAESC53B02G466800,TRAESC53D02G123900,TRAESC53D02G169600,TRAESC53D02G179800,TRAESC53D02G190200,TRAESC53D02G249100,TRAESC53D02G319200,TRAESC53D02G424600,TRAESC54A02G013700,TRAESC54B02G090000,TRAESC54B02G138800,TRAESC54B02G239700,TRAESC54B02G268800,TRAESC54D02G011300,TRAESC54D02G086600,TRAESC54D02G249900,TRAESC54D02G228900,TRAESC54D02G316900,TRAESC54D02G49100,TRAESC55A02G155900,TRAESC55A02G174200,TRAESC55A02G295700,TRAESC55B02G054200,TRAESC55B02G154100,TRAESC55B02G384600,TRAESC55D02G161000,TRAESC55D02G178800,TRAESC55D02G403200,TRAESC55D02G495300,TRAESC56A02G017600,TRAESC56A02G017700,TRAESC56A02G117100,TRAESC56A02G154400,TRAESC56A02G223100,TRAESC56A02G235300,TRAESC56A02G418900,TRAESC56D02G019800,TRAESC56D02G023500,TRAESC56D02G036200,TRAESC56D02G121100,TRAESC56D02G362000,TRAESC57A02G070100,TRAESC57A02G076500,TRAESC57A02G120300,TRAESC57A02G133500,TRAESC57A02G250800,TRAESC57A02G250900,TRAESC57A02G287400,TRAESC57A02G304700,TRAESC57A02G357500,TRAESC57A02G516000,TRAESC57A02G549200,TRAESC57A02G549300,TRAESC57B02G048700,TRAESC57B02G094000,TRAESC57B02G183300,TRAESC57B02G205000,TRAESC57B02G214800,TRAESC57B02G308200,TRAESC57B02G403200,TRAESC57B02G432400,TRAESC57B02G472400,TRAESC57D02G072200,TRAESC57D02G117800,TRAESC57D02G133100,TRAESC57D02G232200,TRAESC57D02G248600,TRAESC57D02G265600,TRAESC57D02G284900,TRAESC57D02G310800,TRAESC57D02G365200,TRAESC57D02G506300,TRAESC57D02G535500,TRAESC57D02G535600
GO:BP	cellular macromolecule biosynthetic process	GO:0034645	2.89E-02	1.54	7697	547	127	42199	TRAESC51B02G222500,TRAESC51D02G211700,TRAESC51D02G419800,TRAESC52B02G098000,TRAESC52D02G196900,TRAESC54A02G089800,TRAESC54D02G327300,TRAESC55A02G131300,TRAESC55A02G177400,TRAESC55A02G250100,TRAESC55B02G403400,TRAESC55D02G139800,TRAESC56B02G302500 TRAESC51A02G169000,TRAESC51A02G333000,TRAESC51A02G445500,TRAESC52A02G252400,TRAESC52B02G358800,TRAESC52D02G090300,TRAESC52D02G295300,TRAESC53A02G154400,TRAESC53B02G180800,TRAESC53D02G231600,TRAESC53D02G366200,TRAESC54A02G197000,TRAESC54A02G281500,TRAESC54A02G495800,TRAESC54B02G031500,TRAESC54D02G029200,TRAESC56A02G2223100,TRAESC56A02G227900,TRAESC56B02G253900,TRAESC56D02G212100,TRAESC57A02G028000,TRAESC57D02G003300
GO:BP	cellular amino acid biosynthetic process	GO:0008652	3.01E-02	1.52	391	547	13	42199	TRAESC51A02G043600,TRAESC51A02G137000,TRAESC51A02G160900,TRAESC51A02G190100,TRAESC51A02G304400,TRAESC51A02G333000,TRAESC51A02G403500,TRAESC51B02G085500,TRAESC51B02G129400,TRAESC51B02G155600,TRAESC51B02G169100,TRAESC51B02G173200,TRAESC51B02G298100,TRAESC51B02G343500,TRAESC51B02G433100,TRAESC51B02G464000,TRAESC51D02G041200,TRAESC51D02G044200,TRAESC51D02G132500,TRAESC51D02G142900,TRAESC51D02G158000,TRAESC51D02G198800,TRAESC51D02G303900,TRAESC51D02G332200,TRAESC51D02G383400,TRAESC51D02G411400,TRAESC52A02G044800,TRAESC52A02G189700,TRAESC52A02G292700,TRAESC52A02G297400,TRAESC52A02G352100,TRAESC52B02G074300,TRAESC52B02G192900,TRAESC52B02G309100,TRAESC52B02G355400,TRAESC52D02G012000,TRAESC52D02G074300,TRAESC52D02G198100,TRAESC52D02G295300,TRAESC52D02G503900,TRAESC52D02G565300,TRAESC53A02G080700,TRAESC53A02G111500,TRAESC53A02G163300,TRAESC53A02G169900,TRAESC53A02G179700,TRAESC53A02G186200,TRAESC53A02G325800,TRAESC53A02G414700,TRAESC53A02G429100,TRAESC53A02G480900,TRAESC53B02G141100,TRAESC53B02G180800,TRAESC53B02G204200,TRAESC53B02G238800,TRAESC53B02G278000,TRAESC53B02G330800,TRAESC53B02G354900,TRAESC53B02G412300,TRAESC53B02G466800,TRAESC53D02G123900,TRAESC53D02G169600,TRAESC53D02G179800,TRAESC53D02G190200,TRAESC53D02G249100,TRAESC53D02G319200,TRAESC53D02G424600,TRAESC53D02G475600,TRAESC54A02G089800,TRAESC54A02G2256800,TRAESC54A02G281500,TRAESC54B02G031500,TRAESC54B02G138800,TRAESC54B02G239700,TRAESC54B02G268800,TRAESC54D02G011300,TRAESC54D02G044700,TRAESC54D02G249900,TRAESC54D02G316900,TRAESC54D02G327300,TRAESC55A02G049100,TRAESC55A02G100300,TRAESC55A02G155900,TRAESC55A02G174200,TRAESC55A02G295700,TRAESC55A02G466100,TRAESC55A02G501900,TRAESC55B02G054200,TRAESC55B02G062900,TRAESC55B02G154100,TRAESC55B02G196800,TRAESC55B02G384600,TRAESC55B02G485300,TRAESC55D02G112800,TRAESC55D02G161000,TRAESC55D02G178800,TRAESC55D02G292700,TRAESC55D02G327300,TRAESC55D02G403200,TRAESC56A02G017600,TRAESC56A02G017700,TRAESC56A02G119000,TRAESC56A02G223100,TRAESC56A02G235300,TRAESC56A02G254900,TRAESC56A02G277800,TRAESC56A02G418900,TRAESC56B02G0201100,TRAESC56B02G0270600,TRAESC56B02G302500,TRAESC56B02G351900,TRAESC56D02G019800,TRAESC56D02G021800,TRAESC56D02G023500,TRAESC56D02G036200,TRAESC56D02G121100,TRAESC56D02G362000,TRAESC57A02G076500,TRAESC57A02G250800,TRAESC57A02G250900,TRAESC57A02G304700,TRAESC57A02G357500,TRAESC57A02G516000,TRAESC57B02G048700,TRAESC57B02G094000,TRAESC57B02G248100,TRAESC57B02G308200,TRAESC57B02G365200,TRAESC57B02G403200,TRAESC57D02G432400,TRAESC57D02G072200,TRAESC57D02G232200,TRAESC57D02G248600,TRAESC57D02G265600,TRAESC57D02G344100,TRAESC57D02G344200,TRAESC57D02G365200,TRAESC57D02G506300,TRAESC5U02G068900
GO:BP	heterocycle metabolic process	GO:0046483	3.09E-02	1.51	8925	547	144	42199	
GO:BP	imidazole-containing compound metabolic process	GO:0052803	3.14E-02	1.50	21	547	3	42199	TRAESC54D02G327300,TRAESC55A02G0501900,TRAESC56B02G302500
GO:BP	cell wall hydroxyproline-rich glycoprotein metabolic process	GO:0010404	3.14E-02	1.50	6	547	2	42199	TRAESC54A02G013700,TRAESC54D02G289100
GO:BP	regulation of cation transmembrane transport	GO:1904062	3.14E-02	1.50	6	547	2	42199	TRAESC55B02G185200,TRAESC55D02G192100
GO:BP	positive regulation of ubiquitin protein ligase activity	GO:1904668	3.14E-02	1.50	21	547	3	42199	TRAESC53A02G527600,TRAESC53B02G595600,TRAESC56B02G296200
GO:BP	cell wall proteoglycan metabolic process	GO:0010384	3.14E-02	1.50	6	547	2	42199	TRAESC54A02G013700,TRAESC54D02G289100
GO:BP	histidine biosynthetic process	GO:0000105	3.14E-02	1.50	21	547	3	42199	TRAESC54D02G327300,TRAESC55A02G0501900,TRAESC56B02G302500
GO:BP	regulation of polysaccharide biosynthetic process	GO:0032885	3.14E-02	1.50	6	547	2	42199	TRAESC51B02G343500,TRAESC51D02G332200
GO:BP	regulation of proton transport	GO:0010155	3.14E-02	1.50	6	547	2	42199	TRAESC55B02G185200,TRAESC55D02G192100
GO:BP	regulation of glucan biosynthetic process	GO:0010962	3.14E-02	1.50	6	547	2	42199	TRAESC51B02G343500,TRAESC51D02G332200
GO:BP	proteoglycan metabolic process	GO:0006029	3.14E-02	1.50	6	547	2	42199	TRAESC54A02G013700,TRAESC54D02G289100
GO:BP	histidine metabolic process	GO:0006547	3.14E-02	1.50	21	547	3	42199	TRAESC54D02G327300,TRAESC55A02G0501900,TRAESC56B02G302500



GO:BP	disaccharide biosynthetic process	GO:0046351	3.15E-02	1.50	72	547	5	42199	TRAESC51A02G339300,TRAESC51B02G351600,TRAESC51D02G341100,TRAESC52A02G161000,TRAESC52B02G187000 TRAESC51A02G425100,TRAESC51B02G169100,TRAESC51B02G222500,TRAESC51D02G41200,TRAESC51D02G198800,TRAESC51D02G21700,TRAESC51D02G419800,TRAESC51D02G43410 0,TRAESC52B02G192900,TRAESC52B02G345400,TRAESC52B02G609800,TRAESC52D02G196900,TRAESC53A02G111500,TRAESC53A02G163300,TRAESC53A02G186200,TRAESC53A02G305 400,TRAESC53A02G480900,TRAESC53B02G525400,TRAESC53D02G169600,TRAESC53D02G190200,TRAESC53D02G475600,TRAESC54A02G501900,TRAESC55B02G102400,TRAESC55D02G109200,TRAESC54D02G2 43900,TRAESC54D02G327300,TRAESC55A02G131300,TRAESC55A02G177400,TRAESC55A02G501900,TRAESC55B02G102400,TRAESC55B02G403400,TRAESC55D02G109200,TRAESC55D02 G139800,TRAESC56A02G017500,TRAESC56B02G299700,TRAESC56B02G302500,TRAESC57A02G304700,TRAESC57A02G512300,TRAESC57A02G516000,TRAESC57B02G205000,TRAESC57B0 2G432400,TRAESC57D02G506300
GO:BP	oxoacid metabolic process	GO:0043436	3.53E-02	1.45	2001	547	41	42199	2G432400,TRAESC57D02G506300
GO:BP	microtubule-based process	GO:0007017	3.53E-02	1.45	407	547	13	42199	TRAESC51B02G118200,TRAESC51B02G123200,TRAESC51D02G407800,TRAESC52A02G305500,TRAESC52B02G229500,TRAESC52B02G409900,TRAESC53A02G169800,TRAESC53A02G60980 0,TRAESC56B02G365300,TRAESC57A02G066300,TRAESC57B02G304900,TRAESC57D02G313500,TRAESC57D02G399400 TRAESC51A02G043600,TRAESC51A02G087500,TRAESC51A02G14900,TRAESC51A02G137000,TRAESC51A02G16900,TRAESC51A02G190100,TRAESC51A02G233200,TRAESC51A02G30344 0,TRAESC51A02G333000,TRAESC51A02G339300,TRAESC51A02G402400,TRAESC51A02G403500,TRAESC51A02G419600,TRAESC51A02G425100,TRAESC51A02G445500,TRAESC51B02G08 5500,TRAESC51B02G094700,TRAESC51B02G106400,TRAESC51B02G129400,TRAESC51B02G134800,TRAESC51B02G154400,TRAESC51B02G155600,TRAESC51B02G169100,TRAESC51B02G17 3200,TRAESC51B02G200200,TRAESC51B02G222500,TRAESC51B02G236600,TRAESC51B02G294900,TRAESC51B02G298100,TRAESC51B02G343500,TRAESC51B02G351600,TRAESC51B02G34 1400,TRAESC51B02G433100,TRAESC51B02G449700,TRAESC51B02G464000,TRAESC51D02G041200,TRAESC51D02G044200,TRAESC51D02G133500,TRAESC51D02G136400,TRAESC51D02G 142900,TRAESC51D02G158800,TRAESC51D02G198800,TRAESC51D02G211700,TRAESC51D02G235500,TRAESC51D02G303900,TRAESC51D02G332200,TRAESC51D02G341100,TRAESC51D0 2G366800,TRAESC51D02G371200,TRAESC51D02G383400,TRAESC51D02G410100,TRAESC51D02G411400,TRAESC51D02G419800,TRAESC51D02G427400,TRAESC51D02G434100,TRAESC52 A02G043100,TRAESC52A02G044800,TRAESC52A02G104500,TRAESC52A02G135400,TRAESC52A02G159300,TRAESC52A02G161000,TRAESC52A02G189700,TRAESC52A02G192200,TRAESC 52A02G248200,TRAESC52A02G252400,TRAESC52A02G266700,TRAESC52A02G292700,TRAESC52A02G297400,TRAESC52A02G310300,TRAESC52A02G334300,TRAESC52A02G352100,TRA ESC52A02G382000,TRAESC52A02G386700,TRAESC52A02G389600,TRAESC52A02G400900,TRAESC52A02G507300,TRAESC52B02G159400,TRAESC52B02G217500,TRAESC52B02G187000,TRA ESC52B02G192900,TRAESC52B02G213800,TRAESC52B02G214100,TRAESC52B02G49900,TRAESC52B02G62900,TRAESC52B02G630900,TRAESC52B02G6327300,TRAESC52B02G634500,TRA ESC52B02G635400,TRAESC52B02G638800,TRAESC52B02G690800,TRAESC52D02G012000,TRAESC52D02G022900,TRAESC52D02G0294300,TRAESC52D02G090300,TRAESC52D02G155200,TRA ESC52D02G196900,TRAESC52D02G198100,TRAESC52D02G253800,TRAESC52D02G295300,TRAESC52D02G328900,TRAESC52D02G386400,TRAESC52D02G402200,TRAESC52D02G42670 0,TRAESC52D02G503900,TRAESC52D02G565300,TRAESC53A02G080700,TRAESC53A02G111500,TRAESC53A02G154400,TRAESC53A02G159900,TRAESC53A02G163300,TRAESC53A02G186200 980,TRAESC53A02G197900,TRAESC53A02G186200,TRAESC53A02G216100,TRAESC53A02G222200,TRAESC53A02G229900,TRAESC53A02G292700,TRAESC53A02G305400,TRAESC53A02 G325800,TRAESC53A02G351800,TRAESC53A02G365800,TRAESC53A02G400600,TRAESC53A02G414700,TRAESC53A02G429100,TRAESC53A02G480900,TRAESC53A02G527600,TRAESC53B 02G018100,TRAESC53B02G141100,TRAESC53B02G148500,TRAESC53B02G180800,TRAESC53B02G204200,TRAESC53B02G233800,TRAESC53B02G252500,TRAESC53B02G2646800,TRAESC53B 02G327500,TRAESC53B02G330800,TRAESC53B02G354900,TRAESC53B02G412300,TRAESC53B02G413800,TRAESC53B02G415800,TRAESC53B02G437300,TRAESC53B02G446800,TRAESC53B 02G510500,TRAESC53B02G525400,TRAESC53B02G530500,TRAESC53B02G544700,TRAESC53B02G595600,TRAESC53D02G018000,TRAESC53D02G123900,TRAESC53D02G124100,TRAESC53D 02G164700,TRAESC53D02G167100,TRAESC53D02G169600,TRAESC53D02G179800,TRAESC53D02G190200,TRAESC53D02G230100,TRAESC53D02G231600,TRAESC53D02G249100,TRAESC53D 02G295200,TRAESC53D02G305500,TRAESC53D02G319200,TRAESC53D02G354500,TRAESC53D02G358900,TRAESC53D02G366200,TRAESC53D02G366900,TRAESC53D02G376500,TRA ESC53D02G424600,TRAESC53D02G466500,TRAESC53D02G475600,TRAESC54A02G013700,TRAESC54A02G089800,TRAESC54A02G149300,TRAESC54A02G197000,TRAESC54A02G226800,TRA ESC54A02G263900,TRAESC54A02G281500,TRAESC54A02G422300,TRAESC54A02G485600,TRAESC54A02G495800,TRAESC54B02G031500,TRAESC54B02G090000,TRAESC54B02G137100, TRAESC54B02G138800,TRAESC54B02G151800,TRAESC54B02G157300,TRAESC54B02G216900,TRAESC54B02G239700,TRAESC54B02G268800,TRAESC54B02G381900,TRAESC54D02G007200, TRAESC54D02G011300,TRAESC54D02G029200,TRAESC54D02G04700,TRAESC54D02G074900,TRAESC54D02G082500,TRAESC54D02G088600,TRAESC54D02G171500,TRAESC54D02G1852 00,TRAESC54D02G199800,TRAESC54D02G204900,TRAESC54D02G243900,TRAESC54D02G281900,TRAESC54D02G316900,TRAESC54D02G318400,TRAESC54D02G327300,TRAESC54D02G3 57800,TRAESC55A02G0049100,TRAESC55A02G006300,TRAESC55A02G100100,TRAESC55A02G100300,TRAESC55A02G131300,TRAESC55A02G155900,TRAESC55A02G174200,TRAESC55A02 G177400,TRAESC55A02G200800,TRAESC55A02G236300,TRAESC55A02G244000,TRAESC55A02G295700,TRAESC55A02G321500,TRAESC55A02G392500,TRAESC55A02G466100,TRAESC55 A02G501900,TRAESC55A02G547900,TRAESC55B02G054200,TRAESC55B02G062900,TRAESC55B02G102400,TRAESC55B02G105200,TRAESC55B02G154100,TRAESC55B02G198800,TRAESC5 5B02G199400,TRAESC55B02G297300,TRAESC55B02G299900,TRAESC55B02G384600,TRAESC55B02G400000,TRAESC55B02G403400,TRAESC55B02G405700,TRAESC55B02G485300,TRAESC5 5D02G071900,TRAESC55D02G109200,TRAESC55D02G112800,TRAESC55D02G139800,TRAESC55D02G161000,TRAESC55D02G178800,TRAESC55D02G238500,TRAESC55D02G292700,TRAESC 55D02G327300,TRAESC55D02G328300,TRAESC55D02G403200,TRAESC55D02G404500,TRAESC55D02G411100,TRAESC55D02G495500,TRAESC56A02G017500,TRAESC56A02G017600,TRAESC56A02G022 100,TRAESC56A02G267000,TRAESC56A02G099600,TRAESC56A02G117100,TRAESC56A02G119000,TRAESC56A02G154400,TRAESC56A02G200300,TRAESC56A02G207600,TRAESC56A02G223100,TRA ESC56A02G227600,TRAESC56A02G227900,TRAESC56A02G235300,TRAESC56A02G254900,TRAESC56A02G277800,TRAESC56A02G418900,TRAESC56B02G009600,TRAESC56B02G221000, TRAESC56B02G253900,TRAESC56B02G256000,TRAESC56B02G270600,TRAESC56B02G279300,TRAESC56B02G291800,TRAESC56B02G292600,TRAESC56B02G297000,TRAESC56B02G302500, TRAESC56B02G305400,TRAESC56B02G3351900,TRAESC56B02G410400,TRAESC56B02G473000,TRAESC56B02G019800,TRAESC56B02G021800,TRAESC56B02G0235900,TRAESC56B02G026200, 0,TRAESC56D02G041100,TRAESC56D02G082900,TRAESC56D02G164900,TRAESC56D02G168400,TRAESC56D02G189600,TRAESC56D02G210400,TRAESC56D02G212100,TRAESC56D02G22 6200,TRAESC56D02G234600,TRAESC56D02G244300,TRAESC56D02G258000,TRAESC56D02G288200,TRAESC56D02G326200,TRAESC57A02G028900,TRAESC57A02G065700,TRAESC57A02G0657200, G066300,TRAESC57A02G070100,TRAESC57A02G076500,TRAESC57A02G105100,TRAESC57A02G120300,TRAESC57A02G133500,TRAESC57A02G226800,TRAESC57A02G250800,TRAESC57 A02G250900,TRAESC57A02G251400,TRAESC57A02G287400,TRAESC57A02G299300,TRAESC57A02G304700,TRAESC57A02G3357500,TRAESC57A02G3502400,TRAESC57A02G3504800,TRAESC 57A02G511900,TRAESC57A02G512300,TRAESC57A02G516000,TRAESC57A02G549200,TRAESC57A02G549300,TRAESC57B02G022900,TRAESC57B02G029400,TRAESC57B02G048700,TRAESC 57B02G094000,TRAESC57B02G131700,TRAESC57B02G139700,TRAESC57B02G183300,TRAESC57B02G192200,TRAESC57B02G197600,TRAESC57B02G205000,TRAESC57B02G214800,TRAESC 57B02G248100,TRAESC57B02G260600,TRAESC57B02G308200,TRAESC57B02G320200,TRAESC57B02G363200,TRAESC57B02G403200,TRAESC57B02G423200,TRAESC57B02G432400,TRAESC 57B02G455500,TRAESC57B02G472400,TRAESC57D02G003300,TRAESC57D02G054800,TRAESC57D02G072200,TRAESC57D02G117800,TRAESC57D02G133100,TRAESC57D02G199200,TRA ESC57D02G212800,TRAESC57D02G232200,TRAESC57D02G248600,TRAESC57D02G249500,TRAESC57D02G265600,TRAESC57D02G282200,TRAESC57D02G284900,TRAESC57D02G310800, TRAESC57D02G344100,TRAESC57D02G344200,TRAESC57D02G365200,TRAESC57D02G487900,TRAESC57D02G500300,TRAESC57D02G500900,TRAESC57D02G506300,TRAESC57D02G5247 00,TRAESC57D02G535500,TRAESC57D02G535600,TRAESC57D02G608900,TRAESC57D02G10700,TRAESC57D02G138100,TRAESC57D02G27000
GO:BP	vesicle fusion	GO:0006906	3.75E-02	1.43	23	547	3	42199	TRAESC55A02G094300,TRAESC55B02G100400,TRAESC55D02G106600
GO:BP	mRNA modification	GO:0016556	3.75E-02	1.43	23	547	3	42199	TRAESC56A02G254900,TRAESC56B02G270600,TRAESC56D02G236200 TRAESC51A02G425100,TRAESC51B02G169100,TRAESC51B02G222500,TRAESC51D02G41200,TRAESC51D02G198800,TRAESC51D02G419800,TRAESC51D02G434100,TRAESC51D02G34310 0,TRAESC52B02G192900,TRAESC52B02G345400,TRAESC52B02G609800,TRAESC52D02G196900,TRAESC53A02G111500,TRAESC53A02G163300,TRAESC53A02G186200,TRAESC53A02G305 400,TRAESC53A02G480900,TRAESC53B02G525400,TRAESC53D02G169600,TRAESC53D02G190200,TRAESC53D02G475600,TRAESC54A02G501900,TRAESC55B02G102400,TRAESC55D02G109200,TRAESC54D02G2 43900,TRAESC54D02G327300,TRAESC55A02G131300,TRAESC55A02G177400,TRAESC55A02G501900,TRAESC55B02G102400,TRAESC55B02G403400,TRAESC55D02G109200,TRAESC55D02G2 43900,TRAESC54D02G327300,TRAESC55A02G131300,TRAESC55A02G177400,TRAESC55A02G501900,TRAESC55B02G102400,TRAESC55B02G403400,TRAESC55D02G109200,TRAESC55D02G2 43900,TRAESC56A02G017500,TRAESC56B02G299700,TRAESC56B02G302500,TRAESC57A02G304700,TRAESC57A02G512300,TRAESC57A02G516000,TRAESC57B02G205000,TRAESC57B0 2G432400,TRAESC57D02G506300
GO:BP	organic acid metabolic process	GO:0006082	3.75E-02	1.43	2011	547	41	42199	2G432400,TRAESC57D02G506300
GO:BP	coenzyme A metabolic process	GO:0015936	3.75E-02	1.43	23	547	3	42199	TRAESC53A02G169900,TRAESC55D02G327300,TRAESC56A02G277800
GO:BP	vesicle docking	GO:0048278	3.76E-02	1.42	47	547	4	42199	TRAESC54A02G025100,TRAESC54D02G277500,TRAESC55A02G106300,TRAESC55B02G11300
GO:BP	alpha-amino acid biosynthetic process	GO:1901607	4.47E-02	1.35	327	547	11	42199	TRAESC51D02G419800,TRAESC52B02G609800,TRAESC52D02G196900,TRAESC54A02G089800,TRAESC54D02G327300,TRAESC55A02G131300,TRAESC55A02G16900,TRAESC55A02G3205019 00,TRAESC55B02G403400,TRAESC55D02G139800,TRAESC56B02G302500



									A02G38960, TRAESCS2A02G40090, TRAESCS2A02G40270, TRAESCS2A02G45500, TRAESCS2A02G521300, TRAESCS2B02G17550, TRAESCS2B02G19120, TRAESCS2B02G27180, TRAESCS2B02G30910, TRAESCS2B02G32510, TRAESCS2B02G35400, TRAESCS2B02G40990, TRAESCS2D02G07430, TRAESCS2D02G15520, TRAESCS2D02G19810, TRAESCS2D02G26900, TRAESCS2D02G25270, TRAESCS2D02G26900, TRAESCS2D02G29530, TRAESCS2D02G30390, TRAESCS2D02G35600, TRAESCS2D02G35630, TRAESCS3A02G07760, TRAESCS3A02G15440, TRAESCS3A02G17970, TRAESCS3A02G19030, TRAESCS3A02G19520, TRAESCS3A02G21010, TRAESCS3A02G21640, TRAESCS3A02G23280, TRAESCS3A02G24170, TRAESCS3A02G242150, TRAESCS3B02G01810, TRAESCS3B02G08520, TRAESCS3B02G12410, TRAESCS3B02G14110, TRAESCS3B02G14230, TRAESCS3B02G19200, TRAESCS3B02G24060, TRAESCS3B02G27800, TRAESCS3B02G28120, TRAESCS3B02G31340, TRAESCS3B02G35490, TRAESCS3B02G40020, TRAESCS3B02G41230, TRAESCS3B02G41380, TRAESCS3B02G43730, TRAESCS3D02G01800, TRAESCS3D02G24910, TRAESCS3D02G31920, TRAESCS4A02G01370, TRAESCS4A02G05020, TRAESCS4A02G14930, TRAESCS4A02G25680, TRAESCS4A02G28150, TRAESCS4A02G40900, TRAESCS4B02G03150, TRAESCS4B02G04260, TRAESCS4B02G23970, TRAESCS4B02G25440, TRAESCS4B02G26880, TRAESCS4B02G38520, TRAESCS4D02G00720, TRAESCS4D02G01130, TRAESCS4D02G16710, TRAESCS4D02G22940, TRAESCS4D02G24060, TRAESCS4D02G25420, TRAESCS4D02G28910, TRAESCS4D02G30450, TRAESCS5A02G04910, TRAESCS5A02G06980, TRAESCS5A02G09430, TRAESCS5A02G11230, TRAESCS5A02G11590, TRAESCS5A02G29570, TRAESCS5A02G39530, TRAESCS5A02G44610, TRAESCS5A02G48570, TRAESCS5B02G05420, TRAESCS5B02G08540, TRAESCS5B02G10040, TRAESCS5B02G15410, TRAESCS5B02G17780, TRAESCS5B02G18520, TRAESCS5B02G23540, TRAESCS5B02G29990, TRAESCS5B02G37570, TRAESCS5B02G40000, TRAESCS5B02G40570, TRAESCS5B02G42600, TRAESCS5D02G04840, TRAESCS5D02G05120, TRAESCS5D02G07610, TRAESCS5D02G10660, TRAESCS5D02G16100, TRAESCS5D02G19210, TRAESCS5D02G23850, TRAESCS5D02G24290, TRAESCS5D02G32830, TRAESCS5D02G39660, TRAESCS5D02G40320, TRAESCS5D02G40450, TRAESCS5D02G41100, TRAESCS5D02G52180, TRAESCS6A02G07540, TRAESCS6A02G11710, TRAESCS6A02G22640, TRAESCS6A02G22790, TRAESCS6A02G23400, TRAESCS6A02G25490, TRAESCS6A02G25860, TRAESCS6A02G41890, TRAESCS6A02G42080, TRAESCS6B02G10010, TRAESCS6B02G20110, TRAESCS6B02G25390, TRAESCS6B02G25600, TRAESCS6B02G26280, TRAESCS6B02G27060, TRAESCS6B02G29970, TRAESCS6B02G35210, TRAESCS6B02G40880, TRAESCS6B02G45070, TRAESCS6D02G04050, TRAESCS6D02G04060, TRAESCS6D02G15940, TRAESCS6D02G16840, TRAESCS6D02G21040, TRAESCS6D02G23620, TRAESCS6D02G25800, TRAESCS7A02G06630, TRAESCS7A02G07010, TRAESCS7A02G07650, TRAESCS7A02G12030, TRAESCS7A02G13500, TRAESCS7A02G13500, TRAESCS7A02G13510, TRAESCS7A02G13520, TRAESCS7A02G14640, TRAESCS7A02G25080, TRAESCS7A02G25090, TRAESCS7A02G25140, TRAESCS7A02G28700, TRAESCS7A02G28740, TRAESCS7A02G35750, TRAESCS7B02G02850, TRAESCS7B02G07140, TRAESCS7B02G09400, TRAESCS7B02G13970, TRAESCS7B02G18330, TRAESCS7B02G19760, TRAESCS7B02G36320, TRAESCS7D02G01780, TRAESCS7D02G11780, TRAESCS7D02G12740, TRAESCS7D02G12800, TRAESCS7D02G13310, TRAESCS7D02G13500, TRAESCS7D02G24950, TRAESCS7D02G26560, TRAESCS7D02G28490, TRAESCS7D02G31350, TRAESCS7D02G34400, TRAESCS7D02G36520, TRAESCS7D02G38340, TRAESCS7D02G32590, TRAESCSU02G06890, TRAESCSU02G27000
GO:CC	amyloplast	GO:0009501	5.66E-06	5.25	22	393	6	32923	TRAESCS1D02G42740, TRAESCS7A02G07010, TRAESCS7A02G12030, TRAESCS7B02G18330, TRAESCS7D02G11780, TRAESCS7D02G28490
GO:CC	isoamylase complex	GO:0043033	4.20E-05	4.38	3	393	3	32923	TRAESCS7A02G25140, TRAESCS7B02G13970, TRAESCS7D02G24950
GO:CC	chloroplast isoamylase complex	GO:0010368	4.20E-05	4.38	3	393	3	32923	TRAESCS7A02G25140, TRAESCS7B02G13970, TRAESCS7D02G24950
GO:CC	RNAi effector complex	GO:0031332	1.36E-04	3.87	4	393	3	32923	TRAESCS4A02G28150, TRAESCS4B02G03150, TRAESCS4D02G02920
GO:CC	RISC complex	GO:0016442	1.36E-04	3.87	4	393	3	32923	TRAESCS4A02G28150, TRAESCS4B02G03150, TRAESCS4D02G02920
GO:CC	CCR4-NOT core complex	GO:0030015	2.87E-04	3.54	27	393	5	32923	TRAESCS1A02G30440, TRAESCS1D02G30390, TRAESCS4D02G01130, TRAESCS5A02G29570, TRAESCS6A02G41890
GO:CC	DNA polymerase III complex	GO:0009360	4.24E-04	3.37	15	393	4	32923	TRAESCS4B02G09000, TRAESCS4D02G08600, TRAESCS7B02G12480, TRAESCS7D02G31080
GO:CC	Set1C/COMPASS complex	GO:0048188	2.86E-03	2.54	10	393	3	32923	TRAESCS4A02G05020, TRAESCS4B02G25440, TRAESCS4D02G25420
GO:CC	CCR4-NOT complex	GO:0030014	2.86E-03	2.54	70	393	6	32923	TRAESCS1A02G30440, TRAESCS1D02G30390, TRAESCS4A02G22580, TRAESCS4D02G01130, TRAESCS5A02G29570, TRAESCS6A02G41890, TRAESCS1B02G23350, TRAESCS2A02G27000, TRAESCS2D02G26900, TRAESCS3A02G29270, TRAESCS3B02G32750, TRAESCS4A02G05020, TRAESCS4B02G04260, TRAESCS4B02G09000, TRAESCS4B02G25440, TRAESCS4D02G08600, TRAESCS4D02G24060, TRAESCS4D02G25420, TRAESCS5D02G40320, TRAESCS6A02G15440, TRAESCS7B02G04870, TRAESCS7B02G21800, TRAESCS7D02G28220, TRAESCS7D02G31080
GO:CC	transferase complex	GO:1990234	3.45E-03	2.46	580	393	18	32923	TRAESCS4B02G09000, TRAESCS4D02G08600, TRAESCS7B02G12480, TRAESCS7D02G31080
GO:CC	DNA polymerase complex	GO:0042575	3.72E-03	2.43	27	393	4	32923	TRAESCS4B02G09000, TRAESCS4D02G08600, TRAESCS7B02G12480, TRAESCS7D02G31080
GO:CC	AP-type membrane coat adaptor complex	GO:0030119	4.01E-03	2.40	50	393	5	32923	TRAESCS2B02G19120, TRAESCS2D02G17210, TRAESCS3A02G24150, TRAESCS6D02G25110, TRAESCS7D02G12890
GO:CC	clathrin adaptor complex	GO:0030131	7.34E-03	2.13	33	393	4	32923	TRAESCS2B02G19120, TRAESCS2D02G17210, TRAESCS6D02G25110, TRAESCS7D02G12890
GO:CC	catalytic complex	GO:1902494	8.81E-03	2.05	1156	393	27	32923	TRAESCS1A02G13700, TRAESCS1A02G23320, TRAESCS1B02G15560, TRAESCS1B02G23350, TRAESCS2A02G27000, TRAESCS2A02G32070, TRAESCS2D02G26900, TRAESCS3A02G29270, TRAESCS3B02G32750, TRAESCS4A02G05020, TRAESCS4B02G04260, TRAESCS4B02G09000, TRAESCS4B02G25440, TRAESCS4D02G00720, TRAESCS4D02G08600, TRAESCS4D02G24060, TRAESCS4D02G25420, TRAESCS5B02G29990, TRAESCS5D02G40320, TRAESCS6A02G15440, TRAESCS7A02G25140, TRAESCS7B02G04870, TRAESCS7B02G13970, TRAESCS7B02G21800, TRAESCS7D02G24950, TRAESCS7D02G28220, TRAESCS7D02G31080
GO:CC	AP-4 adaptor complex	GO:0030124	8.95E-03	2.05	4	393	2	32923	TRAESCS2B02G19120, TRAESCS3A02G24150
GO:CC	histone methyltransferase complex	GO:0035097	1.23E-02	1.91	18	393	3	32923	TRAESCS4A02G05020, TRAESCS4B02G25440, TRAESCS4D02G25420
GO:CC	1,3-beta-D-glucan synthase complex	GO:0000148	1.27E-02	1.90	40	393	4	32923	TRAESCS3A02G29270, TRAESCS3B02G32750, TRAESCS6A02G15440, TRAESCS7B02G04870
GO:CC	elongin complex	GO:0070449	1.93E-02	1.71	6	393	2	32923	TRAESCS1A02G04360, TRAESCS1D02G04420
GO:CC	signal recognition particle	GO:0048500	1.97E-02	1.71	46	393	4	32923	TRAESCS1D02G32930, TRAESCS3A02G40260, TRAESCS3B02G43600, TRAESCS3D02G39760



GO:CC	membrane coat	GO:0030117	2.01E-02	1.70	154	393	7	32923	TRAECS1B02G257500,TRAECS2B02G191200,TRAECS2D02G172100,TRAECS3A02G195200,TRAECS3A02G421500,TRAECS6D02G251100,TRAECS7D02G128900
GO:CC	clathrin coat	GO:0030118	2.01E-02	1.70	47	393	4	32923	TRAECS2B02G191200,TRAECS2D02G172100,TRAECS6D02G251100,TRAECS7D02G128900
GO:CC	coated membrane	GO:0048475	2.01E-02	1.70	154	393	7	32923	TRAECS1B02G257500,TRAECS2B02G191200,TRAECS2D02G172100,TRAECS3A02G195200,TRAECS3A02G421500,TRAECS6D02G251100,TRAECS7D02G128900 TRAECS1A02G158800,TRAECS1A02G165700,TRAECS1A02G19600,TRAECS1A02G425100,TRAECS1B02G094700,TRAECS1B02G169100,TRAECS1B02G173200,TRAECS1B02G20020,TRAECS1B02G224200,TRAECS1B02G233500,TRAECS1B02G257500,TRAECS1B02G449700,TRAECS1D02G198800,TRAECS1D02G281000,TRAECS1D02G329500,TRAECS1D02G4427400,TRAECS1D02G434100,TRAECS2A02G104500,TRAECS2A02G159300,TRAECS2A02G359000,TRAECS2A02G386700,TRAECS2A02G400900,TRAECS2A02G455000,TRAECS2A02G517100,TRAECS2B02G175500,TRAECS2B02G191200,TRAECS2B02G213800,TRAECS2B02G225100,TRAECS2B02G469800,TRAECS2D02G155300,TRAECS2D02G172100,TRAECS2D02G196900,TRAECS2D02G338000,TRAECS2D02G426700,TRAECS2D02G565300,TRAECS3A02G077600,TRAECS3A02G111500,TRAECS3A02G154400,TRAECS3A02G186200,TRAECS3A02G195200,TRAECS3A02G402600,TRAECS3A02G421500,TRAECS3B02G281200,TRAECS3B02G376000,TRAECS3B02G413800,TRAECS3B02G436000,TRAECS3B02G437300,TRAECS3D02G190200,TRAECS3D02G397600,TRAECS4A02G013700,TRAECS4A02G025100,TRAECS4A02G149300,TRAECS4A02G281500,TRAECS4B02G031500,TRAECS4D02G007200,TRAECS4D02G240600,TRAECS4D02G277500,TRAECS4D02G289100,TRAECS4D02G318400,TRAECS4D02G327300,TRAECS5A02G094300,TRAECS5A02G106300,TRAECS5A02G132300,TRAECS5A02G395200,TRAECS5A02G466100,TRAECS5A02G485700,TRAECS5A02G501900,TRAECS5B02G058400,TRAECS5B02G100400,TRAECS5B02G111300,TRAECS5B02G185200,TRAECS5B02G299900,TRAECS5B02G400000,TRAECS5B02G405700,TRAECS5D02G106600,TRAECS5D02G192100,TRAECS5D02G328300,TRAECS5D02G388600,TRAECS5D02G404500,TRAECS5D02G411100,TRAECS5D02G521800,TRAECS5D02G550700,TRAECS6A02G234000,TRAECS6A02G250800,TRAECS6A02G254900,TRAECS6B02G148100,TRAECS6B02G262800,TRAECS6B02G270600,TRAECS6B02G274600,TRAECS6B02G299700,TRAECS6B02G450700,TRAECS6D02G109700,TRAECS6D02G164900,TRAECS6D02G236200,TRAECS6D02G251100,TRAECS6D02G258000,TRAECS7A02G070100,TRAECS7A02G120300,TRAECS7A02G133500,TRAECS7A02G251400,TRAECS7A02G287400,TRAECS7A02G287400,TRAECS7B02G071400,TRAECS7B02G139700,TRAECS7B02G166900,TRAECS7B02G183300,TRAECS7B02G363200,TRAECS7D02G17800,TRAECS7D02G128900,TRAECS7D02G133100,TRAECS7D02G249500,TRAECS7D02G284900
GO:CC	cytoplasm	GO:0005737	2.01E-02	1.70	7523	393	114	32923	TRAECS1A02G043600,TRAECS1A02G137000,TRAECS1A02G233200,TRAECS1A02G304400,TRAECS1B02G155600,TRAECS1B02G233500,TRAECS1B02G237500,TRAECS1D02G044200,TRAECS1D02G030900,TRAECS1D02G329300,TRAECS1D02G383400,TRAECS2A02G270000,TRAECS2A02G297400,TRAECS2A02G330700,TRAECS2B02G191200,TRAECS2B02G409900,TRAECS2D02G172100,TRAECS2D02G269000,TRAECS2D02G295300,TRAECS3A02G195200,TRAECS3A02G292700,TRAECS3A02G402600,TRAECS3A02G414700,TRAECS3A02G421500,TRAECS3B02G327500,TRAECS3B02G436000,TRAECS3D02G397600,TRAECS4A02G025100,TRAECS4A02G050200,TRAECS4A02G225800,TRAECS4D02G256800,TRAECS4A02G281500,TRAECS4B02G031500,TRAECS4B02G042600,TRAECS4B02G090000,TRAECS4B02G254400,TRAECS4B02G268800,TRAECS4D02G007200,TRAECS4D02G11300,TRAECS4D02G029200,TRAECS4D02G086000,TRAECS4D02G240600,TRAECS4D02G254200,TRAECS4D02G277500,TRAECS5A02G106300,TRAECS5A02G295700,TRAECS5B02G111300,TRAECS5B02G299900,TRAECS5D02G048400,TRAECS5D02G238500,TRAECS5D02G242900,TRAECS5D02G388600,TRAECS5D02G403200,TRAECS5D02G521800,TRAECS6A02G075400,TRAECS6A02G117100,TRAECS6A02G154400,TRAECS6A02G418900,TRAECS6B02G100100,TRAECS6B02G148100,TRAECS6B02G201100,TRAECS6D02G0201800,TRAECS6D02G040500,TRAECS6D02G040600,TRAECS6D02G109700,TRAECS6D02G251100,TRAECS7A02G251400,TRAECS7B02G028500,TRAECS7B02G048700,TRAECS7B02G139700,TRAECS7D02G128900,TRAECS7D02G127400,TRAECS7D02G128900,TRAECS7D02G249500,TRAECS7D02G282200,TRAECS7D02G310800,TRAECS7D02G338400,TRAECS7D02G0268900
GO:CC	protein-containing complex	GO:0032991	2.70E-02	1.57	4863	393	78	32923	
GO:CC	nuclear exosome (RNase complex)	GO:0000176	2.71E-02	1.57	8	393	2	32923	TRAECS1A02G137000,TRAECS1B02G155600
GO:CC	nuclear speck	GO:0016607	2.71E-02	1.57	27	393	3	32923	TRAECS6A02G254900,TRAECS6B02G270600,TRAECS6D02G236200
GO:CC	methyltransferase complex	GO:0034708	2.71E-02	1.57	27	393	3	32923	TRAECS4A02G050200,TRAECS4B02G254400,TRAECS4D02G254200
GO:CC	nucleoplasm	GO:0005654	4.20E-02	1.38	489	393	13	32923	TRAECS1A02G043600,TRAECS1D02G044200,TRAECS2A02G270000,TRAECS2D02G269000,TRAECS4A02G050200,TRAECS4A02G256800,TRAECS4B02G042600,TRAECS4B02G254400,TRAECS4D02G254200,TRAECS5D02G403200,TRAECS6A02G254900,TRAECS6B02G270600,TRAECS6D02G236200



Supplementary Table 4.10. Papain family cysteine proteases (CysProt) genes and inhibitors of CysProt genes differentially expressed (DE) in E82 leaf tissue.

Gene name	Gene ID	Leaf WT vs E82	
		Log ₂ FC	FDR
CysProt inhibitor genes			
<i>Icy1-2</i>	<i>TraesCS1B02G322100</i>	-1.49	*
<i>Icy1-2</i>	<i>TraesCS1D02G310300</i>	-1.48	**
<i>Icy8 (HvCPI8)</i>	<i>TraesCS2A02G126100</i>	-2.33	***
<i>Icy8 (HvCPI8)</i>	<i>TraesCS2D02G128900</i>	-2.23	***
<i>Icy2 (HvCPI2)</i>	<i>TraesCS5D02G502200</i>	7.20	*
CysProt genes			
<i>RD19D</i>	<i>TraesCS2A02G187400</i>	4.37	***
Triticain beta (<i>HvPap7</i>)	<i>TraesCS2A02G545700</i>	-3.82	***
<i>Psy1-D1</i>	<i>TraesCS3D02G232500</i>	7.86	***
<i>HvPap14</i> ortholog	<i>TraesCS3A02G348500</i>	6.15	**
<i>HvPap1</i> ortholog	<i>TraesCS5A02G326100</i>	-1.03	*
NA	<i>TraesCS4A02G252800</i>	-1.26	*
NA	<i>TraesCS4B02G062500</i>	-1.35	*
<i>HvPap4</i> ortholog	<i>TraesCS5B02G138300</i>	7.15	*
<i>HvPap4</i> ortholog	<i>TraesCS5D02G152400</i>	7.84	**

NA: non-assigned.

FC: Fold-Change.

FDR: False Discovery Rate.

*, FDR < 0.05; **, FDR < 0.01; ***, FDR < 0.001.

Supplementary Table 4.11. Differentially expressed (DE) genes encoding serine-type proteases in E82 leaf against the wild type (WT).

Gene ID	Leaf WT vs E82	
	Log ₂ FC	FDR
<i>TraesCS1A02G178200</i>	-1.47	*
<i>TraesCS1A02G178300</i>	-1.33	*
<i>TraesCS1A02G187300</i>	6.75	***
<i>TraesCS1B02G110400</i>	-10.12	***
<i>TraesCS1B02G246100</i>	6.91	*
<i>TraesCS1B02G434100</i>	-2.97	***
<i>TraesCS1B02G472100</i>	1.35	*
<i>TraesCS2A02G398900</i>	5.76	*
<i>TraesCS2A02G401000</i>	7.64	**
<i>TraesCS3A02G181400</i>	5.85	*
<i>TraesCS3A02G182700</i>	6.89	**
<i>TraesCS3B02G211400</i>	6.90	***
<i>TraesCS3B02G212700</i>	8.18	**
<i>TraesCS3B02G337400</i>	-2.06	**
<i>TraesCS3B02G431900</i>	8.16	**



Gene ID	Leaf WT vs E82	
	Log ₂ FC	FDR
<i>TraesCS4B02G352100</i>	-8.37	**
<i>TraesCS4D02G076000</i>	-1.28	*
<i>TraesCS4D02G350300</i>	7.33	**
<i>TraesCS5A02G318800</i>	6.21	*
<i>TraesCS5A02G526000</i>	-7.63	*
<i>TraesCS5B02G404600</i>	9.58	***
<i>TraesCS6A02G161600</i>	-1.38	*
<i>TraesCS6A02G220500</i>	1.11	*
<i>TraesCS6B02G041200</i>	7.60	**
<i>TraesCS6D02G319900</i>	-3.18	***
<i>TraesCS7A02G124000</i>	3.82	*
<i>TraesCS7A02G510000</i>	7.42	*
<i>TraesCS7B02G128300</i>	9.54	**
<i>TraesCS7B02G168100</i>	-1.62	**
<i>TraesCS7D02G121200</i>	-1.40	*
<i>TraesCS7D02G470100</i>	8.04	***
<i>TraesCS1A02G178200</i>	-1.47	*
<i>TraesCS1A02G178300</i>	-1.33	*
<i>TraesCS1A02G187300</i>	6.75	***
<i>TraesCS1B02G110400</i>	-10.12	***

FC: Fold-Change.

FDR: False Discovery Rate.

*, FDR < 0.05; **, FDR < 0.01; ***, FDR < 0.001.

Supplementary Table 4.12. Differentially expressed (DE) genes encoding sugar and amino acid transporters in the leaf of E82.

Gene name	Gene ID	Leaf WT vs E82	
		Log ₂ FC	FDR
<i>SUT1A</i>	<i>TraesCS4A02G016400</i>	-1.67	***
<i>SUT1C</i>	<i>TraesCS4B02G287800</i>	-1.65	**
<i>SUT1D</i>	<i>TraesCS4D02G286500</i>	-1.60	**
<i>TaCWIN</i>	<i>TraesCS2B02G311900</i>	5.18	*
<i>TaCWIN</i>	<i>TraesCS2D02G293200</i>	5.20	*
<i>TaSWEET13b-6D</i>	<i>TraesCS6D02G367300</i>	-2.09	***
<i>TaSWEET13d-6B</i>	<i>TraesCS6B02G421700</i>	-2.57	***
<i>TaSWEET13e-6A</i>	<i>TraesCS6A02G382600</i>	-2.88	**
<i>TaSWEET13f-6D</i>	<i>TraesCS6D02G367400</i>	-2.87	***
<i>TaSWEET13g-6B</i>	<i>TraesCS6B02G421800</i>	-2.76	**
<i>TaSWEET13h-6B</i>	<i>TraesCS6B02G421500</i>	-2.94	**
<i>TaSWEET13i-6B</i>	<i>TraesCS6B02G421600</i>	-2.38	***
<i>TaSWEET15a-7D</i>	<i>TraesCS7D02G149000</i>	-10.77	***
<i>TaSWEET15b-7A</i>	<i>TraesCS7A02G147300</i>	-2.76	**
<i>TaSWEET15c-7B</i>	<i>TraesCS7B02G050500</i>	-9.18	***
<i>TaSWEET1a1-3B</i>	<i>TraesCS3B02G441800</i>	4.76	**
<i>TaSWEET4a-6A</i>	<i>TraesCS6A02G218800</i>	-1.26	*
<i>TaSWEET4b-6B</i>	<i>TraesCS6B02G248300</i>	-1.51	**



Gene name	Gene ID	Leaf WT vs E82	
		Log ₂ FC	FDR
<i>Ta</i> SWEET6b17-7D	<i>TraesCS7D02G160600</i>	2.58	*
<i>Ta</i> AAP1-1D	<i>TraesCS1D02G264700</i>	-1.83	*
<i>Ta</i> AAP12-4A	<i>TraesCS4A02G215300</i>	-1.81	**
<i>Ta</i> AAP12-4B	<i>TraesCS4B02G100800</i>	-1.51	*
<i>Ta</i> AAP14-5A	<i>TraesCS5A02G115100</i>	-1.64	*
<i>Ta</i> AAP14-5B	<i>TraesCS5B02G116100</i>	-1.73	*
<i>Ta</i> AAP14-5D	<i>TraesCS5D02G125700</i>	-2.00	**
<i>Ta</i> AAP15-5B	<i>TraesCS5B02G120800</i>	-2.66	**
<i>Ta</i> AAP15-Un	<i>TraesCSU02G134900</i>	-2.77	**
<i>Ta</i> AAP17-6D	<i>TraesCS6D02G265800</i>	-2.47	*
<i>Ta</i> AAP2-2A	<i>TraesCS2A02G268200</i>	-2.73	**
<i>Ta</i> AAP22-7A	<i>TraesCS7A02G356639</i>	6.46	**
<i>Ta</i> AAP3-2D	<i>TraesCS2D02G331800</i>	-2.08	*
<i>Ta</i> AAP6-2A	<i>TraesCS2A02G499800</i>	9.02	***
<i>Ta</i> AAP8-3A	<i>TraesCS3A02G388000</i>	8.11	*
<i>Ta</i> AAP9-3A	<i>TraesCS3A02G388100</i>	-2.85	*
<i>Ta</i> AAP9-3B	<i>TraesCS3B02G420600</i>	-3.32	***
<i>Ta</i> AAP9-3D	<i>TraesCS3D02G381400</i>	-2.76	***
<i>Ta</i> ANT6-6D	<i>TraesCS6D02G235500</i>	10.55	***
<i>Ta</i> ATLa1-3D	<i>TraesCS3D02G340400</i>	-4.26	**
<i>Ta</i> ATLa6-7A	<i>TraesCS7A02G517100</i>	6.12	**
<i>Ta</i> ATLb2-2A	<i>TraesCS2A02G052500</i>	8.22	**
<i>Ta</i> ATLb3-2D	<i>TraesCS2D02G280800</i>	-2.42	*
<i>Ta</i> AUX2-1B	<i>TraesCS1B02G287300</i>	6.51	*
<i>Ta</i> AUX3-3A	<i>TraesCS3A02G369200</i>	-1.51	*
<i>Ta</i> AUX3-3B	<i>TraesCS3B02G401000</i>	-2.00	**
<i>Ta</i> BAT4-3A	<i>TraesCS3A02G484600</i>	-11.34	***
<i>Ta</i> CAT5-4B	<i>TraesCS4B02G013800</i>	9.07	***
<i>Ta</i> LHT3-2D	<i>TraesCS2D02G402900</i>	8.99	***
<i>Ta</i> LHT5-5A	<i>TraesCS5A02G089400</i>	9.68	***
<i>Ta</i> LHT5-5B	<i>TraesCS5B02G095400</i>	9.09	***
<i>Ta</i> ProT1-2B	<i>TraesCS2B02G287200</i>	1.45	*
<i>Ta</i> ProT2-3A	<i>TraesCS3A02G414300</i>	-3.29	**
<i>Ta</i> ProT2-3B	<i>TraesCS3B02G449100</i>	-4.17	***
<i>Ta</i> ProT2-3D	<i>TraesCS3D02G408800</i>	-3.76	***
<i>Ta</i> ProT3-4A	<i>TraesCS4A02G179500</i>	-2.01	***
<i>Ta</i> TTP3-4A	<i>TraesCS4A02G083100</i>	2.73	***

FC: Fold-Change.

FDR: False Discovery Rate.

ns: non-significant; *, FDR < 0.05; **, FDR < 0.01; ***, FDR < 0.001.





CHAPTER V

Unraveling the celiac disease-related immunogenic complexes in a set of wheat and tritordeum genotypes: implications for low-gluten precision breeding in cereal crops

Under review in *Frontiers in Plant Science*, submitted in 2023.



Abstract

The development of low-gluten immunogenic cereal varieties is a suitable way to fight the increment of pathologies associated with the consumption of cereals. Although RNAi and CRISPR/Cas technologies were effective in providing low-gluten wheat, the regulatory framework, particularly in the European Union, is an obstacle to the short- or medium-term implementation of such lines. In the present work, we carried out a high throughput amplicon sequencing of two highly immunogenic complexes of wheat gliadins in a set of bread and durum wheat, and tritordeum genotypes. Bread wheat lines harboring the 1BL/1RS translocation were included in the analysis and their amplicons successfully identified. The number of celiac disease (CD) epitopes and their abundances were determined in the alpha- and gamma-gliadin amplicons, including 40k- γ -secalin ones. Bread wheat genotypes not containing the 1BL/1RS translocation showed a higher average number of both alpha- and gamma-gliadin epitopes than those containing such translocation. Interestingly, alpha-gliadin amplicons not containing CD epitopes accounted for the highest abundance (around 53%), and the alpha- and gamma-gliadin amplicons with the highest number of epitopes were present in the D-subgenome. The durum wheat and tritordeum lines showed the lowest number of alpha- and gamma-gliadin CD epitopes. Our results allow progress in unraveling the immunogenic complexes of alpha- and gamma-gliadins and can contribute to the development of low-immunogenic varieties within precision breeding programs, by crossing or by CRISPR/Cas gene editing.

Keywords: amplicon, alpha-gliadins, gamma-gliadins, precision breeding, wheat, CRISPR/Cas



Resumen

El desarrollo de variedades de cereales con bajo contenido en gluten inmunogénico es una vía para combatir el incremento de patologías asociadas al consumo de cereales. Si bien las tecnologías ARNi y CRISPR/Cas fueron efectivas para proporcionar trigo bajo en gluten, el marco regulatorio, particularmente en la Unión Europea, es un obstáculo para la implementación a corto o medio plazo de dichas líneas. En el presente trabajo, llevamos a cabo la secuenciación de alto rendimiento de amplicones de dos complejos altamente inmunogénicos de gliadinas en un conjunto de genotipos de trigo harinero y duro, y tritórdeo. Las líneas de trigo harinero que albergaban la translocación 1BL/1RS se incluyeron en el análisis, y sus amplicones se identificaron con éxito. Se determinó el número de epítomos relacionados con la enfermedad celíaca (EC) y sus abundancias en los amplicones de alfa- y gamma-gliadinas, incluidos los de 40k- γ -secalinas. Los genotipos de trigo harinero que no contenían la translocación 1BL/1RS mostraron un número promedio más alto de epítomos de alfa- y gamma-gliadinas que aquellos que contenían dicha translocación. Curiosamente, los amplicones de alfa-gliadinas que no contenían epítomos de la EC representaron la mayor abundancia (alrededor del 53%), y los amplicones con el mayor número de epítomos estaban presentes en el subgenoma D. Las líneas de trigo duro y tritórdeo mostraron el número más bajo de epítomos de la EC de alfa- y gamma-gliadinas. Nuestros resultados permiten avanzar en el desentrañamiento de los complejos inmunogénicos de las gliadinas y pueden contribuir al desarrollo de variedades de baja inmunogenicidad dentro de los programas de mejora de precisión, mediante cruce o edición de genes por CRISPR/Cas.

Palabras clave: amplicón, alfa-gliadinas, gamma-gliadinas, mejora de precisión, trigo, CRISPR/Cas



1. Introduction

Bread wheat (*Triticum aestivum* L.) is the dominant crop in temperate countries widely used in the human diet and livestock feed. It is an essential source of nutrients and other beneficial compounds. World production of this crop reaches 760 million tons per year, representing 30% of all harvested cereals (FAO, 2022). Wheat grains contain around 9-15% of protein, and derived wheat flour is processed into a great variety of food products that humans consume daily, such as bread, cakes, noodles, biscuits, etc. Grain proteins can be classified according to their extractability in a series of solvents: albumins (soluble in water), globulins (soluble in salt), gliadins (soluble in alcohol), and glutelins (insoluble in other solvents and extracted in alkali). In wheat, glutelins are mainly composed of glutenins which are, in fact, prolamin subunits not extractable with alcohol/water solvents due their polymeric nature by inter-chain disulphide bonds (Shewry, 2019). Both groups, gliadins and glutenins, comprising gluten, and they are also divided into different fractions; the first one into three structural groups: omega-, alpha-, and gamma-gliadins; and the second one into two types of subunits: the High Molecular Weight (HMW), and the Low Molecular Weight (LMW) glutenin subunits (Shewry, 2009). Despite the ability to provide breadmaking properties, these proteins present epitopes that can trigger the immune response in patients suffering from celiac disease (CD), an autoimmune enteropathy provoked by the presence of certain gluten peptides which are deamidated by the tissue transglutaminase 2 (tTG2) in the human gut. These peptides are presented by the human leukocyte antigen (HLA)-DQ2 and DQ8 haplotypes to the CD4 T cells, triggering the immunogenic response in the small intestine and leading to a cascade of symptoms like intestine inflammation, the atrophy of intestinal villus, malabsorption of nutrients, among others (Sollid, 2002; Ludvigsson et al., 2013).

However, not all the HLA-DQ molecules are linked to the same level of risk. The HLA-DQ2.5 molecule binds to more gluten peptides compared to HLA-DQ2.2 and forms stable complexes with the gluten T-cell epitopes (Vader et al., 2003; Fallang et al., 2009). Over the years, many HLA-DQ2.5 epitopes have been characterized in gluten proteins; showing that they are mainly present in the gliadin fraction of gluten (Sollid et al., 2012; 2020). Particularly, the alpha-gliadins of wheat contain the 33-mer peptide, one of the most immunogenic peptides known so far. This peptide contains six overlapping copies of three epitopes and shows high stimulatory properties in relation to CD (Shan et al., 2002). In addition, the alpha-gliadins of wheat also present the p31-43 peptide, which is associated with the innate immune response in CD non-T cell-dependent (Maiuri et al., 2003). In the wheat gamma-gliadins, it was identified another highly immunogenic peptide denoted as 26-mer, which is also highly resistant to intestinal brush border membrane proteolysis and contains many overlapping epitopes (Shan et al., 2005).

In wheat, the alpha- and gamma-gliadin proteins are encoded by two multigene families in the *Gli-2* and *Gli-1* loci located, respectively, in the short arm of the



homoeologous chromosome 6 (Huo et al., 2018b) and 1 (Altenbach et al., 2020; Anderson et al., 2012). Both loci contain a large number of duplicated genes arranged in tandem that, in addition to the wheat polyploid nature, the high content of repetitive DNA and the large size of the genome, make it difficult to obtain high-quality sequences for these gene families. In fact, for the alpha-gliadins between 25 to 150 copies in wheat cultivars and ancestral species were proposed (Harberd et al., 1985). Recently, a droplet digital PCR (ddPCR) assay revealed the copy number variation of different wheat lines, obtaining 70-76 copies of alpha-gliadins in tetraploid wheat lines (AABB), and 86-95 copies in synthetic hexaploid wheat lines (AABBDD) (Jouanin et al., 2020).

There is a high variation of the immunogenicity among wheat varieties (Spaenij-Dekking et al., 2005), which is tightly linked to the variation of CD-immunogenic epitopes present in the gluten proteins. The characterization of the epitope content and the identification of lines with low immunogenic potential is the starting point in precision breeding programs and would facilitate the development of non-genetically modified low-gluten wheat lines. In addition, these CD low-epitope lines could be the targets for genomic editing as their reduced structural complexity will facilitate the use of Cas enzymes to develop low-immunogenic lines.

In previous work, we analyzed di-, tetra-, and hexaploid wheat accessions by amplicon-based Next-Generation Sequencing (NGS) to study the origin and evolution of different types of alpha-gliadins (Ozuna et al., 2015). This amplicon expands the p31-43 and 33-mer, allowing the identification of both peptides along with the DQ2.5_glia_α3, which overlaps partially with the 33-mer peptide. This approach provided an excellent tool for addressing the structural complexity of the immunogenic alpha-gliadins, and for the selection of wheat varieties with very low CD immunogenic epitopes. In this work, we report the characterization of the gamma-gliadin immunogenic complex by NGS amplicon sequencing, and the utility of both alpha- and gamma-gliadins amplicons to study a wide range of cereal lines comprising bread, durum wheat, and the amphiploid tritordeum (AABBH^{ch}H^{ch}), a fertile cross between the wild barley *Hordeum chilense* and durum wheat, which showed lower immunogenic properties than bread wheat (Haro et al., 2022). Based on their alpha- and gamma-gliadins CD epitopes abundance, we can anticipate their immunogenic potential concerning gluten-related disorders (GRDs). More interestingly, we can identify wheat varieties containing rye translocations and their associated epitopes. Results reported in this work are important not only for selecting potential low-immunogenic cereals but also to select wheat varieties with low gliadin complexity, which can be used as targets for genome editing by CRISPR/Cas (Sánchez-León et al., 2018).



2. Materials and methods

2.1. Plant material

A total of 44 lines were analyzed comprising 37 bread wheat lines, two lines of durum wheat, and five lines of tritordeum (Supplementary Table 5.1). Among the bread wheat lines, 14 of them presented rye translocation (1BL/1RS translocation) as denoted by the presence of the allele *Gli-B1l* in the mature grains gliadin proteins profiles by A-PAGE (Metakovsky et al., 2018) (Supplementary Figures 5.1 and 5.2). The protocol followed for A-PAGE gels of total gliadin proteins is described in Gil-Humanes et al. (2012). For the lines BIOINTA2004 and INIACondor, the allele A-PAGE protein profile was not as clear as the other bread wheat lines with rye translocation, but both presented amplicons annotated as 40k- γ -secalins from genes located in the 1RS chromosome of rye (Qi et al., 2013). Per duplicate, all the plants were grown in a greenhouse and the leaves were collected in the first stages of the plant and frozen at -80 °C.

2.2. DNA extraction, PCR and NGS sequencing

The total DNA was extracted for each leaf sample following the CTAB protocol (Murray and Thompson, 1980). The DNA concentration was determined by NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA, United States). The primers aGli900F1 (5'-GTTAGAGTTCCAGTGCCACAA-3') as forward and 33mer1R2-Ok (5'-GGTTGTTGTGGTTGCGRATA-3') as reverse, and γ gliF3 (5'-GCCAATATRCAGGTCGACCC-3') as forward and γ gliR3 (5'-GGGTTCAWCTGTTGTTGTAG-3') as reverse were used to amplify the alpha and gamma-gliadin amplicons, respectively. The primers of the alpha-gliadin amplicon were previously designed by Ozuna et al. (2015) in a conserved region of the genes coding the first repetitive domain of the alpha-gliadin proteins which comprises the p31-43 region associated with the innate response in CD and the 33-mer peptide (Figure 5.1A). The primers for the gamma-gliadin amplicon were designed in the present work to amplify the gene sequence coding part of the N-terminal and the first repetitive domain of gamma-gliadin proteins comprising a high number of CD epitopes, including the 26-mer peptide (Figure 5.1B). These primers also amplified the 40k- γ -secalins (located in the short arm of chromosome 1R) which have high homology with gamma-gliadins of wheat (Qi et al., 2013). The amplicon sequencing was carried out by Fundación Parque Científico de Madrid (Cantoblanco, Madrid) by the MiSeq system (<https://www.illumina.com>) producing paired-end reads of 2x280 bp. The amplicon lengths were checked by the Agilent 2100 Bioanalyzer system (Agilent Technologies, Santa Clara, CA 95051).



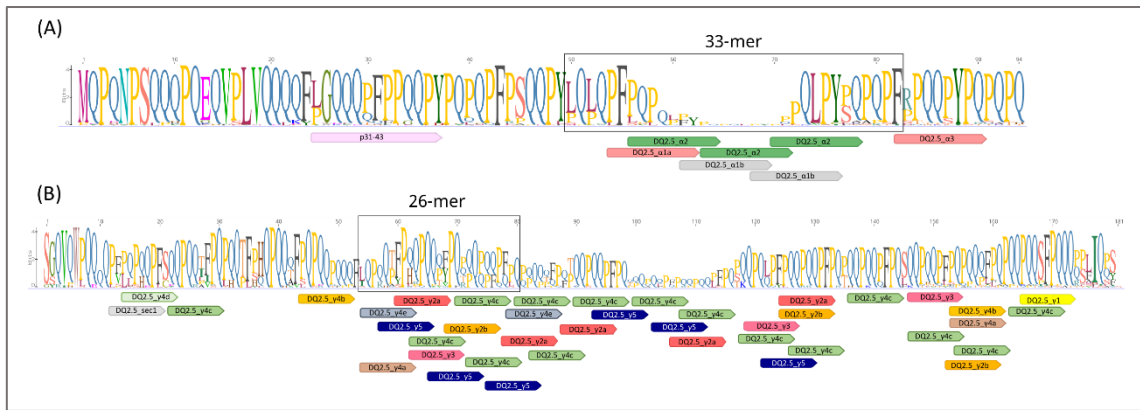


Figure 5.1. Schematic diagram of amplicon-encoded proteins of the (A) alpha and (B) gamma gliadins. The amplicons are located in the first repetitive domain of both gliadin proteins. Mapping of different DQ2.5 celiac disease (CD) epitopes were shown using the Geneious v2019.0.3 software (<https://www.geneious.com>). The consensus sequences of alpha- and gamma-gliadins from BW208 clones, the NCBI, and RefSeq v1.1 reference genome annotation of bread wheat were used for the representation.

2.3. Amplicon analysis

A GNU/Linux Ubuntu v 18.04 (<https://ubuntu.com>) server with 64 cores and 128 GB of RAM was used for sequences analysis. For the amplicon, the raw reads were analyzed by Usearch v9.2.64 (Edgar, 2010). Briefly, the raw paired-end reads were merged (-fastq_mergepairs), filtered (-fastq_filter), de-replicated (-fastx_uniques), and denoised (-unoise2) to obtain the unique denoised amplicons for all the lines (Figure 5.2). From now on, the unique denoised amplicons will be referred to as ‘amplicons’ for ease of reading. This protocol was run independently for alpha- and gamma-gliadins/secalins.

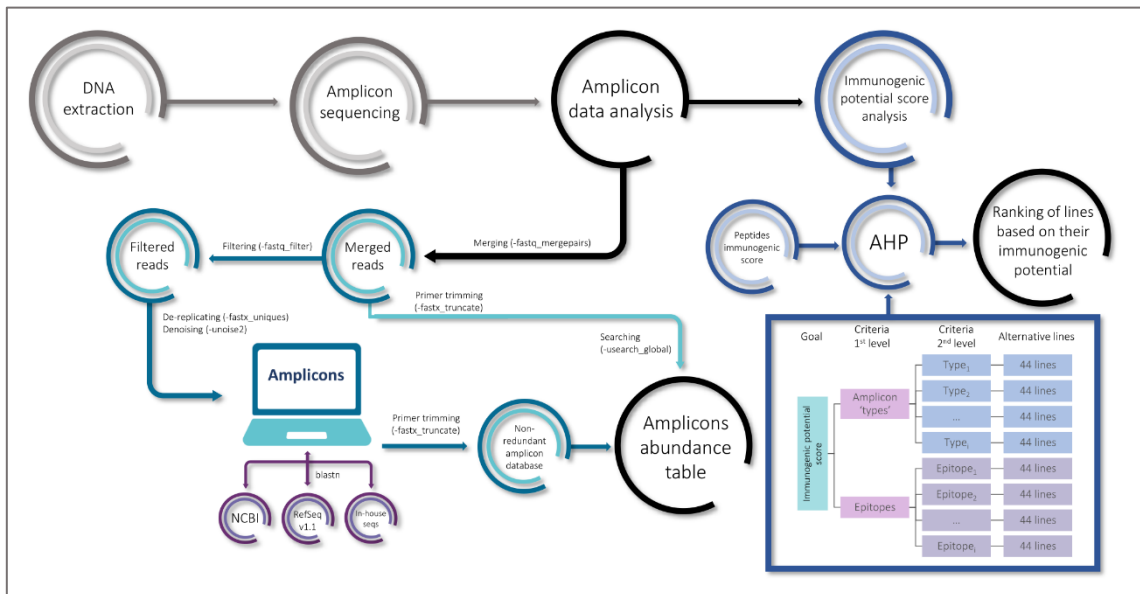


Figure 5.2. The pipeline of the amplicon analysis.



A non-redundant amplicon database was constructed for alpha- and gamma-gliadins separately. In addition, alpha- and gamma-gliadin gene sequences obtained from the National Center for Biotechnology Information (NCBI), the RefSeq v1.1 bread wheat reference genome annotation (Appels et al., 2018), and in-house sequences of the wheat variety BW208 clones by Sanger sequencing were included. The primers were aligned to these sequences to extract the corresponding amplicon sequence and, subsequently, those primers were removed from all the database entries by Usearch (-fastx_truncate) (Figure 5.2). For the gamma-gliadin amplicon analysis, seven 40k- γ -secalin sequences were included from the NCBI. The amplicons sequences obtained by Usearch were aligned in the NCBI database, the RefSeq v1.1, and the BW208 clones' sequences by blastn (Johnson et al., 2008) to avoid redundancy. Unique sequences were kept for the alpha- and gamma-gliadin/secalin amplicon databases.

After primer trimming, the alpha- and gamma-gliadin/secalin reads for each line were searched in both databases respectively with an identity score of 0.99 (-usearch_global -id 0.99) to calculate the abundance of the amplicons. The amplicons with less than 0.25% of abundance in at least four samples were discarded for the analysis. The abundance of each amplicon was normalized by the total number of reads of the amplicons present in each line. The mean of the two biological replicates per line was calculated for further analyzes.

2.3.1. Epitope search and amplicon types

For analyzing the potential immunogenicity of wheat and tritordeum lines based on the NGS data, the amplicons were translated to peptides for searching the non-deamidated CD epitopes in their sequence with Python custom scripts (<https://github.com/MiriamMarinS/wheatAHP>). As reported in Sollid et al., (2012), the DQ2.5_glia_ ω 1, DQ2.5_hor_1, and DQ2.5_sec1 share the same epitope sequence; like the DQ2.5_hor_2 and DQ2.5_sec2. Therefore, they were denoted as 'sec1' and 'sec2' (Supplementary Table 5.2). In addition, 1-mismatch epitopes were considered and characterized in the alpha-gliadin amplicon analysis. The abundance of CD epitopes was calculated for each line according to the abundance of the amplicons containing them. We have also classified the alpha- and gamma-gliadin amplicons into 'types' considering the total number of DQ epitopes present in them, e.g. the Alpha_0 'type' comprises all the alpha-gliadin amplicons with no epitopes, while the Alpha_7 amplicons contain 7 epitopes. For the alpha-gliadin amplicon, the p31-43 peptide was analyzed independently, and it was not included in the amplicons 'types' classification. The abundance of amplicon 'types' was also calculated as described previously.



2.4. Ranking of cereal lines' immunogenic potential by Analytic Hierarchy Process (AHP)

In order to study the immunogenic potential of the lines assayed in the present work, we got a synthetic value (score) to rank the lines crossing their epitope abundance with a previous immunogenicity screening of prolamin peptides encompassing CD epitopes (Tye-Din et al., 2010). We searched for the known CD epitopes in the list of oligopeptides and crossed their scores with our values of epitope abundance in each line based on the AHP method (Saaty, 2008) implemented with AHPy library. We added the abundance data from the amplicons 'types' to the process, getting a scheme with one criterion of first level and two criteria of second level (Figure 5.2). The complete methods were described on <https://github.com/MiriamMarinS/wheatAHP>. To calculate the final score for the alternative lines (Figure 5.2), the $\log_2(\text{FC})$ of the abundance was calculated for each pair-wise comparison between each pair of lines, and this measure was transformed into intensity values according to the AHP method. In the first criterion comprising the epitopes vs amplicons 'types' comparison, we gave equal or more weight to the first one in order to highlight the epitope sequence rather than the number of copies across the amplicon. Once the intensity values favoring the amplicons 'types' were removed, the remaining nine possible values were used for this criterion, giving nine immunogenic potential scores for each line. From this, the mean and the standard error were used in the ranking representation. The source code is available at <https://github.com/MiriamMarinS/wheatAHP>.

3. Results

A summary of the alpha and gamma-gliadin amplicons is presented in Table 5.1. The abundance of each amplicon for each line is provided in Supplementary Tables 5.3 and 5.4. As shown, for the set of genotypes analyzed, we have found a higher number of different amplicons for the alpha- than for the gamma-gliadins. Stop codons were found in 21 of 76 alpha-gliadin amplicons and in 9 of 41 gamma-gliadin ones (Table 5.1). All the epitopes listed in Sollid et al., (2020) were searched in the translated amplicon sequences, and 17 non-deamidated ones were found. The number of CD epitopes found in the same amplicon varied between 0 and 7 for the alpha-gliadins, and between 0 and 14 for the gamma-gliadins/secalins (Table 5.1). Interestingly, 39% of the alpha-gliadin amplicons do not contain CD epitopes while only 15% of the gamma-gliadin/secalin ones were free of them. The total number of amplicons per species, and their distribution in putative genes and pseudogenes are in Supplementary Table 5.5. It is remarkable that bread wheat genotypes not containing the 1BL/1RS translocation showed a higher average number of both alpha- and gamma-gliadin epitopes than those containing such translocation.



Table 5.1. General description of the alpha and gamma-gliadin amplicons in the set of wheat and tritordeum genotypes used in this study.

Features	Alpha-gliadin	Gamma-gliadin
Total number of different amplicons	76	41 ^a
Putative genes. Amplicons with no-stop codons	55	32 ^b
Pseudogenes. Amplicons with stop codons	21	9
Range of amplicon length (bp)	201 - 279	216 - 507
Range of CD epitopes per amplicon. Only putative genes	0 - 7	0 - 14
Number of amplicons not containing CD epitopes (putative genes)	31 (25)	6 (3) ^c

^a 7 of them are secalins

^b 4 of them are secalins

^c all of them are secalins

3.1. Epitope abundance in the set of cereal genotypes

Next, the abundance of the CD epitopes in the three species was calculated considering the abundance of the amplicons which carry them in each genotype (Supplementary Tables 5.6 and 5.7). The list of epitopes/peptides found comprises the p31-43 and four DQ2.5 epitopes for the alpha-gliadin amplicons; and 12 DQ2.5 epitopes, including two secalin ones, for the gamma-gliadin/secalin amplicons. Although the bread wheat lines showed the highest abundance of epitopes, they also had wider heterogeneity on their abundance profiles than the other cereals (Figure 5.3). In contrast, durum wheat lines had the lowest number of epitope matches.

In the alpha-gliadin amplicons, there were notable differences between the three cereal species for the abundance of the CD immunogenic epitopes. In the bread wheat lines, the DQ2.5_glia_α1a was the highest abundant among the three epitopes included in the 33-mer peptide (Figure 5.3A) while DQ2.5_glia_α1b was the less one. If bread wheat lines with and without the rye translocation were compared, the significant effects in the alpha-gliadin epitopes were shown, involving a less abundance of DQ2.5_glia_α3 epitope and p31-43 peptide (Figure 5.3A). In contrast to bread wheat, durum wheat and tritordeum lines had null or closer to zero abundance of the DQ2.5_glia_α1b and DQ2.5_glia_α2 epitopes (Figure 5.3A).

In the case of the gamma-gliadin/secalin amplicons, the DQ2.5_glia_γ4c epitope, included in the 26-mer peptide, stood up in all three cereal species (Figure 5.3B). In addition, the DQ2.5_glia_γ1, DQ2.5_glia_γ2b, and DQ2.5_glia_γ3 epitopes had also high abundance in all the cereal crops studied. The last two epitopes had the same abundance because they were present in the same gamma-gliadin amplicons (Supplementary Table 5.4). The rest of the epitopes were poorly represented in durum wheat and tritordeum lines, while for bread wheat, other epitopes such as



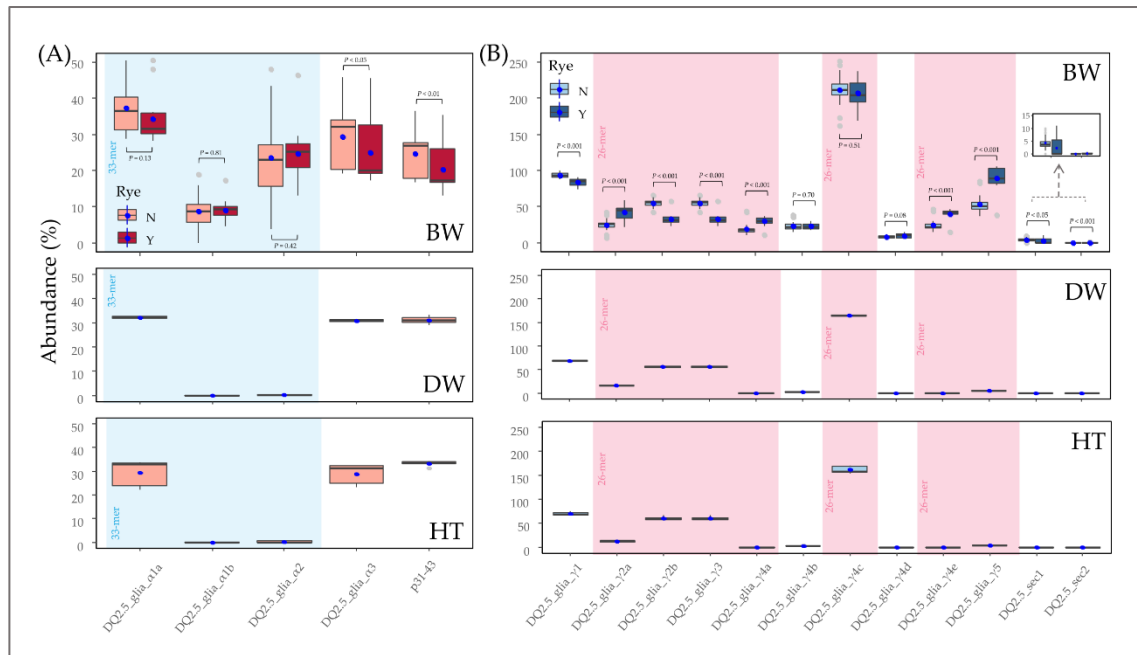


Figure 5.3. The abundance of CD epitopes in the (A) alpha- and (B) gamma-gliadin amplicons of putative genes. For bread wheat lines, the presence (Y) or not (N) of the rye translocation is indicated. The blue dots indicate the mean, the grey dots indicate the outliers, and the black horizontal line represents the median value. The statistical analysis for each epitope and the immunogenic peptide between the lines containing (Y) or not (N) the rye translocation was performed by the non-parametric Mann-Whitney-Wilcoxon test, and by the t-test for the DQ2.5_glia_α1b, DQ2.5_glia_γ1, DQ2.5_glia_γ2a, DQ2.5_glia_γ4c and DQ2.5_glia_γ4d epitopes that were normally distributed across the lines. The epitopes contained in the 33-mer and 26-mer regions are shaded. BW: bread wheat; DW: durum wheat; HT: tritordeum.

the DQ2.5_glia_γ5 included in the 26-mer peptide, had also a high abundance (Figure 5.3B). Comparing the abundance of bread wheat lines with and without the rye translocation, the most abundant epitope DQ2.5_glia_γ4c was equally represented in both. Interestingly, the epitopes DQ2.5_sec1 can also be found in the gamma-gliadins amplicons (Figure 5.1B), in fact, there were no differences in the abundance of this epitope between bread wheat with and without the rye translocation (Figure 5.3B). On the other hand, the abundance of the DQ2.5_sec2 epitope was significantly higher in the lines containing the rye translocation (Figure 5.3B), since this epitope was only present in one 40-k-γ-secalin (Supplementary Table 5.4), Moreover, epitopes DQ2.5_glia_γ4b and DQ2.5_glia_γ4c were also present in rye secalins.

3.2. Gliadin amplicons are grouped into different types depending on the number of CD epitopes

The 33-mer is one of the most immunogenic peptides described for the alpha-gliadins, and it is related to the presence of six overlapping copies of three CD epitopes (Shan et al., 2002). However, the 33-mer region showed a high sequence



heterogeneity in the alpha-gliadin amplicons and not all of them presented the full 33-mer with the six epitopes (Figure 5.1A). This is clear for the durum wheat and tritordeum lines, which only contained 2 of the 3 epitopes of the complete 33-mer. Similarly, the gamma-gliadin amplicons presented regions that comprise many overlapping epitopes with a high sequence variability (Figure 5.1B). The total amplicons for both the alpha and gamma-gliadins/secalins differed in the number of CD epitopes in their peptide sequences, comprising a range of 0 - 7 and 0 - 14 epitopes respectively (Table 5.1). Based on these differences, we were able to group the alpha and gamma-gliadin/secalin amplicons into 7 and 14 'types', respectively, named accordingly as Alpha_0 to Alpha_7 and Gamma_0 to Gamma_14 (Figures 5.4 and 5.5). In addition, their abundance was calculated for each genotype and summarized in Supplementary Tables 5.8 and 5.9.

As shown, the Alpha_0 - not containing CD epitopes - stood out as the alpha-gliadin amplicon 'type' with the highest overall abundance (53.4%), being comprised mainly of putative genes, followed by the Alpha_2 and Alpha_1 'types' (Figure 5.4A). On the other hand, only two alpha-gliadin amplicons contained seven CD epitopes (Alpha_7) with an abundance of 1.1% in all the lines (Figure 5.4A). The Alpha_7 amplicons presented the six overlapping epitopes comprising the 33-mer and the DQ2.5_glia_α3 downstream epitope (Figure 5.4E), showing the same epitope distribution represented in Figure 5.1A. Next, we assigned each amplicon to the A, B, and D-subgenomes by finding specific motifs in their translated peptide sequence as described in van Herpen et al. (2006). In addition, the Chinese Spring (CS) alpha-gliadin genes (Huo et al., 2018b) were also included in the phylogenetic tree as references. However, not all the sequences in a cluster were assigned to the same subgenome, i.e. there was no homogeneity in the subgenome assignment despite the sequence similarity of the elements in the cluster (Figure 5.4B). The most abundant alpha-gliadin amplicon in almost all the lines from the three genotypes belonged to the B-subgenome (Figure 5.4B,D). This amplicon, and most of the B-subgenome ones - except three of them which contain one epitope - were classified as Alpha_0 'type'. In addition, only one B-subgenome amplicon had the p31-43 peptide, but this peptide was mainly overrepresented in the A-subgenome amplicons (Figure 5.4B). On the other hand, amplicons with a high number of epitopes, including the Alpha_7, belong to the D-subgenome (Figure 5.4B). The lines OlaetaArtillero and ACA321 showed the highest abundance for the Alpha_7 'type', being both bread wheat lines without the rye translocation (Figure 5.4F, Supplementary Table 5.8). Interestingly, the abundance of the Alpha_3 to Alpha_7 'types' was very low or even absent for the durum wheat and tritordeum lines. In contrast, the Alpha_2 stood out for these two crop species and comprised amplicons harboring the DQ2.5_glia_α1a and DQ2.5_glia_α3 epitopes (Figure 5.4F,E). About 27% of the alpha-gliadin amplicons presented premature stop codons and were considered pseudogenes (Table 5.1). Some of them proceed from the mutation of C → T in the glutamine codons (Figure 5.4C). These pseudogene amplicons were mainly presented in the A-subgenome as well, followed by the B-subgenome (Figure 5.4B).



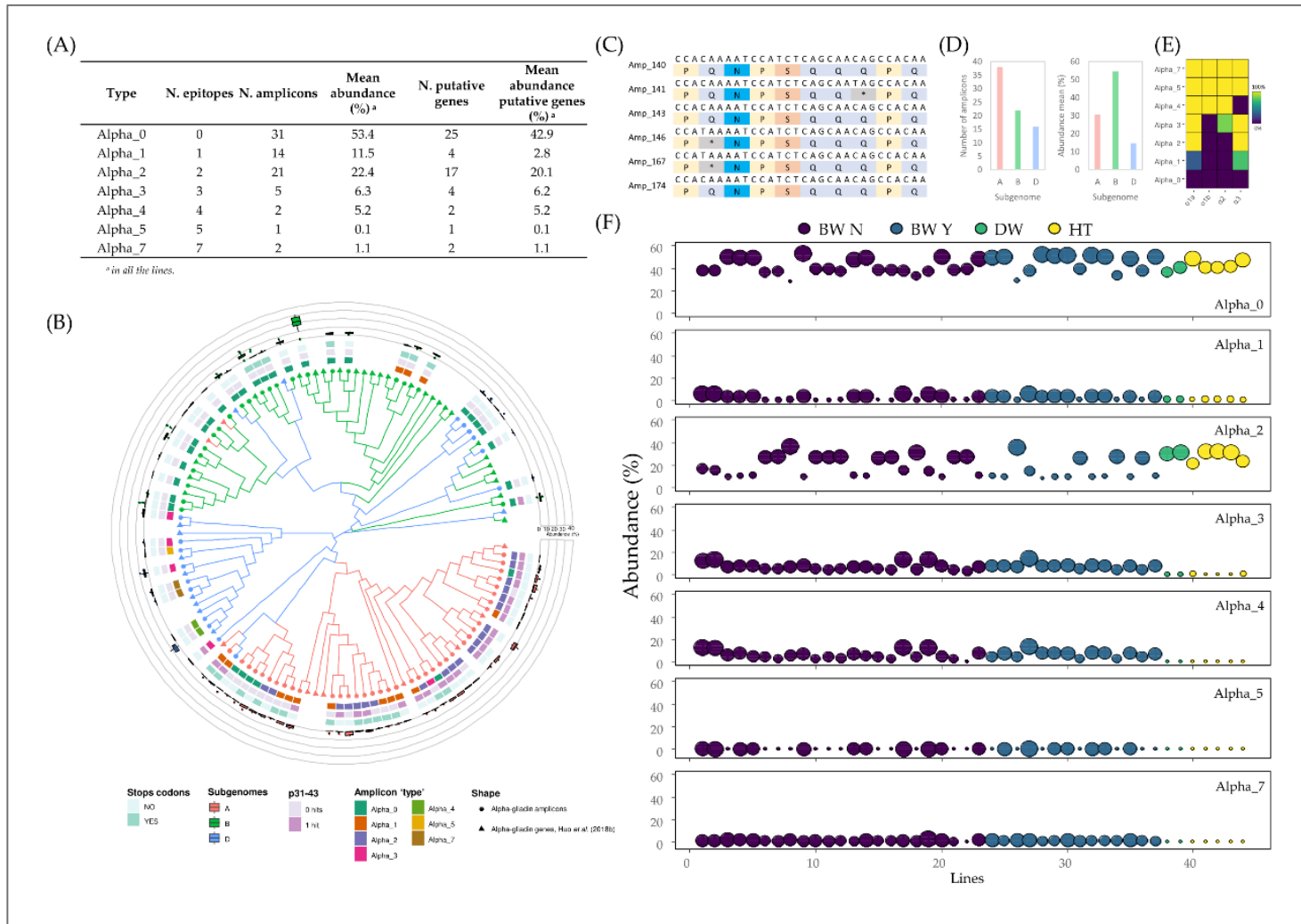


Figure 5.4. Characterization of the alpha-gliadin's amplicons 'types'. (A) The abundance of each alpha-gliadin 'type'. (B) Phylogenetic tree using the maximum-likelihood method in MEGA 10.1.7 from a MUSCLE sequence alignment (default parameters). From outside to inside: (i) the abundance for all the lines of each amplicon 'type', (ii) presence of premature stop codons, (iii) hits of the p31-43 peptide, and (iv) the number of DQ2.5 epitopes in their translated peptide sequence. (C) Stop codons generated by the C → T mutation. (D) Number and mean abundance (%) per subgenome. (E) Frequency with which epitope appears in each amplicon 'type': the number of amplicons containing each epitope divided by the total amplicons classified in this 'type'. (F) The abundance of each alpha-gliadin's amplicon 'type' for each wheat and tritordeum line based on putative genes. BW N: bread wheat without rye translocation; BW Y: bread wheat with rye translocation; DW: durum wheat; HT: tritordeum. The dot sizes ranged from the lowest (smallest) to the highest (biggest) abundance of all lines within each amplicon 'type'.



Regarding the gamma-gliadins/secalins analysis, there were only six amplicons without CD epitopes with a low mean abundance of 1.47 %, being only present in bread wheat lines with rye translocation as these amplicons were the 40k- γ -secalins (Figure 5.5A). However, these lines presented also the Gamma_14 'type', which contained the highest number of epitopes matches in the set of amplicons 'types'. This 'type' consisted of five DQ2.5 gamma-gliadin epitopes repeated across the amplicon sequences and had a very low abundance (Figure 5.5B,D; Supplementary Table 5.9). The most abundant gamma-gliadin 'types' were Gamma_5 and Gamma_7, with a high variability of DQ2.5 epitopes combinations in their amplicon sequences, and comprising 9 and 6 putative genes respectively (Figure 5.5A,D). The first one was found in all the lines, including bread and durum wheat, and tritordeum lines, however, their abundances were slightly lower for many of the bread wheat lines with the rye translocation (Figure 5.5E). In the case of Gamma_7, it was only present in bread wheat lines, except for the tritordeum line Aucan, being proINTAGaicho – not containing the rye translocation – the line with the highest abundance for this gamma-gliadin amplicon 'type' (Supplementary Table 5.9; Figure 5.5E). As for the alpha-gliadins, the gamma-gliadin amplicons were assigned to subgenomes, using the CS genes as references. The most abundant amplicons were from the D-subgenome, followed by B and A ones (Figure 5.5B,C). Interestingly, the Gamma_14 was only present in the D-subgenome, and the Gamma_0 was found only in the secalin amplicons, as stated before (Figure 5.5B). The pseudogenes were found in all three subgenomes, with a slightly higher proportion in the B one and secalin amplicons (Figure 5.5B). The Gamma_10, Gamma_13 and Gamma_4, only found in the B-subgenome, were absent in bread wheat lines with rye translocation. As an exception, the Gamma_13 presented only a low abundance in BIONTA2004, which had not the complete *Gli-B11* allele in the A-PAGE profile (Supplementary Table 5.9; Supplementary Figure 5.1).

3.3. Clustering of wheat and tritordeum lines based on CD epitopes profile

A Principal Component Analysis (PCA) and unsupervised hierarchical clustering were carried out for the alpha- and gamma-gliadin/secalin amplicons separately, comprising variables such as the abundance of the CD epitopes, the p31-43 peptide, and the amplicons 'types'. Other parameters, like the number of CD epitopes and the number of putative genes and pseudogenes, were included in the PCA (Figure 5.6).

The lines assayed were separated into four clusters based on the alpha-gliadin amplicon results (Figure 5.6B). One of the clusters comprised the durum wheat and tritordeum lines. Interestingly, two bread wheat lines – RelmoSiriri and THA85 – were in this cluster (Figure 5.6B). The main features of lines from this cluster were a lower abundance for all the CD epitopes and alpha-gliadin amplicons 'types', but not for the p31-43 peptide, the Alpha_2, and the DQ2.5_glia_ α 3. The other three clusters were comprised of bread wheat lines and were clearly separated in the PCA



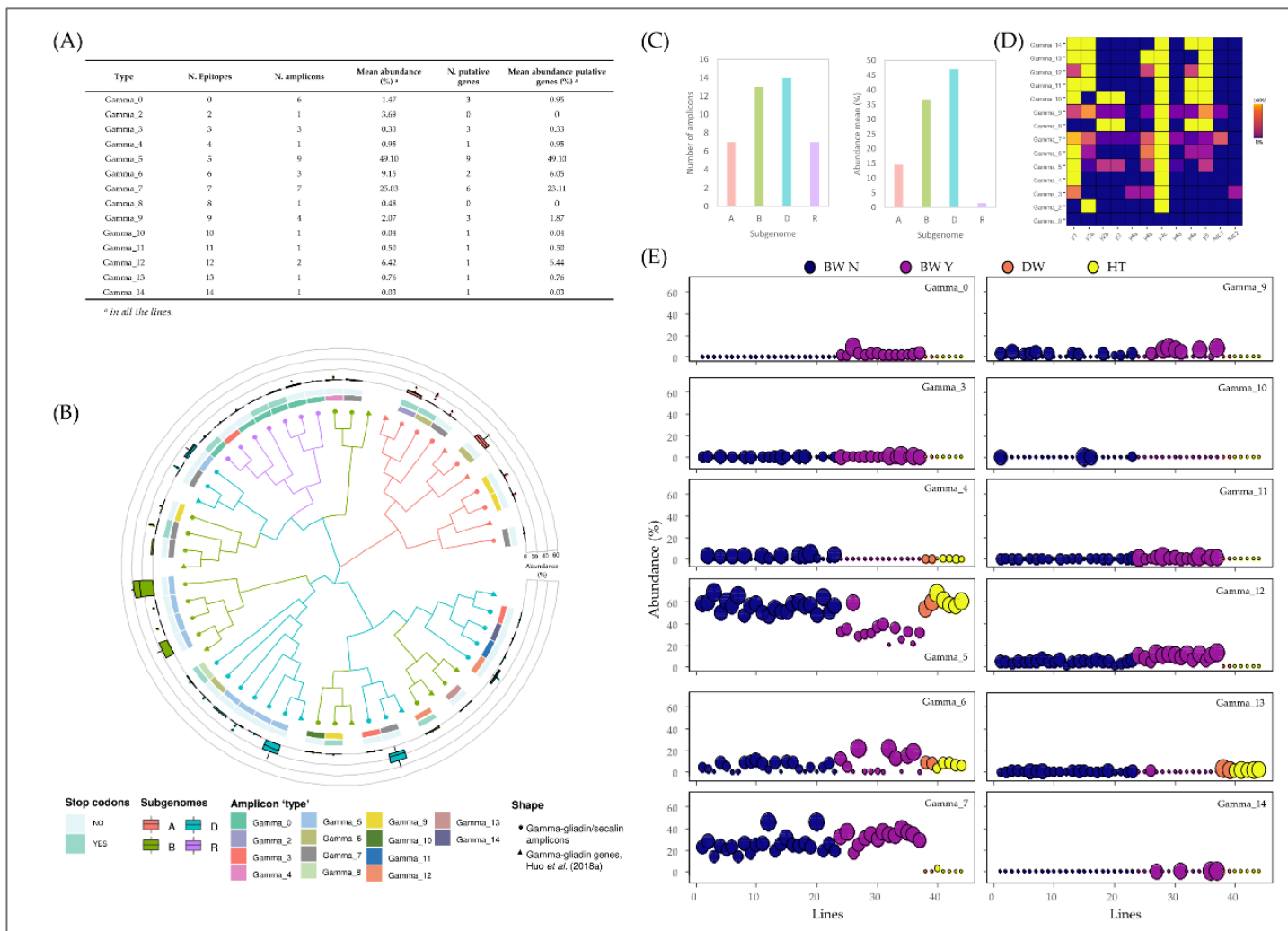


Figure 5.5. Characterization of the gamma-gliadin's amplicons 'types' (A) The abundance of each gamma-gliadin 'type'. (B) Phylogenetic tree using the maximum-likelihood method in MEGA 10.1.7 from a MUSCLE sequence alignment (default parameters). The amplicons were assigned to each subgenome by their similarity with the gamma-gliadin genes identified in Huo et al. (2018a) for the Chinese Spring cultivar. From outside to inside: (i) the abundance for all the lines of each amplicon 'type', (ii) presence of premature stop codons, (iii) hits of the p31-43 peptide, and (iv) the number of DQ2.5 epitopes in their translated peptide sequence. (C) Number and mean abundance (%) per subgenome. (D) Frequency with which epitope appears in each amplicon 'type': the number of amplicons containing each epitope divided by the total amplicons classified in this 'type'. (E) The abundance of each gamma-gliadin's amplicon 'type' for each wheat and tritordeum line based on putative gene amplicons. BW N: bread wheat without rye translocation; BW Y: bread wheat with rye translocation; DW: durum wheat; HT: tritordeum. The dot sizes are ranged from the lowest (smallest) to the highest (biggest) abundance of all lines within each amplicon 'type'.



analysis (Figure 5.6A,B). The second and third ones - from left to right side in Figure 5.6B - had an opposite abundance profile concerning epitopes like the DQ2.5_glia_α1a, DQ2.5_glia_α3, the p31-43 peptide, and the Alpha_2 amplicon 'type', which were abundant in the third cluster and poorly represented in the second one. In contrast, the Alpha_0 was less abundant in the third cluster than in the second one (Figure 5.6B). The remaining five bread wheat lines were clustered together and presented a high abundance for all of the CD epitopes and amplicon 'types', except for the Alpha_2, and Alpha_0 (Figure 5.6B). OlaetaArtillero and ACA321 stood out because of their high abundance of Alpha_7, which contains the full 33-mer peptide (Figure 5.6B).

Regarding the gamma-gliadin/secalin amplicon, the hierarchical clustering provided four clusters (Figure 5.6D). The durum wheat and tritordeum lines were grouped together in one of them and presented lower abundance for almost all the CD epitopes and gamma-gliadin amplicon 'types' (Figure 5.6D). Interestingly, the bread wheat lines were well separated by the presence of rye translocation, and only BIOINTA2004 line (containing the rye translocation) was grouped with the bread wheat lines without secalins (Figure 5.6C,D). In general, bread wheat lines grouped in the second and third clusters - from left to right side in Figure 5.6D - had a high variability of profiles. However, the first cluster located on the far left, comprising bread wheat lines with the rye translocation, had a contrasting profile than the durum wheat and tritordeum lines, particularly, the Gamma_0 amplicon 'type' was higher abundant in this cluster than in the rest of them, as for the DQ2.5_sec2 epitope and Gamma_14 in some of the lines (Figure 5.6D).

The linkage of the abundance of the epitopes and the amplicons 'type' of all the lines with previous immunogenic screening of oligopeptides encompassing CD epitopes gave as a result the ranking of the lines based on their immunogenic potential score (Figure 5.6E). As it is shown, the durum wheat-based lines had the lower scores, far away from the first-positioned bread wheat lines (Figure 5.6D). Comparing this ranking with the lines clusterization in Figure 5.6B and D, OlaetaArtillero was maintained as one of the most potential immunogenic lines as it was present among the first places in the ranking, while ACA321 had a much lower score despite having a 33-mer abundance similar to OlaetaArtillero. This was due to ACA321 presenting low overall epitope abundance for the gamma-gliadin amplicons, close to the durum wheat-based lines (Figure 5.6E). Many of the first-ranked lines had a high content of alpha-gliadin epitopes and the 33-mer peptide, and a low abundance of the Alpha_0 'type' (Figure 5.6E). Interestingly, the presence of the rye translocation did not seem to have a significant role in this ranking, as all the bread wheat lines with secalins were intercalated with the others (Figure 5.6E).



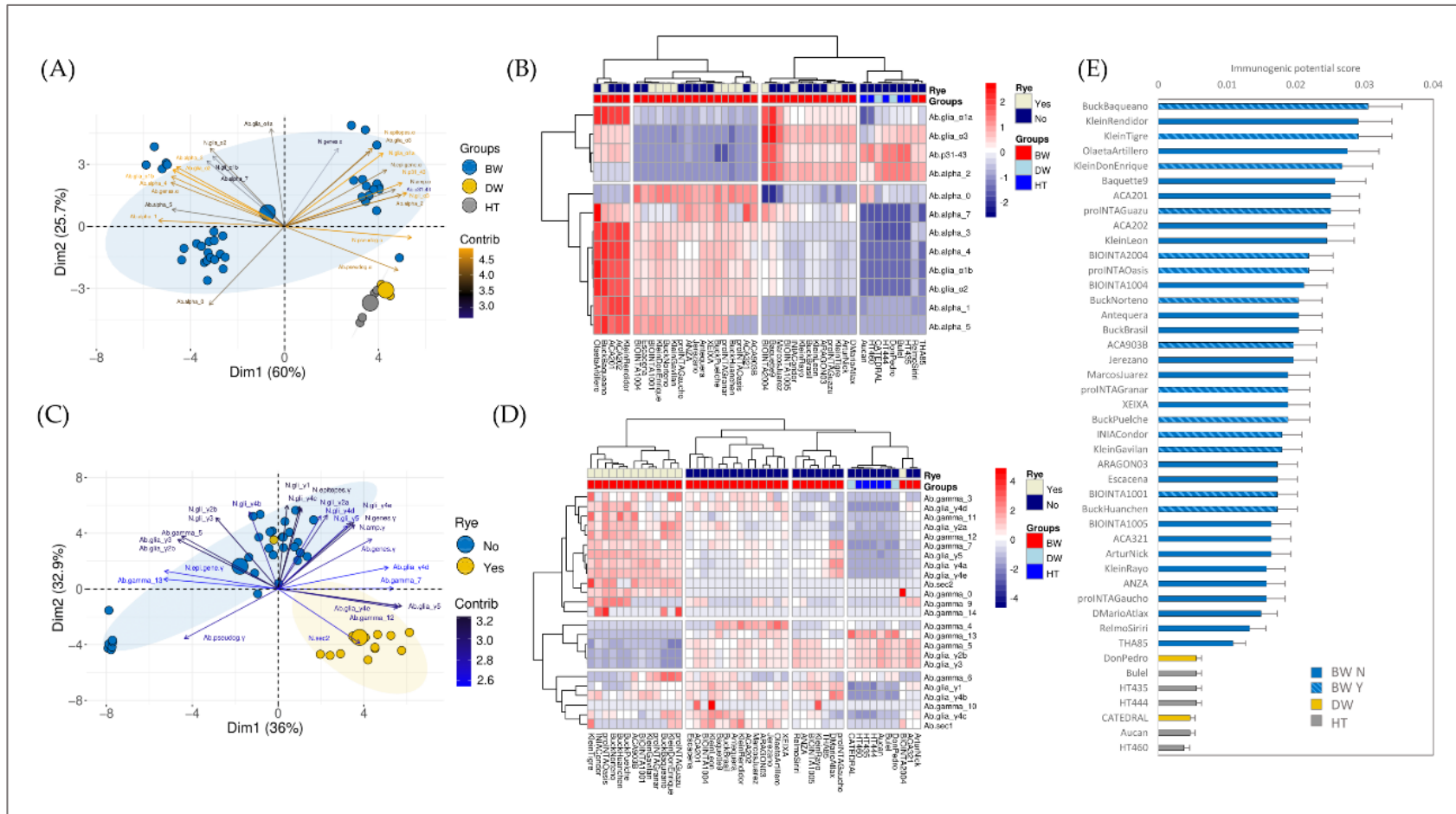


Figure 5.6. Principal Component Analysis (PCA) for the (A) alpha- and (C) gamma-gliadin amplicons. The contribution of each variable to the variance of the model is indicated by a color code. The big size dots indicate the center of the ellipses with 95 % of confidence per each group of lines. Because of the high number of variables for the gamma-gliadin amplicon, the top 25 variables that most contribute to the variance of the model were included to improve the visualization of the results. Heatmap of the scaled abundance of the CD epitopes and the amplicons 'types' for alpha- (B) and gamma-gliadin (D) amplicons. The dendrogram based on Pearson and Spearman correlation distance, for alpha- and gamma-gliadin amplicons respectively, is represented for the epitopes and amplicons 'types', and for the bread wheat (BW), durum wheat (DW), and tritordeum (HT) lines. The rye translocation is indicated with Yes/No in the Rye legend. (E) Ranking of lines based on their immunogenic potential score. The bars indicate the mean plus the standard error.



3.4. Identification of CD alpha-gliadin epitopes with one mismatch

The CD alpha-gliadin epitopes containing one mismatch with respect to the canonical ones were identified, and their abundance was also calculated (Figure 5.7). Next, the $\log_2(\text{FC})$ between them and their canonical epitopes was measured (Supplementary Table 5.10). Only for three of them, the $\log_2(\text{FC})$ mean was higher than zero, which means that these epitope variants were more abundant than the canonical ones in most of the lines (Figure 5.7A). The S variant of DQ2.5_glia_α2 (PQPQLPYSQ) had higher abundance in durum wheat and tritordeum lines, while the P variant of DQ2.5_glia_α3 and the p31-43 P variant were more abundant in bread wheat lines, being the last one overrepresented in the lines containing the rye translocation (Figure 5.7B).

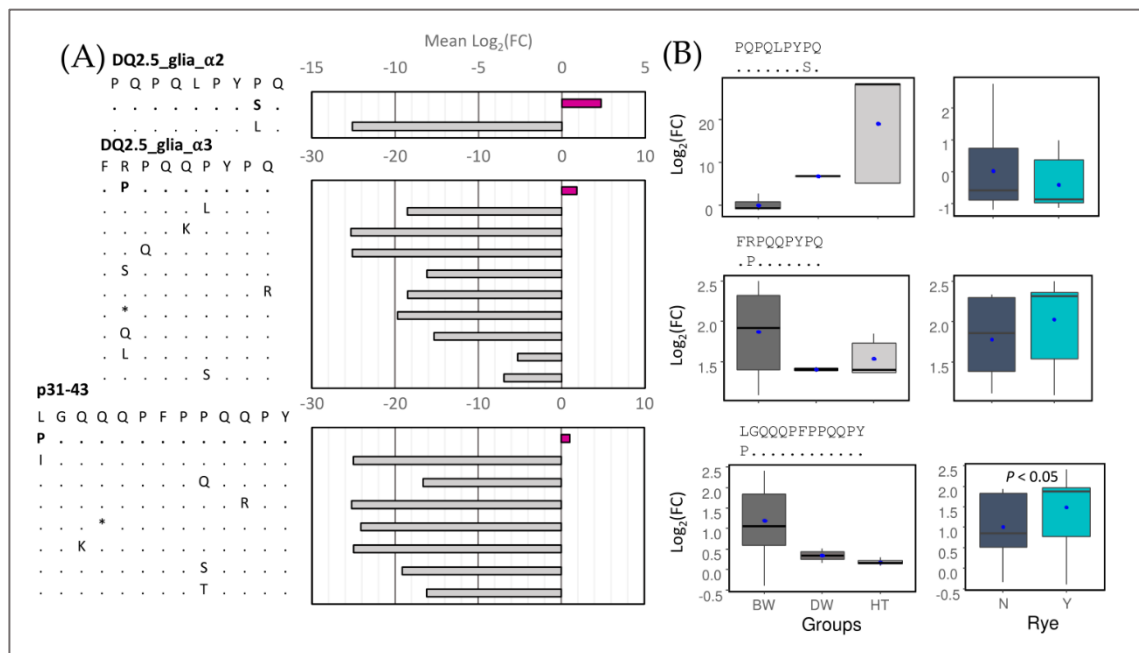


Figure 5.7. Analysis of epitopes' variants (one mismatch). (A) Mean of the $\log_2(\text{FC})$ between the abundance of the epitopes' variants and the canonical epitopes, or the immunogenic peptide in the case of p31-43. (B) Boxplots for the three epitopes' variants with the $\log_2(\text{FC}) > 0$ for wheat and tritordeum lines (left) and for the presence of the rye translocation only in bread wheat lines (right). The $\log_2(\text{FC})$ was calculated as follows: the abundance of the epitope' variant/abundance of the canonical epitope. For the rye boxplots, the Mann-Whitney-Wilcoxon test was performed, and only the significant P -value was indicated. The blue dots indicate the mean value.

4. Discussion

RNAi technology has been used to down-regulate the genes encoding for the gliadins, the major responsible for triggering CD. As a result, a reduction of up to 98 % for these proteins was reported in some wheat lines (Gil-Humanes et al., 2010). More recently, CRISPR gene editing was used to knock out the alpha-gliadins multigene family (Sánchez-León et al., 2018). The use of these techniques is not without controversy, particularly in the European Union, where they are subject to



very strong regulation, which greatly limits the development and marketing of such products. The identification of wheat varieties with low content of immunogenic epitopes related to CD is still an important objective to develop final food products or to use them as parents in future breeding programs, either by classical breeding or by biotechnological tools. In particular, for the use of gene editing techniques, it is essential to use wheat varieties with low gliadin structural complexity, either with a low number of genes or with a low number of epitopes. This is now facilitated by using embryogenesis-related transcription factors such as BABY BOOM (Bbm), GROWTH-REGULATING FACTORS (GRFs) or Wuschel-related homeobox 5 (WOX5), which improve embryogenesis, tissue culture regeneration and plant transformation efficiency, and allowing the genotype-independent transformation in plants (Bilichak et al., 2018; Debernardi et al., 2020; Wang et al., 2022).

Diverse approaches involving long-read sequencing, RNA and DNA amplicon sequencing, and proteomic analysis have been used to study the CD epitopes in the gliadin sequences of wheat. Most of them were focused on the potential immunogenicity of alpha-gliadins in *Triticum turgidum*, *Triticum aestivum* and their ancestors, being studied to look for the T cell stimulatory CD epitopes (Huo et al., 2018b; Ozuna et al., 2015; Salentijn et al., 2013; Schaart et al., 2021; Halstead-Nussloch et al., 2021). However, there are fewer studies involving other gliadin groups such as the gamma-gliadins in terms of the presence and abundance in CD epitopes (Wang et al., 2017; Cho et al., 2018). In this work, we have analyzed the immunogenic gliadin complexes located on the repetitive domain of the alpha- and gamma-gliadins in a set of bread and durum wheat, and tritordeum lines by high throughput amplicon sequencing. The two amplicon databases obtained were compared to the gene sequences of the alpha- and gamma-gliadins published for the CS cultivar (Huo et al., 2018b; 2018a). Assignment to the three wheat subgenomes of the gamma-gliadin amplicons by their proximity to the gene sequence from CS in the phylogenetic tree resulted in a total of 7, 13, and 14 amplicons assigned to A-, B- and D-subgenomes respectively. In addition, seven 40k- γ -secalin amplicons were identified, as they are located in the short arm of the 1R chromosome of rye (Kozub et al., 2014), present on the bread wheat lines with the 1BL/1RS rye translocation. However, the subgenome assignation of alpha-gliadin amplicons either by searching subgenome-specific motifs around the 33-mer region (van Herpen et al., 2006) or by their similarity to CS reference genes (Huo et al., 2018b) had discrepancies. Actually, the motifs employed for subgenome assignation were searched in the CS reference genes, aligning themselves in genes from different subgenomes to that of the motif itself. However, most of the A-subgenome amplicons by alpha-gliadin motifs were clustered together, as were many of the D- and B-subgenomes.

Both the alpha- and gamma-gliadins presented a high rate of pseudogenes due to the high percentage of glutamine residues encoded by CAA and CAG codons that can be mutated into stop codons through C \rightarrow T transition (Huo et al., 2018b; Paris et al., 2021). We report 27.6% and 22.0% of pseudogenes for the alpha- and gamma-



gliadin amplicons, and many of them presented the mutation of the glutamine into a stop codon, mainly in the alpha-gliadins. However, the number of pseudogenes could be underestimated because the amplicon does not cover the entire sequence of both gliadins genes as reported by Sánchez-León et al. (2021). In fact, there was previously described 45% of pseudogenes in alpha-gliadins, while the gamma-gliadin-genes presented 21% of them, close to our results for this family (Huo et al., 2018a; Huo et al., 2018b). The alpha-gliadin amplicons that presented premature stop codons in the first repetitive domain could be regulated by a mechanism called nonsense-mediated mRNA decay (NMD) at the post-translational level so that the production of truncated proteins is prevented (Hug et al., 2016). We have shown that pseudogenes were strongly ligated to the subgenomes A and B, mainly for the alpha-gliadins. Moreover, this high rate of pseudogenes in both subgenomes could partly explain the less immunogenicity of durum wheat and tritordeum genotypes (Auricchio et al., 1982; Vaquero et al., 2018).

Recently, Vríz et al. (2021) proposed a ranking of CD epitopes based on their enzymatic degradation, HLA-DQ binding affinity, and T cell activation, including the data from Tye-Din et al. (2010), which allows for determining their potential to trigger CD. In this ranking, the alpha-gliadin epitopes were the most immunodominant. However, there was an exception for the DQ2.5_glia_α3 epitope because of its chymotrypsin degradation decreasing its adequate binding to HLA-DQ receptors. Interestingly, for the set of genotypes analyzed in this work, about 40% of the alpha-gliadin amplicons which an overall abundance of 53.4%, mainly located in the B-subgenome, had no CD epitopes. This is in agreement with previous works which reported that alpha-gliadin genes containing few CD epitopes or a few copies of them were in the B-subgenome (Wang et al., 2017; Halstead-Nussloch et al., 2021). Among the amplicons with CD epitopes, there were two epitopes included in the 33-mer peptide with a low or null abundance in durum wheat and tritordeum lines. In those lines, the epitope with a high abundance was DQ2.5_glia_α3, which is the less immunogenic alpha DQ2.5 epitope as mentioned before (Anderson et al., 2006). The fact that the durum wheat and tritordeum lines lack two of the epitopes included in the 33-mer implies that this peptide cannot be found in these lines, and therefore they also lack the Alpha_7 amplicon 'type', which comprises the six overlapping epitopes of the 33-mer and one DQ2.5_glia_α3 downstream. The 33-mer was only present in alpha-gliadin amplicons from the D-subgenome, and at low frequency, which means that there are only few copies of the gene coding this peptide. The low frequency of the 33-mer in bread wheat was reported previously (Ozuna et al., 2015; Huo et al., 2018b). The same works proposed that the origin of the 33-mer was probably due to variation events of *Aegilops tauschii* alpha-gliadin genes, supporting the fact that this immunogenic peptide can only be found on the D-subgenome.

In addition to the 33-mer, the p31-43 peptide was reported to activate the innate immune response inducing a non-HLA-mediated inflammatory reaction. Its



immunogenicity was proved in both in vitro and in vivo studies (Gianfrani et al., 2005; Barone et al., 2014; Maiuri et al., 1996). This peptide was mainly present in the alpha-gliadin amplicons from the A-subgenome, in agreement with the work of Halstead-Nussloch et al. (2021). This peptide had a conserved high abundance in most of the lines compared to some of the alpha-gliadin DQ2.5 epitopes. In addition, the p31-43P epitope variant, which has been described as an immunogenic peptide for CD patients (Maiuri et al., 1996, 2003), was more represented overall than the canonical peptide in bread wheat lines, especially for the lines containing the rye translocation. There were also other epitope variants more abundant than the canonical epitopes. For example, the DQ2.5_glia_α2 PQPQLPYSQ variant (P to S substitution) showed comparable T-cell proliferation and anti-33-mer binding capacity than its canonical CD epitopes (Ruiz-Carnicer et al., 2019), one of the major epitopes recognized by most of CD patients (Tye-Din et al., 2010). Interestingly, this variant was more abundant on tritordeum lines and was found only in the A-subgenome (Ruiz-Carnicer et al., 2019). Overall, there is a strong relationship between the CD epitope subgenomes and their immunogenicity, those with the highest immunogenic potential the ones from the A- and, mainly, the D-subgenome (Halstead-Nussloch et al., 2021). Thus, durum wheat and tritordeum lines would have a weak immune response in CD patients and fewer gluten immunogenic peptides (GIPs) are expected in the excretion of healthy individuals (Auricchio et al., 1982; Vaquero et al., 2018). However, there was a high heterogeneity of the alpha-gliadin amplicon abundance profile from the three subgenomes in bread wheat lines, so there are lines with a low-immunogenic potential as they have clustered together with the durum wheat and tritordeum lines: RelmoSiriri and THA85, for example. Actually, THA85 is the only bread wheat line in which we have not found the complete 33-mer peptide. Although this will have to be corroborated in subsequent analyses, it is ranked as one of the best candidates for precision breeding programs in order to obtain low-immunogenic wheat. This clusterization also allowed us to identify five bread wheat lines with the highest immunogenic potential based on the alpha-gliadin amplicon results, including only one line with the rye translocation. Clearly these lines should be avoided in breeding programs to obtain varieties with low immunogenicity.

As mentioned before, there are very few studies that have simultaneously categorized the CD epitope content in both the alpha- and gamma-gliadins in wheat genotypes. Actually, the gamma-gliadin genes have more diverse and numerous CD epitopes, as described in Wang et al. (2017). A total of 10 gamma-gliadin HLA-DQ2.5 epitopes and two epitopes of secalins were mapped to the gamma-gliadin/secalin amplicons which include the 26-mer region. This peptide is highly resistant to intestinal brush border membrane proteolysis and, as it is multivalent, the intact 26-mer peptide is more antigenic compared to its smaller monovalent counterparts (Shan et al., 2005). Like the 33-mer, the 26-mer is only present in the D-subgenome, and at a low frequency in the transcripts of bread wheat (Salentijn et al., 2012). Of the epitopes included in this region, the most abundant was DQ2.5_glia_γ4c in all the



lines, independently of their genetic background, as it presents a high number of copies along the gamma-gliadin amplicon. In addition to this, other epitopes also presented in the 26-mer region had a higher abundance in bread wheat, with significant differences between those with and without the rye translocation. The bread wheat lines with the 1BL/1RS translocation had a higher abundance of the epitope DQ2.5_glia_γ5, reported as one of the gamma-gliadin epitopes with higher immunological relevance (Vriz et al., 2021). Only one 40k-γ-secalin amplicon showed the epitope DQ2.5_sec2, in combination to DQ2.5_glia_γ4b and DQ2.5_glia_γ4c, while the rest of secalin amplicons do not contain any epitope matches. As shown in previous research, the 40k-γ-secalins present few epitope hits, including the DQ2.5_glia_γ4c between them. In fact, this secalin type has less epitope than the rest of rye prolamins such as ω-secalins or 75k-γ-secalins, whose genes are present in the 2R chromosome (Lexhaller et al., 2019; Kozub et al., 2014). The other 40k-γ-secalins found on these bread wheat lines comprised the Gamma_0 'type', only present in wheat lines containing the rye translocation. However, these lines were also the only ones to present the gamma-gliadin amplicon with the highest number of CD epitopes (Gamma_14), which belonged to the D-subgenome. There were three amplicons 'types' present only in the B-subgenome. All of them were absent in the bread wheat lines with the 1BL/1RS translocation, as 1BS chromosome were fully or partially depleted in those lines. However, the BIOINTA2004 line was the exception, as it had one of those amplicons 'types' at low abundance. The *Gli-B11* allele, associated with this translocation, was not complete in its A-PAGE protein profile, so a partial translocation could explain the differences between BIOINTA2004 and the other lines with the rye translocation.

The durum wheat and tritordeum genotypes had a lower abundance for most of the epitopes and amplicon 'types', including the Gamma_14. Like alpha-gliadins, the gamma-gliadins presented less immunogenic elements for those lines, which makes them good candidate for obtaining varieties with low-immunogenic gluten by precision breeding programs. This also applies to the bread wheat lines clustered with the durum wheat and tritordeum lines which have a low epitopes' abundance profile, including one line with the rye translocation (BIOINTA2004). The rest of the bread wheat lines containing the translocation were clustered together and presented a high abundance of many epitopes and amplicon 'types', including Gamma_0.

The clusterization of lines by the alpha- and gamma-gliadin amplicons allowed us to identify groups of lines that could be potential candidates for precision breeding programs. However, the hierarchical clustering reflects only the abundance of the set of epitopes and amplicons 'types', and does not consider information on how much immunogenic each epitope is as they do not contribute equally to trigger the CD response. Connecting the results of the amplicon analysis with the scores presented for the 20-mer oligopeptides in the work of Tye-Din et al. (2010) gave rise to the ranking of the lines assayed in this work, ordered by their immunogenic



potential score. Furthermore, this score comprises the data from both immunogenic complexes, taking into account the differences in their immunoreactivity. In this ranking, the durum wheat and tritordeum lines were in the last positions, which means that they presented a lower immunogenic potential. Close to these, the THA85 and RelmoSiriri bread wheat lines also stood out for their low scores. The score was more influenced by the epitope data from alpha-gliadins, since this complex has more immunological relevance than the gamma-gliadins (Vriz et al., 2021), also reflected in the results of Tye-Din et al. (2010). In line with the alpha-gliadins immunodominance, the bread wheat lines with the rye translocation, despite of the full or partial depletion of the 1BS chromosome containing the gamma-gliadin loci, were intercalated in the ranking with the other bread wheat lines, not having the translocation a decisive role in their immunogenic potential. Despite this ranking being based on previous IFN- γ ELISpot assays with PBMCs from CD patients, a further ELISpot assay is needed to know the immunoreactivity of the lines selected for future screening. In addition, this ranking could be improved with NGS data from the omega-gliadin genes, as they have an important role in the immunogenic response in CD (Vriz et al., 2021).

As shown in this work, there is a high variability of the immunogenic potential between wheat and tritordeum lines. The durum wheat-based lines stand out for their low abundance of epitopes, mainly the high immunogenic alpha-gliadin ones, and therefore they possess a great potential as candidates for precision breeding programs. However, there was a high epitope abundance heterogeneity among bread wheat lines, the lines with an epitope profile comparable to that of durum wheat and tritordeum lines the best to further develop low-immunogenic varieties. The rye translocation in bread wheat lines can be detected and quantified, playing an important role in the classification of genotypes by their gamma-gliadin amplicon profiles. Although the rye translocation provides varieties with a lower number of epitopes, the translocation did not imply a lower immunogenic potential compared to the rest of the wheat lines. Even though all the lines described in this work contain gluten, there are marked differences between them, both in the number of epitopes and in the number of amplicons and their abundances, which can be used for effective variety selection towards the development of low-gluten products, or as starting material for the application of biotechnological tools such as CRISPR/Cas.

5. Acknowledgements:

We acknowledge Instituto Nacional de Tecnología Agropecuaria (INTA) for providing the Argentinian wheat varieties.



6. References

- Altenbach, S. B., Chang, H. C., Simon-Buss, A., Mohr, T., Huo, N., and Gu, Y. Q. (2020). Exploiting the reference genome sequence of hexaploid wheat: a proteomic study of flour proteins from the cultivar Chinese Spring. *Funct. Integr. Genomics* 20, 1–16. doi:10.1007/s10142-019-00694-z.
- Anderson, O. D., Dong, L., Huo, N., and Gu, Y. Q. (2012). A new class of wheat gliadin genes and proteins. *PLoS One* 7, e52139.
- Anderson, R. P., van Heel, D. A., Tye-Din, J. A., Jewell, D. P., and Hill, A. V. S. (2006). Antagonists and non-toxic variants of the dominant wheat gliadin T cell epitope in coeliac disease. *Gut* 55, 485–491.
- Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*. 361, eaar7191.
- Auricchio, S., De Ritis, G., De Vincenzi, M., Occorsio, P., and Silano, V. (1982). Effects of gliadin-derived peptides from bread and durum wheats on small intestine cultures from rat fetus and coeliac children. *Pediatr. Res.* 16, 1004–1010.
- Barone, M. V., Troncone, R., and Auricchio, S. (2014). Gliadin peptides as triggers of the proliferative and stress/innate immune response of the celiac small intestinal mucosa. *Int. J. Mol. Sci.* 15, 20518–20537.
- Bilichak, A., Luu, J., Jiang, F., and Eudes, F. (2018). Identification of BABY BOOM homolog in bread wheat. *Agri Gene* 7, 43–51.
- Cho, K., Beom, H.-R., Jang, Y.-R., Altenbach, S. B., Vensel, W. H., Simon-Buss, A., et al. (2018). Proteomic profiling and epitope analysis of the complex α -, γ -, and ω -gliadin families in a commercial bread wheat. *Front. Plant Sci.* 9, 818.
- Debernardi, J. M., Tricoli, D. M., Ercoli, M. F., Hayta, S., Ronald, P., Palatnik, J. F., et al. (2020). A GRF–GIF chimeric protein improves the regeneration efficiency of transgenic plants. *Nat. Biotechnol.* 38, 1274–1279.
- Edgar, R. (2010). Usearch. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).
- Fallang, L.-E., Bergseng, E., Hotta, K., Berg-Larsen, A., Kim, C.-Y., and Sollid, L. M. (2009). Differences in the risk of celiac disease associated with HLA-DQ2. 5 or HLA-DQ2. 2 are related to sustained gluten antigen presentation. *Nat. Immunol.* 10, 1096–1101.
- FAO (2022). FAOSTAT. Available at: <https://www.fao.org/faostat/>.
- Gianfrani, C., Auricchio, S., and Troncone, R. (2005). Adaptive and innate immune responses in celiac disease. *Immunol. Lett.* 99, 141–145.
- Gil-Humanes, J., Pistón, F., Rosell, C. M., and Barro, F. (2012). Significant down-regulation of γ -gliadins has minor effect on gluten and starch properties of bread wheat. *J. Cereal Sci.* 56, 161–170. doi:10.1016/j.jcs.2012.02.009.
- Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L. M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proc. Natl. Acad. Sci. U. S. A.* 107, 17023–17028. doi:10.1073/pnas.1007773107.
- Halstead-Nussloch, G., Tanaka, T., Copetti, D., Paape, T., Kobayashi, F., Hatakeyama, M., et al. (2021). Multiple Wheat Genomes Reveal Novel *Gli-2* Sublocus Location and Variation of Celiac Disease Epitopes in Duplicated α -Gliadin Genes. *Front. Plant Sci.*, 12:715985.
- Harberd, N. P., Bartels, D., and Thompson, R. D. (1985). Analysis of the gliadin multigene loci in bread wheat using nullisomic-tetrasomic lines. *Mol. Gen. Genet.* MGG 198, 234–242.



- Haro, C., Guzmán-López, M. H., Marín-Sanz, M., Sánchez-León, S., Vaquero, L., Pastor, J., et al. (2022). Consumption of Triticum Bread Reduces Immunogenic Gluten Intake without Altering the Gut Microbiota. *Foods* 11, 1439.
- Hug, N., Longman, D., and Cáceres, J. F. (2016). Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res.* 44, 1483–1495.
- Huo, N., Zhang, S., Zhu, T., Dong, L., Wang, Y., Mohr, T., et al. (2018a). Gene duplication and evolution dynamics in the homeologous regions harboring multiple prolamin and resistance gene families in hexaploid wheat. *Front. Plant Sci.* 9, 673. doi:10.3389/fpls.2018.00673.
- Huo, N., Zhu, T., Altenbach, S., Dong, L., Wang, Y., Mohr, T., et al. (2018b). Dynamic Evolution of α -Gliadin Prolamin Gene Family in Homeologous Genomes of Hexaploid Wheat. *Sci. Rep.* 8:5181. doi:10.1038/s41598-018-23570-5.
- Jouanin, A., Tenorio-Berrio, R., Schaart, J. G., Leigh, F., Visser, R. G. F., and Smulders, M. J. M. (2020). Optimisation of droplet digital PCR for determining copy number variation of α -gliadin genes in mutant and gene-edited polyploid bread wheat. *J. Cereal Sci.* 92, 102903.
- Kozub, N. A., Motsnyi, I. I., Sozinov, I. A., Blume, Y. B., and Sozinov, A. A. (2014). Mapping a new secalin locus on the rye 1RS arm. *Cytol. Genet.* 48, 203–207.
- Lexhaller, B., Colgrave, M. L., and Scherf, K. A. (2019). Characterization and relative quantitation of wheat, rye, and barley gluten protein types by liquid chromatography–tandem mass spectrometry. *Front. Plant Sci.* 10, 1530.
- Ludvigsson, J. F., Leffler, D. A., Bai, J. C., Biagi, F., Fasano, A., Green, P. H. R., et al. (2013). The Oslo definitions for coeliac disease and related terms. *Gut* 62, 43–52. doi:10.1136/gutjnl-2011-301346.
- Maiuri, L., Ciacci, C., Ricciardelli, I., Vacca, L., Raia, V., Auricchio, S., et al. (2003). Association between innate response to gliadin and activation of pathogenic T cells in coeliac disease. *Lancet* 362, 30–37.
- Maiuri, L., Troncone, R., Mayer, M., Coletta, S., Picarelli, A., Vincenzi, M. De, et al. (1996). In vitro Activities of A-Gliadin-Related Synthetic Peptides Damaging Effect on the Atrophic Coeliac Mucosa and Activation of Mucosal Immune Response in the Treated Coeliac Mucosa. *Scand. J. Gastroenterol.* 31, 247–253.
- Metakovsky, E., Melnik, V., Rodriguez-Quijano, M., Upelniek, V., and Carrillo, J. M. (2018). A catalog of gliadin alleles: Polymorphism of 20th-century common wheat germplasm. *Crop J.* 6, 628–641.
- Murray, M. G., and Thompson, W. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4326.
- Ozuna, C. V., Iehisa, J. C. M., Giménez, M. J., Alvarez, J. B., Sousa, C., and Barro, F. (2015). Diversification of the celiac disease α -gliadin complex in wheat: A 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *Plant J.* 82, 794–805. doi:10.1111/tpj.12851.
- Paris, R., Petruzzino, G., Savino, M., De Simone, V., Ficco, D., and Trono, D. (2021). Genome-Wide Identification, Characterization and Expression Pattern Analysis of the γ -Gliadin Gene Family in the Durum Wheat (*Triticum durum* Desf.) Cultivar Svevo. *Genes (Basel)*. 12, 1743.
- Qi, P.-F., Chen, Q., Ouellet, T., Le, C.-X., Ai, J., Wei, Y.-M., et al. (2013). Identification of 40k γ -secalin genes. *J. Genet.* 92, 151–154.
- Ruiz-Carnicer, Á., Comino, I., Segura, V., Ozuna, C. V., Moreno, M. de L., López-Casado, M. Á., et al. (2019). Celiac immunogenic potential of α -gliadin epitope variants from *Triticum* and *Aegilops* species. *Nutrients* 11, 220.
- Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *Int. J. Serv. Sci.* 1, 83–98.



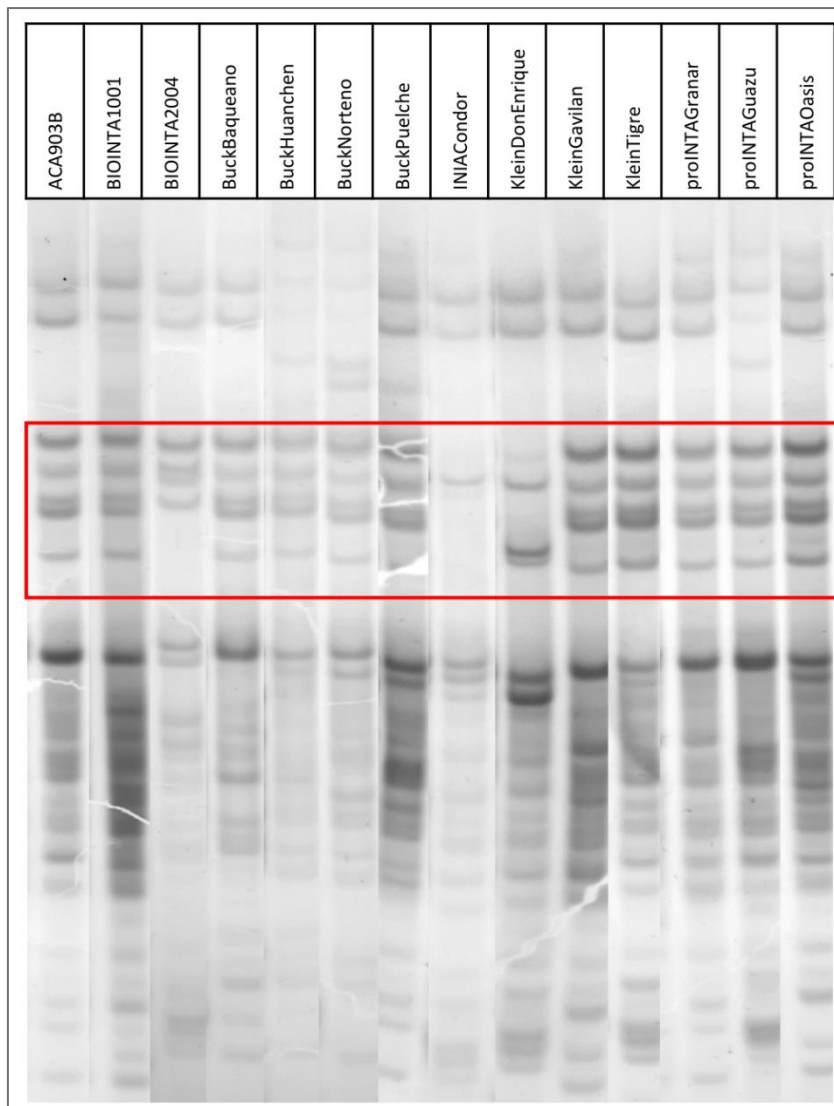
- Salentijn, E. M. J., Esselink, D. G., Goryunova, S. V., van der Meer, I. M., Gilissen, L. J. W. J., and Smulders, M. J. M. (2013). Quantitative and qualitative differences in celiac disease epitopes among durum wheat varieties identified through deep RNA-amplicon sequencing. *BMC Genomics* 14, 1–16.
- Salentijn, E. M. J., Mitea, D. C., Goryunova, S. V., van der Meer, I. M., Padioleau, I., Gilissen, L. J. W. J., et al. (2012). Celiac disease T-cell epitopes from gamma-gliadins: immunoreactivity depends on the genome of origin, transcript frequency, and flanking protein variation. *BMC Genomics* 13, 1–12.
- Sánchez-León, S., Gil-Humanes, J., Ozuna, C. V., Giménez, M. J., Sousa, C., Voytas, D. F., et al. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant Biotechnol. J.* 16, 902–910. doi:10.1111/pbi.12837.
- Sánchez-León, S., Giménez, M. J., and Barro, F. (2021). The α -Gliadins in Bread Wheat: Effect of Nitrogen Treatment on the Expression of the Major Celiac Disease Immunogenic Complex in Two RNAi Low-Gliadin Lines. *Front. Plant Sci.* 12, 742.
- Schaart, J. G., Salentijn, E. M. J., Goryunova, S. V., Chidzanga, C., Esselink, D. G., Gosman, N., et al. (2021). Exploring the alpha-gliadin locus: the 33-mer peptide with six overlapping coeliac disease epitopes in *Triticum aestivum* is derived from a subgroup of *Aegilops tauschii*. *Plant J.* 106, 86–94.
- Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., et al. (2002). Structural basis for gluten intolerance in Celiac Sprue. *Science*. 297, 2275–2279. doi:10.1126/science.1074129.
- Shan, L., Qiao, S.-W., Arentz-Hansen, H., Molberg, Ø., Gray, G. M., Sollid, L. M., et al. (2005). Identification and analysis of multivalent proteolytically resistant peptides from gluten: implications for celiac sprue. *J. Proteome Res.* 4, 1732–1741.
- Shewry, P. (2019). What is gluten—why is it special? *Front. Nutr.*, 101.
- Shewry, P. R. (2009). Wheat. *J. Exp. Bot.* 60, 1537–1553.
- Sollid, L. M. (2002). Coeliac disease: Dissecting a complex inflammatory disorder. *Nat. Rev. Immunol.* 2, 647–655. doi:10.1038/nri885.
- Sollid, L. M., Qiao, S.-W., Anderson, R. P., Gianfrani, C., and Koning, F. (2012). Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics* 64, 455–460.
- Sollid, L. M., Tye-Din, J. A., Qiao, S. W., Anderson, R. P., Gianfrani, C., and Koning, F. (2020). Update 2020: nomenclature and listing of celiac disease-relevant gluten epitopes recognized by CD4+ T cells. *Immunogenetics* 72, 85–88. doi:10.1007/s00251-019-01141-w.
- Spaenij-Dekking, L., Kooy-Winkelaar, Y., van Veelen, P., Drijfhout, J. W., Jonker, H., van Soest, L., et al. (2005). Natural variation in toxicity of wheat: potential for selection of nontoxic varieties for celiac disease patients. *Gastroenterology* 129, 797–806.
- Tye-Din, J. A., Stewart, J. A., Dromey, J. A., Beissbarth, T., Van Heel, D. A., Tatham, A., et al. (2010). Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci. Transl. Med.* 2, 41ra51–41ra51. doi:10.1126/scitranslmed.3001012.
- Vader, W., Stepniak, D., Kooy, Y., Mearin, L., Thompson, A., van Rood, J. J., et al. (2003). The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. *Proc. Natl. Acad. Sci.* 100, 12390–12395.
- van Herpen, T. W. J. M., Goryunova, S. V., van der Schoot, J., Mitreva, M., Salentijn, E., Vorst, O., et al. (2006). Alpha-gliadin genes from the A, B, and D genomes of wheat contain different sets of celiac disease epitopes. *BMC Genomics* 7, 1–13. doi:10.1186/1471-2164-7-1.



- Vaquero, L., Comino, I., Vivas, S., Rodríguez-Martín, L., Giménez, M. J., Pastor, J., et al. (2018). Tritordeum: a novel cereal for food processing with good acceptability and significant reduction in gluten immunogenic peptides in comparison with wheat. *J. Sci. Food Agric.* 98, 2201–2209. doi:10.1002/jsfa.8705.
- Vriz, R., Moreno, F. J., Koning, F., and Fernandez, A. (2021). Ranking of immunodominant epitopes in celiac disease: Identification of reliable parameters for the safety assessment of innovative food proteins. *Food Chem. Toxicol.* 157, 112584.
- Wang, D.-W., Li, D., Wang, J., Zhao, Y., Wang, Z., Yue, G., et al. (2017). Genome-wide analysis of complex wheat gliadins, the dominant carriers of celiac disease epitopes. *Sci. Rep.* 7:44609.
- Wang, K., Shi, L., Liang, X., Zhao, P., Wang, W., Liu, J., et al. (2022). The gene *TaWOX5* overcomes genotype dependency in wheat genetic transformation. *Nat. plants* 8, 110–117.

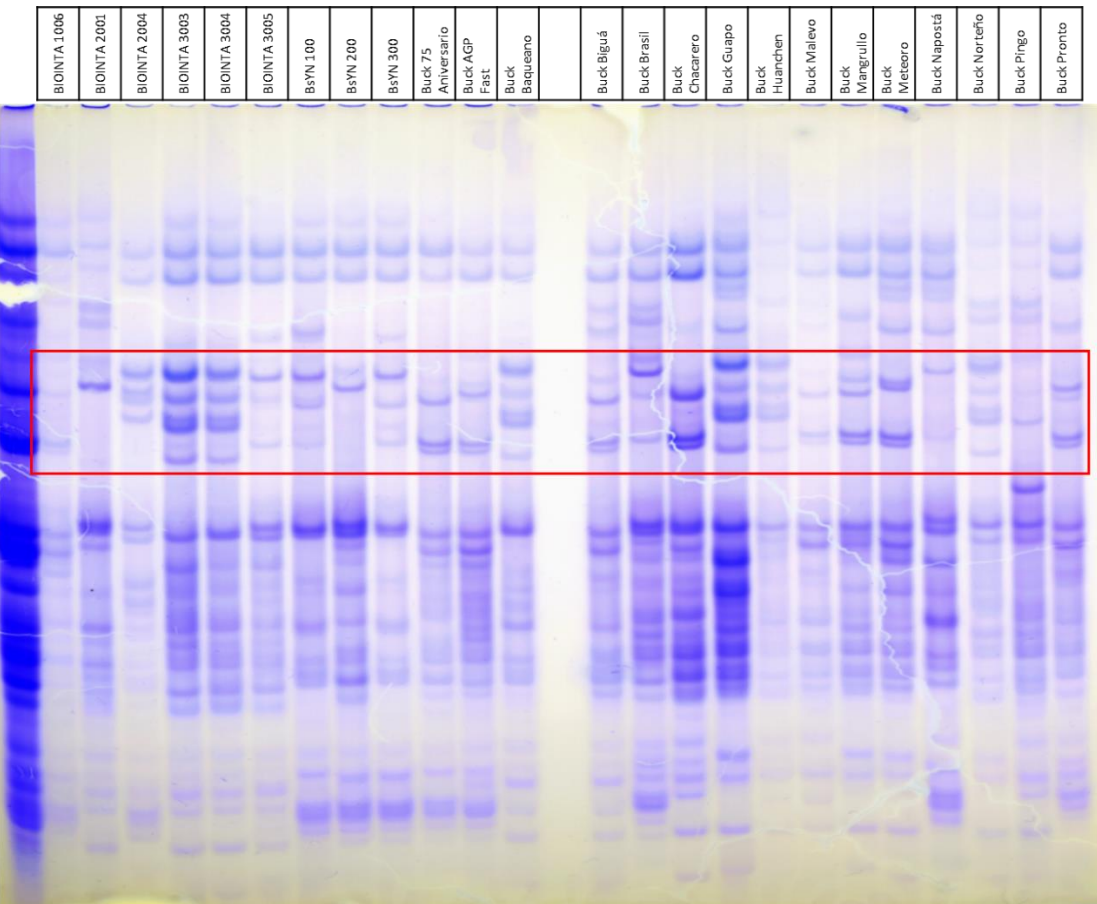
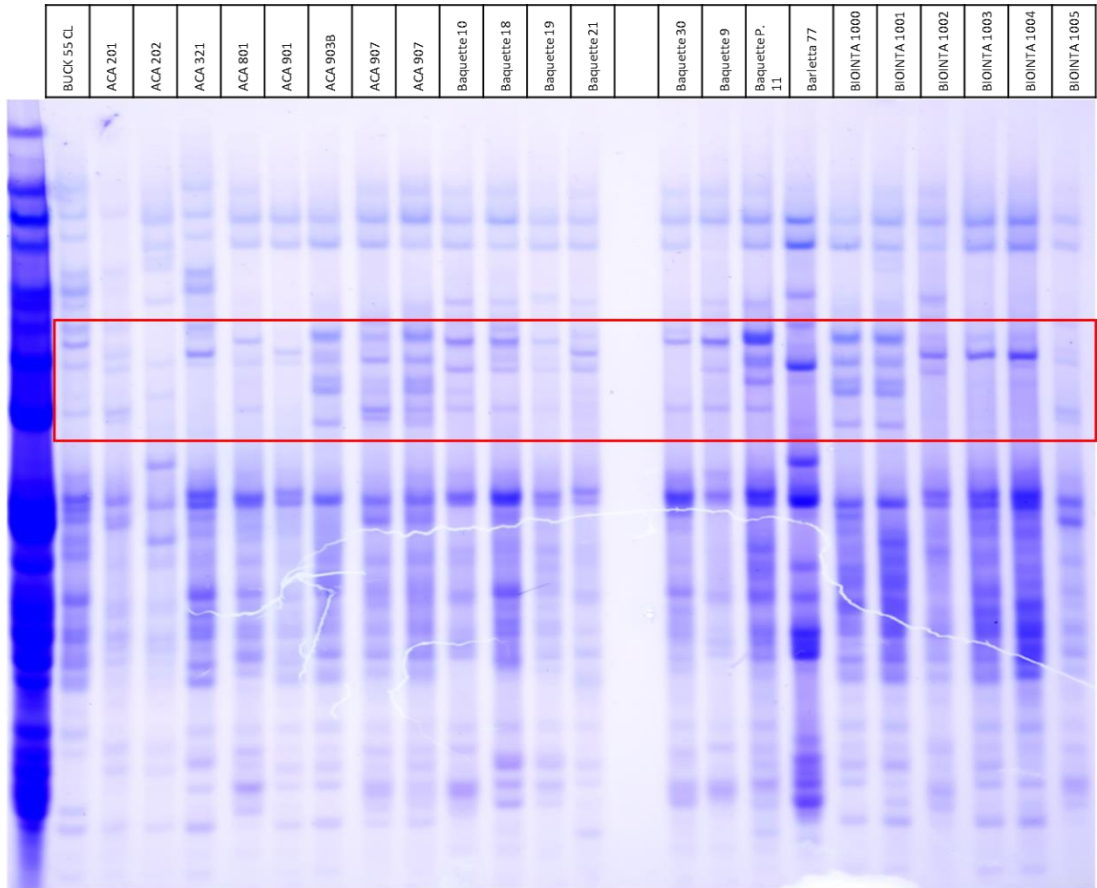


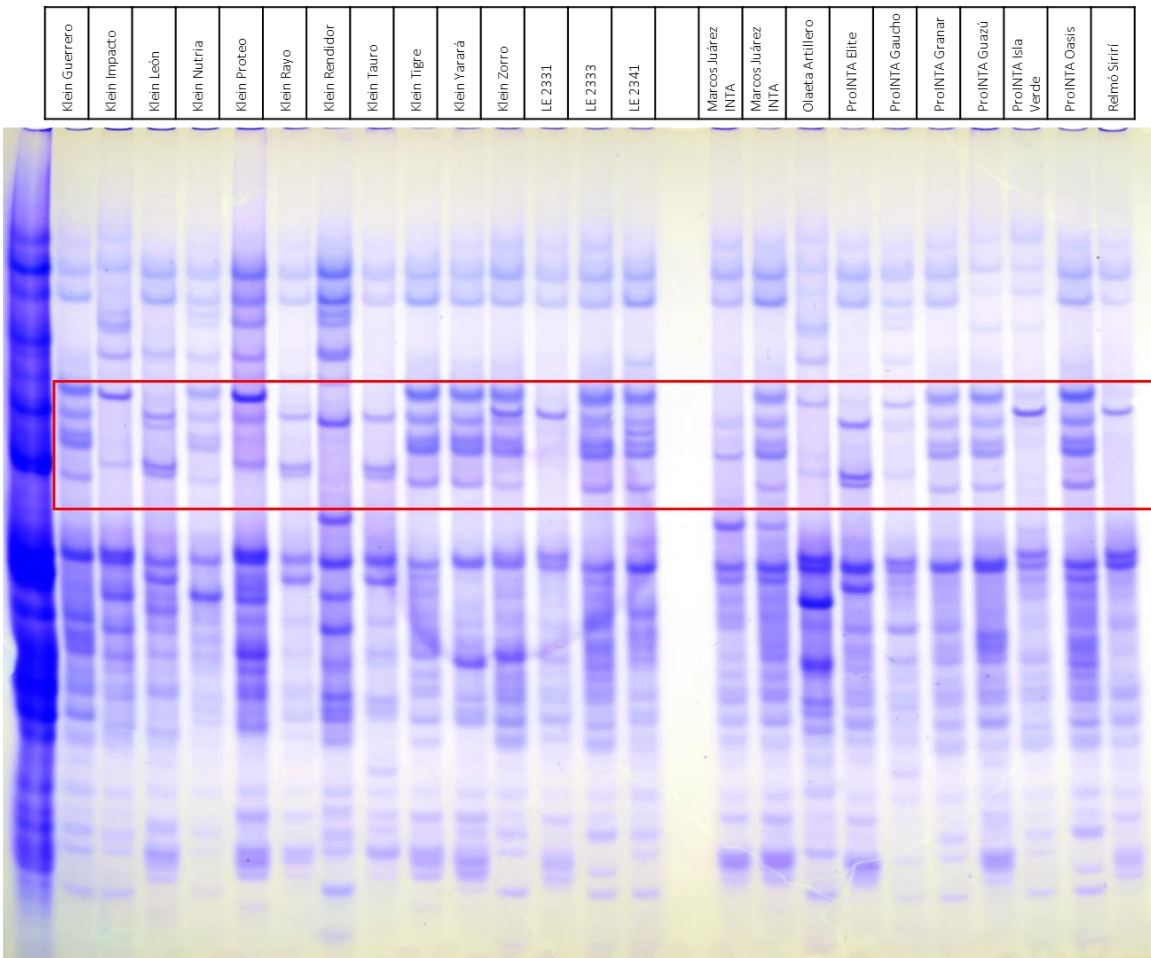
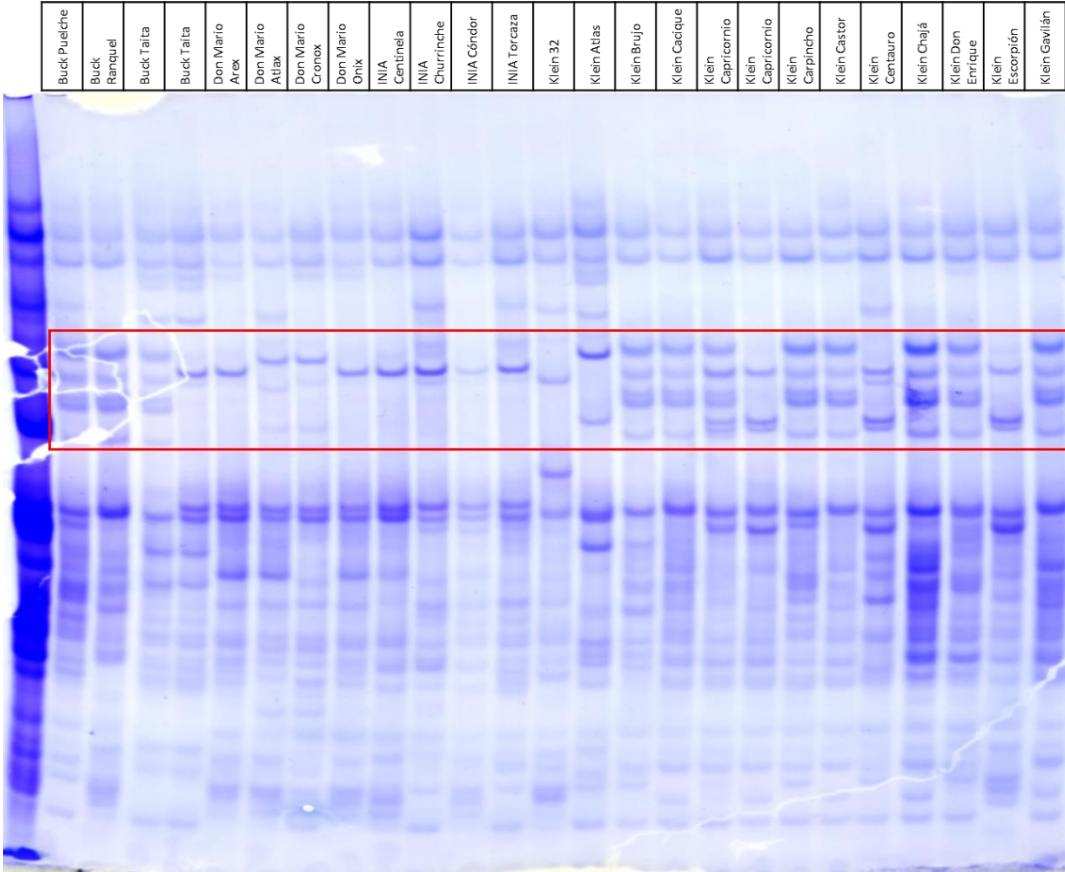
7. Supplementary Materials

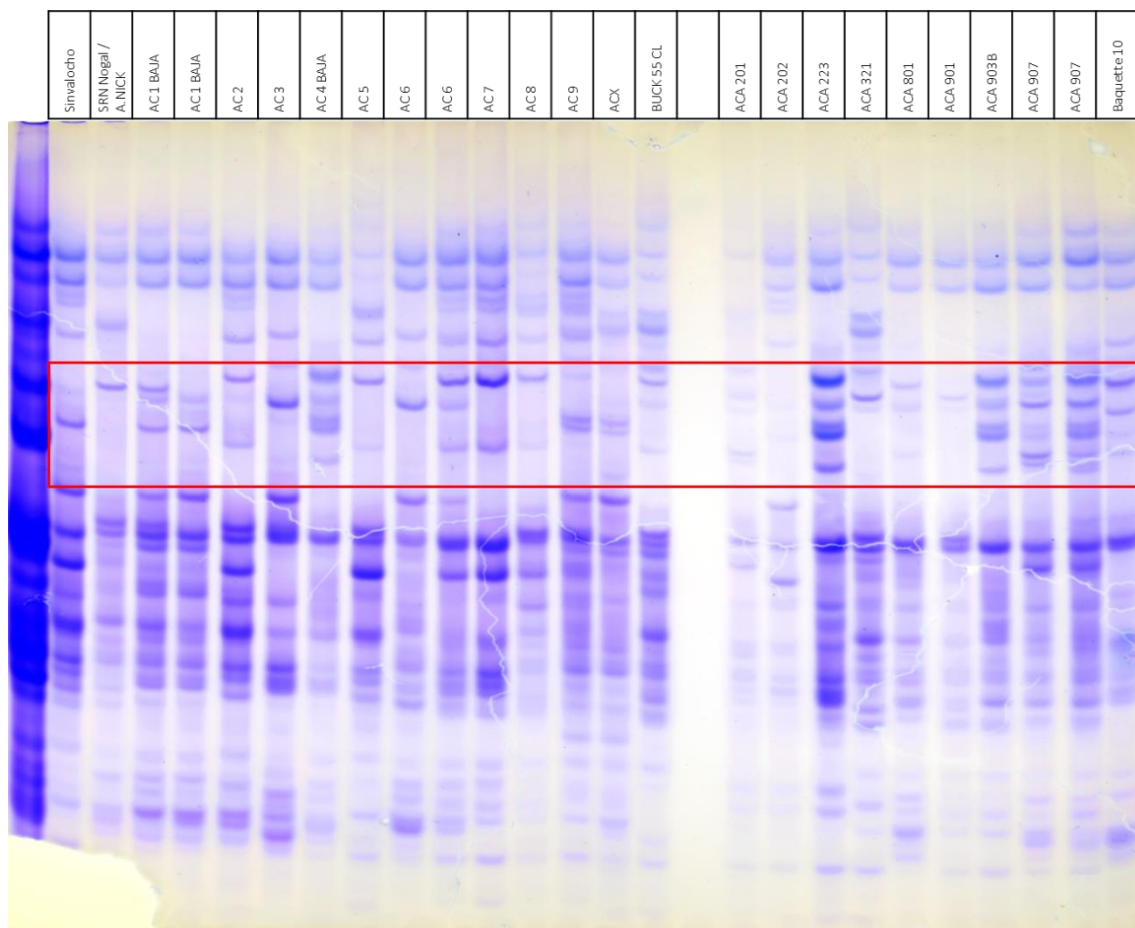


Supplementary Figure 5.1. Composition of A-PAGE gels of lines with rye translocation. The position of the allele *Gli-B1* described in Metakovsky *et al.* (2018) is indicated in a red box. The original gels are in the Supplementary Figure 5.2.









Supplementary Figure 5.2. A-PAGE gels of lines assayed in the present work. The protocol followed for A-PAGE gels of total gliadin proteins is described in Gil-Humanes *et al.* (2012).

Supplementary Table 5.1. Lines of bread wheat with and without rye translocation, durum wheat and tritordeum used in the present study. BW: bread wheat, DW: durum wheat, HT: tritordeum, Y: with rye translocation, N: without rye translocation.

Genotype	Genus	Rye translocation
ACA201	BW	N
ACA202	BW	N
ACA321	BW	N
ACA903B	BW	Y
Antequera	BW	N
ANZA	BW	N
ARAGON03	BW	N
ArturNick	BW	N
Aucan	HT	N
Baquette9	BW	N
BIOINTA1001	BW	Y
BIOINTA1004	BW	N
BIOINTA1005	BW	N



Genotype	Genus	Rye translocation
BIOINTA2004	BW	Y
BuckBaqueano	BW	Y
BuckBrasil	BW	N
BuckHuanchen	BW	Y
BuckNorteno	BW	Y
BuckPuelche	BW	Y
Bulel	HT	N
CATEDRAL	DW	N
DMarioAtlax	BW	N
DonPedro	DW	N
Escacena	BW	N
HT435	HT	N
HT444	HT	N
HT460	HT	N
INIACondor	BW	Y
Jerezano	BW	N
KleinDonEnrique	BW	Y
KleinGavilan	BW	Y
KleinLeon	BW	N
KleinRayo	BW	N
KleinRendidor	BW	N
KleinTigre	BW	Y
MarcosJuarez	BW	N
OlaetaArtillero	BW	N
proINTAGaicho	BW	N
proINTAGranar	BW	Y
proINTAGuazu	BW	Y
proINTAOasis	BW	Y
RelmoSiriri	BW	N
THA85	BW	N
XEIXA	BW	N

Supplementary Table 5.2. List of epitopes searched in the alpha-gliadin and gamma-gliadin/secalin amplicons. The epitopes were collected from the work of Sollid *et al.* (2020). The original and deamidated sequences are represented.

Epitope name	Original sequence	Deamidated sequence
DQ2.5_glia_α1a	PFQPQLPY	PFQPPELPY
DQ2.5_glia_α1b	PYPQLPY	PYPQPELPY
DQ2.5_glia_α2	PQPQLPYPQ	PQPPELPYPQ
DQ2.5_glia_α3	FRPQQPYPQ	FRPEQPYPQ
DQ2.5_glia_γ1	PQQSFPQQQ	PQQSFPEQQ
DQ2.5_glia_γ2	IQPQQPAQL	IQPEQPAQL
DQ2.5_glia_γ2a	FPQQPQQPF	FPEQPEQPF



Epitope name	Original sequence	Deamidated sequence
DQ2.5_glia_γ2b	YPQQPQQPF	YPEQPEQPF
DQ2.5_glia_γ3	QQPQQPYFQ	QQPEQPYFQ
DQ2.5_glia_γ4a	SQPQQQFPQ	SQPEQEFPQ
DQ2.5_glia_γ4b	PQPQQQFPQ	PQPEQEFPQ
DQ2.5_glia_γ4c	QQPQQPFPQ	QQPEQPFPQ
DQ2.5_glia_γ4d	PQPQQPFCQ	PQPEQPFCQ
DQ2.5_glia_γ4e	LQPQQPFPQ	LQPEQPFPQ
DQ2.5_glia_γ5	QQPFPQQPQ	QQPFPEQPQ
DQ2.5_glia_ω1	PFQPPQQPF	PFQPEQPF
DQ2.5_glia_ω2	PQPQQPFPW	PQPEQPFPW
DQ2.5_glut_L1	PFSQQQQPV	PFSEQEQPV
DQ2.5_glut_L2	FSQQQQSPF	FSQQQESPF
DQ2.5_hor_1	PFQPPQQPF	PFQPEQPF
DQ2.5_hor_2	PQPQQPFPQ	PQPEQPFPQ
DQ2.5_hor_3a	PIPQQPQPY	PIPEQPQPY
DQ2.5_hor_3b	PYPQQPQPY	PYPEQPQPY
DQ2.5_sec1	PFQPPQQPF	PFQPEQPF
DQ2.5_sec2	PQPQQPFPQ	PQPEQPFPQ
DQ2.5_sec3	PFQPPQFQI	PFPEQPQFI
DQ2.5_ave_1a	PYPEQQEPF	PYPEQEEPF
DQ2.5_ave_1b	PYPEQQQPF	PYPEQEQPF
DQ2.5_ave_1c	PYPEQQQPI	PYPEQEQPI
DQ2.2_glut_L1	PFSQQQQPV	PFSEQEQPV
DQ2.2_glia_α1	QGSVQPQQL	EGSVQPQEL
DQ2.2_glia_α2	QYSQPQQPI	QYSQPEQPI
DQ8_glia_α1	QGSFQPSQQ	EGSFQPSQE
DQ8_glia_γ1a	QQPQQPFPQ	EQPQQPFPQ
DQ8_glia_γ1b	QQPQQPYFQ	EQPQQPYFQ
DQ8_glia_γ2	PQQSFPQQQ	PQQSFPEQE
DQ8_glut_H1	QGYYPPTSPQ	QGYYPPTSPQ
DQ8.5_glia_α1	QGSFQPSQQ	EGSFQPSQE
DQ8.5_glia_γ1	PQQSFPQQQ	PQQSFPEQE
DQ8.5_glut_H1	QGYYPPTSPQ	QGYYPPTSPQ



Supplementary Table 5.3. Abundance of alpha-gliadin amplicons per line. The DNA and peptide sequence of each amplicon and the number of the CD epitopes matched are indicated.

Amplicon	Amp_131	Amp_116	Amp_112	Amp_107	Amp_68	Amp_38	Amp_28	AJ133607	N_stops_codons
ACA201	0	0	0	0	0	0	0	0	0
ACA202	0.33	0	0	4.87	0	0.38	0	0	0
ACA321	0.38	0	0	4.42	0	0.43	0	0	0
ACA903B	0.27	0	0	2.87	0	0	0	0	0
Antequera	0	0	0	3.08	0.28	0	0.91	0	0
ANZA	0.15	0	0	3.09	0	0.29	0	0	0
ARAGON03	0	0	0	3.3	0.26	0.27	0.89	0	0
ArturNick	0.89	0	0	1.84	0	0	0	2.24	0
Aucan	0.84	0	0	1.92	0	0	0	2.31	0
Baquette9	0.72	0	0	0.44	0	0	0	5.07	0
BIOINTA1001	0.89	0	0	2.59	0.28	0	0.78	2.81	0
BIOINTA1004	0	0	0	2.9	0	0.3	0	0	0
BIOINTA1005	0	0	0	3.08	0	0.32	0	0	0
BIOINTA2004	0.4	0	0	1.94	0	0	0.73	2.19	0
BuckBaqueano	0.85	0	0	2.46	0.31	0	0.87	2.84	0
BuckBrasil	0	0	0	4.73	0	0.5	0	0	0
BuckHuanchen	0.47	0	0	1.99	0	0	0.6	2.08	0
BuckNorteno	0	0	0	2.87	0	0	0	0	0
BuckPueiche	0	0	0	3.23	0	0.27	0	0	0
Bulel	0	0	0	3.22	0	0.28	0	0	0
CATEDRAL	0	2.51	1.33	0.75	0	0	0	0	0
DMarioAtlix	0.31	2.35	1.4	0.71	0	0	0	0	0
DonPedro	0.55	0	0	1.98	0	0	0	2.33	0
Escacena	0.29	2.58	1.48	0.76	0	0	0	0	0
HT435	0	0	0	3.38	0	0.27	0	0	0
HT444	0	2.43	1.33	0.73	0	0	0	0	0
HT460	0	2.54	1.45	0.71	0	0	0	0	0
INIACondor	0.77	0	0	0.4	0	0	0	5.53	0
Jerezano	0.6	0	0	1.9	0	0	0.61	2.17	0
KleinDonEnrique	0	0	0	3.32	0	0.27	0.87	0	0
KleinGavilan	0	0	0	3.19	0	0.3	0	0	0
KleinLeon	0	0	0	3.23	0	0.28	0	0	0
KleinRayo	0.53	0	0	1.89	0	0	0	2.21	0
KleinRendidor	0.53	0	0	1.96	0	0	0.48	2.17	0
KleinTigre	0.28	0	0	4.81	0	0.42	0	0	0
MarcosJuarez	0.62	0	0	1.77	0	0	0.51	2.16	0
OlaetaArtillero	0.84	0	0	2.19	0.14	0	0.63	2.51	0
proINTAGaicho	0.29	0	0	4.54	0	0.4	0	0	0
proINTAGranar	0	0	0	2.96	0	0.28	0	0	0
proINTAGuazu	0	0	0	3.23	0	0.26	0	0	0
proINTAOasis	0.54	0	0	1.84	0	0	0	2.1	0
RelmoSiriri	0.15	0	0	3.09	0	0	0	0	0
THA85	0.57	0	0	1.9	0	0	0	2.14	0
XEIXA	0.84	0	0	1.87	0	0	0	2.17	0
p3I_43	0	0	0	3.01	0	0.27	0	0	0
DQ2.5_glia_α1a	1	0	1	1	0	0	0	1	0
DQ2.5_glia_α1b	1	1	1	1	0	1	1	0	0
DQ2.5_glia_α2	0	0	0	0	0	1	2	0	0
DQ2.5_glia_α3	1	0	0	0	0	2	3	0	0
total_DQ_epitope	1	1	1	1	0	1	1	0	0
	3	2	1	2	0	5	7	0	0



Amp_193	Amp_174	Amp_167	Amp_146	Amp_143	Amp_141	Amp_140	Amp_135	Amp_134	Amp_132
0	0	2	2	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0.41	0	0	0	0	0	0	0	0	0
0.45	0	0	0	0	0	0	0	0	0
0.41	0	0	0	0	0	0	0	0	0
0.38	0	0	0	0	0	0	0	0	0
0.32	2.19	0	0	0	0	0	0	0	0
0.29	2.29	0	0	0	0	0	0	0	0
0.31	0.41	0	0	0	0	0	0	0	0
0	3.13	0	0	0	0	0	0	0	0
0.38	0	0	0	0	0	0	0	0	0
0.38	0	0	0	0	0	0	0	0	0
0	2.44	0	0	0	0	0	0	0	0
0	4.87	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	2.61	0	0	0	0	0	0	0	0
0.28	0	0	0	0	0	0	0	0	0
0.33	0	0	0	0	0	0	0	0	0
0.37	0	0	0	0	0	0	0	0	0
0.34	0	0.4	0.13	0.67	0.65	2.54	0.34	4.58	2.68
0.36	0	0.5	0.33	0.66	0.67	2.25	0.38	4.75	2.37
0.27	2.38	0	0	0	0	0	0	0	0
0.39	0	0.52	0.31	0.74	0.73	2.25	0.41	4.6	2.55
0.36	0	0	0	0	0	0	0	0	0
0.32	0	0.4	0.28	0.64	0.65	2.49	0.3	4.24	2.84
0.34	0	0.52	0.28	0.62	0.65	2.34	0.3	4.39	2.66
0.36	0.43	0	0	0	0	0	0	0	0
0	2.36	0	0	0	0	0	0	0	0
0.4	0	0	0	0	0	0	0	0	0
0.36	0	0	0	0	0	0	0	0	0
0.36	0	0	0	0	0	0	0	0	0
0	4.04	0	0	0	0	0	0	0	0
0	2.44	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0.28	2.5	0	0	0	0	0	0	0	0
0	2.68	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0.38	0	0	0	0	0	0	0	0	0
0.36	0	0	0	0	0	0	0	0	0
0	2.44	0	0	0	0	0	0	0	0
0.39	0	0	0	0	0	0	0	0	0
0.27	2.54	0	0	0	0	0	0	0	0
0.29	2.4	0	0	0	0	0	0	0	0
0.44	0	0	0	0	0	0	0	0	0
0	1	0	0	1	0	1	0	0	1
0	1	0	0	1	1	1	0	0	1
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	2	1	0	0	1
0	2	1	1	2	3	2	0	0	2



Amp_348	Amp_347	Amp_338	Amp_333	Amp_325	Amp_295	Amp_211	Amp_210	Amp_209	Amp_196
0	1	0	1	0	1	3	0	1	1
0	0	0	0	0.32	0	0	0	0	0
0	0	0	0	0.43	0.27	0	0	0	0
0	0.51	0	0	0	0	0	0	0	0
0	0.57	0	0	0	0.27	0	0	0	0
0	0.5	0	0	0.17	0.26	0	0	0	0
0	0.55	0	0	0	0	0	0	0	0
0.4	0.38	4.24	0	0	0	0	0	0	0
0.43	0.37	4.1	0.27	0	0	0	0	0	0
0.52	0.43	0	0.58	0	0	0	0	0	0
0.46	0	5.82	0.29	0	0	0	0	0	0
0	0.55	0	0	0	0	0	0	0	0
0	0.49	0	0	0	0	0	0	0	0
0.35	0.32	4.31	0	0	0	0	0	0	0
0.42	0	5.91	0.28	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0.4	0.32	4.36	0	0	0	0	0	0	0
0	0.47	0	0	0	0	0	0	0	0
0	0.49	0	0	0	0	0	0	0	0
0	0.5	0	0	0.32	0	7.92	0	0	0
0	0.41	0	0	0	0	0	3.18	2.87	2.72
0	0.44	0	0	0	0	0	3	2.77	2.99
0.38	0.37	4.31	0	0	0	0	0	0	0
0	0.43	0	0	0	0	0	3.09	2.77	3.14
0	0.53	0	0	0	0.29	0	0	0	0
0	0.38	0	0	0	0	0	3.39	2.92	2.61
0	0.4	0	0	0	0	0	3.31	2.66	2.88
0.51	0.45	0	0.57	0	0	0	0	0	0
0.36	0.34	4.29	0	0	0	0	0	0	0
0	0.53	0	0	0	0.26	0	0	0	0
0	0.51	0	0	0	0	0	0	0	0
0	0.47	0	0	0	0	0	0	0	0
0.34	0.3	4.37	0	0	0	0	0	0	0
0.35	0.34	4.38	0	0	0	0	0	0	0
0	0	0	0	0.31	0.27	0	0	0	0
0.41	0.4	4.31	0	0	0	6.11	0	0	0
0.39	0.2	5.12	0.15	0	0	0	0	0	0
0	0	0	0	0.39	0.3	0	0	0	0
0	0.54	0	0	0.14	0	0	0	0	0
0	0.5	0	0	0	0	8.14	0	0	0
0.38	0.33	4.35	0	0	0	0	0	0	0
0	0.5	0	0	0.27	0	0	0	0	0
0.38	0.35	4.22	0	0	0	0	0	0	0
0.37	0.4	4.69	0.26	0	0	0	0	0	0
0	0.57	0	0	0	0.29	0	0	0	0
1	0	0	0	0	0	0	0	0	0
0	0	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	1
1	1	2	1	0	0	0	0	0	1



EF569971	DQ246446	Amp_389	Amp_386	Amp_383	Amp_379	Amp_375	Amp_362	Amp_358	Amp_357
0	0	2	2	0	1	1	1	0	1
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	3.67	4.92	0	0	0
0	0	0	0	0	3.84	5.22	0	0	0
0	0	0	0	0	3.56	4.71	0	0	0
0	0	0	0	0	3.82	5.07	0	0	0
1.66	3.01	1.97	1.96	2.13	2.67	3.7	0.64	0.81	2.23
0	3.34	2.05	1.94	2.17	2.47	3.35	0.61	0.8	2.38
0	4.76	0	2.28	5.37	3.24	4.25	2.81	0.57	2.49
2.63	4.68	2.54	0	2.96	0	0	0.76	0.9	3.11
0	0	0	0	0	3.46	4.48	0	0	0
0	0	0	0	0	3.26	4.1	0	0	0
0	3.46	1.98	1.66	2.04	2.09	2.7	0.58	0.66	2.41
2.67	4.55	2.74	2.52	2.86	0	0	0.73	0.86	2.95
0	0	0	0	0	0	0	0	0	0
2.05	3.62	1.99	0	2.15	2.12	3.12	0.6	0.58	2.26
0	0	0	0	0	3.15	4.3	0	0	0
0	0	0	0	0	3.36	4.31	0	0	0
0	0	0	0	0	3.43	4.56	0	0	0
0	6.68	0	0	0	2.93	4.05	0	0	2.92
0	5.72	0	0	0	3.13	4.12	0	0	2.72
2.02	3.54	2.15	1.85	2.2	2.54	3.19	0.54	0.66	2.34
0	5.91	0	0	0	3.02	4.05	0	0	2.68
0	0	0	0	0	3.46	5.01	0	0	0
0	6.36	0	0	0	2.78	4.08	0	0	2.96
0	6.12	0	0	0	2.89	3.98	0	0	2.79
0	5.06	0	2.38	5.79	3.18	4.13	2.83	0.76	2.62
0	3.41	2.04	1.86	2.13	2.34	3.13	0.59	0.68	2.29
0	0	0	0	0	3.69	5.38	0	0	0
0	0	0	0	0	3.41	4.63	0	0	0
0	0	0	0	0	3.37	4.77	0	0	0
2.06	3.5	2.09	1.78	2.09	2.2	3.04	0.56	0.61	2.34
0	3.52	2.09	1.85	1.99	2.29	3.23	0.61	0.63	2.38
0	0	0	0	0	0	0	0	0	0
0	2.06	2.2	3.82	2.23	2.66	3.66	0.63	0.68	2.34
3.08	3.69	2.23	2.17	2.37	1.27	1.66	0.68	0.85	2.57
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	3.61	4.64	0	0	0
0	0	0	0	0	3.51	4.62	0	0	0
1.96	3.44	2.07	1.88	1.96	2.38	3.3	0.64	0.65	2.27
0	0	0	0	0	3.5	4.64	0	0	0
0	3.52	2.04	2.04	1.98	2.47	3.4	0.69	0.68	2.43
2.96	3.37	2.18	2.11	2.17	2.78	4.02	0.55	0.72	2.48
0	0	0	0	0	3.88	5.31	0	0	0
0	1	1	0	0	0	0	1	1	1
0	1	0	1	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	1	0	1	0	1	1	1	1	1
0	2	0	2	0	1	1	1	1	2



KJ410476.1	KC715927.1	KC715915.1	KC715890.1	JX828313.1	JX828263.1	JX828212.1	JN831389.1	EU018293
0	0	1	0	2	0	0	1	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0.29	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0.45	0	1.45	1.88	0	0.89	0	2.39	0.51
1.87	0	1.43	1.98	0	0.86	0	2.35	0.51
1.95	0	0	1.9	0.38	0	0	2.42	0
0.75	0	1.98	2.6	0	0.88	0	3.13	0.57
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
2.29	0	1.48	1.96	0	0.64	0	2.49	0.46
0.71	0	2.02	0.66	0	0.82	0	3.23	0.52
0	0	0	0	0	0	0.28	0	0
0.55	0	1.43	1.88	0	0.4	0	2.36	0.45
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0.85	0.5	0	0	0	0
0	2.55	0	0.92	0.61	0	0	0	0
0	2.36	0	1.96	0	0.63	0	2.32	0.47
0.52	0	1.35	0.96	0	0	0	0	0
0	2.47	0	0.96	0.58	0	0	0	0
0	0	0	0	0	0	0	0	0
0	2.73	0	0.82	0.48	0	0	0	0
0	2.45	0	0.91	0.62	0	0	0	0
1.81	0	0	2.2	0.67	0	0	2.39	0
2.01	0	1.5	1.95	0	0.55	0	2.41	0.45
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0.48	0	1.42	0.58	0	0.48	0	2.19	0.45
2.02	0	1.41	1.97	0	0.51	0	2.36	0.49
0	0	0	0	0	0	0.28	0	0
0	0	1.42	3.45	0	0.63	0	2.52	0.75
0.3	0	1.73	0.68	0	0.85	0	2.78	0.45
0	0	0	0	0	0	0.28	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0.53	0	1.49	2.01	0	0.6	0	2.27	0.45
0	0	0	0	0	0	0	0	0
2.04	0	1.42	1.9	0	0.56	0	2.44	0.51
0.33	0	1.64	0.63	0	0.71	0	2.57	0.48
0	0	0	0	0	0	0	0	0
0	1	1	1	0	1	0	0	1
0	1	0	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0
0	1	1	1	1	1	1	0	1
0	2	1	2	2	2	3	1	2



TraesCS6A02C049400	TraesCS6A02C048900	TAU51310	TAU51309	LT627589.1	LT627586.1	LT627577.1	KX452085.1
0	0	1	0	0	0	0	0
0.43	1.17	0	8.36	7.88	0.82	0	0
0.36	1.17	0	8.63	7.77	0.86	0	0
0.13	0.43	0	0	1.38	0.4	0	0
0.26	0.53	0	0	1.44	0.41	0	0
0.27	0.61	0	0	1.36	0.43	0	0
0.29	0.42	0	0	1.37	0.34	0	0
0	0	0.52	0	1.02	0.33	0.27	0
0	0	0.5	0	0.94	0.3	0.28	0.47
0	0	0.4	0	1.04	0.27	0	0.56
0	0	0.55	4.62	4.34	0.51	0.28	0.31
0.27	0.66	0	0	1.36	0.43	0	0
0	0.5	0	0	1.56	0.34	0	0
0	0	0.4	0	0.87	0	0	0.54
0	0	0.55	4.86	4.34	0.45	0.28	0.29
0.39	0.53	0	8.81	8.17	0.84	0	0
0	0	0.41	0	0.81	0	0	0.27
0	0.3	0	0	1.04	0.27	0	0
0.27	0.44	0	0	1.31	0.31	0	0
0.26	0.49	0	0	1.36	0.34	0	0
0	0	0	0	1.22	0.32	0	0
0	0	0	0	1.17	0.34	0	0
0	0	0.47	0	0.97	0.27	0	0.27
0	0	0	0	1.18	0.33	0	0
0	0	0	0	1.29	0.36	0	0
0.29	0.47	0	0	1.19	0.31	0	0
0	0	0	0	1.25	0.31	0	0
0	0	0.34	0	1.29	0.33	0	0.48
0	0	0.41	0	0.86	0	0	0.48
0.3	0.42	0	0	1.26	0.36	0	0
0.27	0.31	0	0	1.24	0.3	0	0
0.13	0.4	0	0	1.27	0.3	0	0
0	0	0.45	0	0.8	0	0	0.27
0	0	0.52	0	0.81	0	0.27	0.51
0.4	0.53	0	9.15	8.27	0.93	0	0
0	0	0.49	0	0.9	0.27	0	0
0	0	0.47	2.34	2.5	0.27	0.15	0
0.42	0.7	0	8.4	7.21	0.8	0	0
0.26	0.44	0	0	1.4	0.38	0	0
0.27	0.47	0	0	1.36	0.37	0	0
0	0	0.48	0	0.85	0	0	0.27
0.27	0.54	0	0	1.42	0.39	0	0
0	0	0.5	0	0.84	0.13	0	0.49
0	0	0.57	0	0.96	0.28	0.27	0
0.27	0.45	0	0	1.38	0.43	0	0
1	0	1	0	0	0	1	0
1	0	1	0	0	0	1	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
1	0	1	0	0	0	1	0
2	0	2	0	0	0	2	0



TraesCS6B02G065578	TraesCS6B02G065536	TraesCS6A02G049700	TraesCS6A02G049600	TraesCS6A02G049500
0	0	0	0	0
0	0	6.23	10.35	1.33
0	0	6.16	10.07	1.29
0.27	4.78	3.88	5.89	1.05
0.29	4.87	4.06	6.23	1.09
0.17	4.29	3.94	6.36	0.99
0	4.88	4.38	6.5	1.02
0	2.68	0	10.45	0
0	2.53	0	10.7	0
0	3.84	0	13.95	0
0	0	0	13.3	0
0.27	4.3	3.84	6.2	0.93
0	4.61	4.32	6.3	0.81
0	2.8	0	10.07	0
0	0	0	12.9	0
0	0	6	9	1.16
0	2.96	0	10.06	0
0	4.76	3.75	5.18	0.86
0	4.38	4.13	5.81	0.86
0	5.01	4.04	5.73	0.86
0	3.59	0	13.05	0
0	3.09	0	12.35	0
0	2.62	0	10.3	0
0.28	3.21	0	13.05	0
0	4.51	4.19	6.33	1.1
0	3.72	0	13.35	0
0	3.45	0	13.05	0
0	3.58	0	14.95	0
0	2.75	0	9.75	0
0	4.96	4.25	5.97	1.1
0	4.62	4.27	5.71	0.88
0	4.8	4.08	5.6	0.89
0	2.79	0	9.54	0
0	2.75	0	9.77	0
0	0	6.2	9.31	1.22
0	2.9	0	10.35	0
0	1.34	0	13.45	0
0	0	5.68	9.34	1.14
0.28	5.01	4.13	6.11	0.85
0	4.45	3.76	5.9	0.89
0	2.81	0	9.61	0
0.27	4.43	3.92	6.37	0.87
0	2.94	0	10.08	0
0	2.83	0	11.35	0
0.31	4.25	3.94	6.46	1.1
0	0	1	1	0
0	0	1	1	1
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	1	1
0	0	1	2	2



TraesCSU02G108100	TraesCSU02G108000	TraesCS6B02G086522	TraesCS6B02G086500	TraesCS6B02G0865856
0	0	4	0	0
3.98	4.85	11.65	6.78	0
3.9	4.88	11.7	6.75	0
2.21	2.65	7.64	3.62	0.3
2.44	2.38	8.37	3.96	0.34
2.26	2.66	7.17	3.54	0.34
2.4	2.86	8.47	3.95	0
1.26	1.58	5.2	2.3	0
1.18	1.68	5.29	2.36	0
0	0	6.56	2.85	0.29
1.77	2.68	7.39	3.36	0
2.4	2.83	7.16	3.62	0.33
2.83	3.09	7.47	3.78	0
1.48	1.82	5.01	2.28	0
1.92	2.9	6.59	3.39	0
4.19	5.66	11.45	6.56	0
1.31	1.85	5.38	2.17	0
2.35	2.84	8	3.58	0
2.48	3.2	6.99	3.55	0
2.63	3.31	0	0	0.29
0	0	6.16	2.96	0
0	0	11.35	0	0
1.38	2	4.88	2.33	0
0	0	5.93	2.75	0
2.36	2.66	8.63	3.58	0
0	0	6.21	2.96	0
0	0	5.84	2.77	0
0	0	6.1	2.81	0
1.31	1.91	4.8	2.21	0
2.25	2.54	9.01	3.8	0
2.61	2.8	7.64	3.87	0
2.45	3.36	7.71	3.59	0
1.4	2.02	5.06	2.23	0.26
1.22	1.79	5.36	2.17	0
4.21	5.48	12.5	6.64	0
1.24	1.75	0	0	0
1.45	2.19	5.64	2.8	0
4.09	4.97	11.95	6.42	0
2.46	3.29	7.73	3.94	0.29
2.47	2.9	0	0	0.29
1.31	1.85	5.26	2.28	0
2.39	2.95	7.32	3.66	0.31
1.17	1.88	5.46	2.41	0
0	1.8	6.05	2.45	0
2.27	2.53	7.75	3.83	0.26
0	0	0	1	0
1	0	0	0	0
0	0	0	0	0
1	0	0	0	0
1	0	0	0	0
3	0	0	0	0



TraesCSU02G153800	TraesCSU02G149946	TraesCSU02G108700	TraesCSU02G108363	TraesCSU02G108300
0	0	0	0	0
1.02	0	6.04	8.24	0.37
0.97	0	5.89	8.66	0.39
0.52	28.5	3.42	4.67	0
0.48	27	4.11	5.29	0
0.52	27.25	4.01	5.03	0.27
0.44	28.8	2.91	5.25	0
0	17.5	2.25	2.92	0
0	17.7	2.38	2.9	0
0	25	0	0	0
0	0	3.29	4.09	0
0.56	27.2	4.1	5.2	0
0.53	30.55	3.84	5.67	0
0	17.75	2.6	3.08	0
0	0	3.32	4.14	0
0.52	0	6.25	9.39	0.35
0	20.1	2.44	2.87	0
0.28	30.4	4.39	5.18	0.27
0.36	29.9	3.53	5.27	0
0.43	31.65	3.92	5.66	0
0	22.05	0	0	0
0	21.05	0	0	0
0	16.6	2.56	3.07	0
0	21.8	0	0	0
0.43	27.75	3.49	4.81	0.27
0	22.2	0	0	0
0	22.95	0	0	0
0	23.2	0	0	0
0	20.05	2.58	2.94	0
0.39	28.8	3.25	4.83	0.27
0.37	29.6	3.96	5.75	0.3
0.36	29.5	3	5.26	0.28
0	19.15	2.32	3	0
0	18.95	2.12	2.78	0
0.59	0	5.17	8.43	0.42
0	19.4	2.4	2.88	0
0	9.45	2.93	3.21	0
0.75	0	6.94	8.81	0.39
0.5	26.85	3.73	5.36	0
0.43	29.65	4.37	5.11	0.27
0	19.5	2.17	2.82	0
0.46	28.25	3.45	5.02	0
0	19.5	2.33	2.68	0
0	19.6	2.04	2.99	0
0.5	27.2	4.05	4.94	0.27
0	0	0	0	0
0	0	0	1	1
0	0	0	0	1
0	0	0	1	2
0	0	0	1	0
0	0	0	4.94	4



X54517.1	TraesCSU02G255529	TraesCSU02G220200	TraesCSU02G202177	TraesCSU02G188800	TraesCSU02G160200
0	0	0	2	0	0
0	0	12.1	0	1.47	1.06
0	0	11.85	0	1.46	0.99
0	3.09	6.27	3.34	2.31	0.55
0	2.91	4.74	2.8	0.87	0.54
0	3.07	7.17	3.08	1.53	0.53
0	2.48	5.01	2.44	0.86	0.37
2.13	1.8	4.2	1.84	0.86	0
1.76	1.97	2.86	1.98	1.33	0
0	2.14	0	2.27	0	0
2.39	0	5.83	0	0.78	0
0	3.47	7.51	3.77	0.94	0.57
0	3.11	7.05	0	1.31	0.64
1.83	4.21	2.94	2.15	0.68	0
2.54	0	3.98	0	0.88	0
0	0	12.9	0	1.7	0.54
1.72	2.25	4.25	2.27	0.51	0
0	3.6	7.11	3.64	0.81	0.28
0	3.27	7.37	3.28	0.9	0.46
0	3.6	7.59	0	1.86	0.4
0	1.9	0	2.17	0	0
0	2.36	0	2.49	0	0
2.1	2.44	3.28	2.62	1.05	0
0	2.36	0	2.44	0	0
0	2.89	6.88	3.28	0.75	0.43
0	1.9	0	2.05	0	0
0	2.27	0	2.3	0	0
0	2.06	0	2.04	0	0
1.71	2.34	2.79	2.55	0.58	0
0	2.82	4.66	2.91	0.78	0.36
0	3.63	7.84	0.34	0.91	0.4
0	3.1	7.17	3.06	0.86	0.37
1.68	2.11	4.14	2.25	1	0
1.73	4.08	2.55	2.12	0.48	0
0	0	12.15	0	1.36	0.72
1.8	2.09	2.75	2.23	0.5	0
2.19	0.98	4.75	1.03	0.65	0
0	0	12.2	0	2.9	0.74
0	3.27	5.33	3.6	1.73	0.51
0	3.54	7.17	3.51	1.7	0.47
1.8	2.1	4	2.14	0.86	0
0	3.27	7.1	3.38	0.85	0.51
1.8	1.96	2.72	2.11	0.47	0
1.79	1.97	0	2.03	0	0
0	3.06	7.19	1.77	1.55	0.49
0	0	0	0	0	0
1	0	1	0	1	0
0	0	1	0	2	0
0	0	2	0	3	0
1	0	0	0	1	0
2	0	4	0	7	0



Supplementary Table 5.4. Abundance of gamma-gliadin/secalin amplicons per line. The DNA and peptide sequence of each amplicon and the number of the CD epitopes matched are indicated.

Amplicon	BW-32	BW-30	BW-26	BW-9	AT7_3-6
Secalins	N	N	N	N	N
N_stops_codons	0	0	0	0	0
ACA201	3.41	0.46	2.59	0.73	7.21
ACA202	0.39	0.32	0	0	0
ACA321	1.98	0.52	1.56	0.28	4.6
ACA903B	0	1.69	0	0	0
Antequera	0	0.47	0	0	0
ANZA	1.66	0	0	0	0
ARAGON03	0.26	0.31	2.88	0.55	6.18
ArturNick	1.96	0.44	1.59	0.33	5.56
Aucan	3.01	0	0	0	0
Baquette9	0.26	0.4	0	0	7.04
BIOINTA1001	0	0.43	0	0	0
BIOINTA1004	2.25	0.27	0	0	0
BIOINTA1005	2.38	0	0	0	0
BIOINTA2004	2.92	0.32	0	0	0
BuckBaqueano	0	0.92	0	0	0
BuckBrasil	0	0.42	0	0.46	6.42
BuckHuanchen	0	1.77	0.83	1.02	14.8
BuckNorteno	0	0.82	6.42	1.04	15.2
BuckPuelche	0	1.01	5.65	0	0
Bulel	0	0	0	0	0
CATEDRAL	0	0	0	0	0
DMarioAtlas	0.3	0.27	0	0	0
DonPedro	0	0	0	0	0
Escacena	2.29	0.33	0	0	0
HT435	0	0	0	0	0
HT444	0	0	0	0	0
HT460	0	0	0	0	0
INIACondor	0	0.72	0	0	0
Jerezano	0	0.39	0	0	0
KleinDonEnrique	0	0.64	0	0	0
KleinGavilan	0	0.77	0	0	0
KleinLeon	3.3	0	0	0	0
KleinRayo	2.45	0.32	0	0	0
KleinRendidor	0.27	0.39	0	0	0
KleinTigre	0	1.83	0.32	0	0
MarcosJuauez	0	0.41	0	0	0
OlaetaArtillero	0.38	0.53	0	0.88	6.88
proINTAGaicho	0	0	0.83	0	0
proINTAGranar	0	0.35	0	0	0
proINTAGuazu	0	1.27	0.3	0	0
proINTAOasis	0	1.4	0.29	0.73	15
ReinoSiriri	1.34	0.39	0	0	0
THA85	0	0.52	0	0	0
XEIXA	0.39	0.85	2.52	0.35	6.49
DQ2.5_glia_γ2a	1	2	1	0	0
DQ2.5_glia_γ4e	0	1	0	0	0
DQ2.5_sec1	0	0	0	0	0
DQ2.5_glia_γ4d	0	0	0	0	0
DQ2.5_sec2	0	0	0	0	0
DQ2.5_glia_γ4b	1	0	0	1	1
DQ2.5_glia_γ1	1	1	0	1	1
DQ2.5_glia_γ5	0	2	2	1	1
DQ2.5_glia_γ4c	4	5	6	2	2
DQ2.5_glia_γ3	0	0	0	0	0
DQ2.5_glia_γ4a	0	0	0	0	0
DQ2.5_glia_γ2b	0	0	0	0	0
total_DQ_epitopes	7	11	9	5	5



T545-7	FJ006594	EU863818	D64-2_48	D64-2_45	D64-2_38	D64-2_27
N	N	N	N	N	N	N
0	0	0	0	0	1	1
0	0.33	0	0	0	0	0
0	2.8	0	3.54	0.52	0.93	0
0	0	0	0	0	0.3	0
0	11.65	0	0	0	0	0
0	9.4	0	2.69	0	0.58	0
0	4.8	0	0	0	0	0
0	0.31	0	2.94	0	1.23	0
0	0.33	0	0	0	0.41	0
0	3.62	0	0	0	4.54	9.96
0	8.88	7.82	3.56	0.17	1.25	0
0	5.06	0	0	0	0	11.35
0	9.88	0	0	0	0.35	0
0	8.95	0	0	0	0	0
0	0	0	0	0	0.38	0
0.27	21.85	0	0	0.57	0	0
0	8.42	7.77	3.48	0	0.73	0
0	0.68	0	0	0	0	0
0	0.38	0	0	0	0	0
0	0.87	0	0	0	0	0
0	7.62	0	0.36	0	0.53	18.3
0	9.08	0	0.7	0	0.85	21.55
0	1.98	2.12	1.24	0	0.29	0
0	8.21	0	0.66	0	0.73	16.45
0	8.81	0	0	0.27	0.31	0
0	7.82	0	0.5	0	0.6	19.25
0	6.51	0	0.56	0	0.55	21.45
0	5.82	0	0.33	0	0.76	18.05
0.28	0.31	0	0	0	0	0
0	0	0	3.58	0	0.97	0
0	21.1	0	0	0.25	0	0
0	12.95	0	0	0	0	0
0	6.53	7.29	0	0	0	0
0	6.28	0	0	0	0	0
0	2.92	0	4.32	0	1.12	0
0	0	0	0	0	0	0
0	0	0	3.4	0	0.82	0
0	3.8	0	4.95	0.47	0.63	0
0	0	0	0.79	0	0.29	0
0	14.9	0	0	0	0	0
0.38	18.45	0	0	0.31	0	0
0.4	0.29	0	0	0	0	0
0	4.37	0	0	0.27	0	0
0	7.68	0	0.26	0	0.39	0
0	0.49	0	3.94	0	1.4	0
3	1	1	0	0	0	0
1	0	0	0	0	1	0
0	0	0	0	1	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	1	0	0	0	2
1	1	1	1	0	0	1
3	0	0	0	1	1	0
6	4	4	3	5	2	3
0	0	0	0	0	2	0
0	0	0	0	0	0	0
0	0	0	0	0	2	0
14	6	7	4	7	8	6



HQ875989.1	HQ875980.1	HQ875965.1	T545-33	T545-20	T545-19	T545-9
Y	Y	Y	N	N	N	N
2	0	0	0	0	0	0
0	0	0	0	0.41	4.3	5.26
0	0	0	0	0	0	4.47
0	0	0	0	0	0	3.66
0	2.21	0.36	0	0	0	10.16
0	0	0	0	0	0	4.95
0	0	0	0	0	0	3.68
0	0	0	0	0	0	5.54
0	0	0	0	0	0	5.14
0	0	0	0	0	0	0
0	0	0	0	0	0	6.57
0.28	1.5	0.27	0	0	0	7.66
0	0	0	0	0	0	5.02
0	0	0	5.94	0	1.47	3.77
0	8.45	0	0	0	0	4.18
0.28	2.52	0.3	0	0	0	12.05
0	0	0	0	0	0	6.15
0.26	2.1	0	0	0	0	10.27
0	3.07	0.38	0	0	0	11.5
0.29	3.07	0	0	0	0	10.7
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	1.55
0	0	0	0	0	0	0
0	0	0	0	0	0	4.37
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0.39	2.06	0.26	0	0	0	8.65
0	0	0	0	0	0	5.24
0	1.82	0.3	0	0	0	11.45
0	1.94	0.27	0	0	0	6.25
0	0	0	5.48	0.66	3.6	4.8
0	0	0	4.96	0.42	2.53	4.13
0	0	0	0	0	0	6.54
0.26	2.45	0.58	0	0	0	11.8
0	0	0	0	0	0	4.09
0	0	0	0	0	0	4.98
0	0	0	0	0	0	1.75
0	1.48	0.26	0	0	0	5.76
0	2.29	0.28	0	0	0	10.03
0.26	3.47	0.39	0	0	0	12.9
0	0	0	0	0	0	3.21
0	0	0	0	0	0	4.87
0	0	0	0	0.13	0	6.26
0	0	0	0	0	0	2
0	0	0	0	1	1	1
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	1	0	0	0	0
0	0	1	0	0	1	0
0	0	0	1	1	1	1
0	0	0	0	1	1	3
0	0	1	2	3	2	5
0	0	0	1	2	0	0
0	0	0	0	0	0	0
0	0	0	1	2	0	0
0	0	3	5	10	6	12



Q0871775	Q0871774	Q0871770	HQ875998.1	HQ875997.1	HQ875995.1	HQ875991.1
N	N	N	Y	Y	Y	Y
0	0	0	0	0	1	1
0	0.33	7.29	0	0	0	0
3.48	0	5.91	0	0	0	0
2.51	0	6.08	0	0	0	0
0	0.36	11	0.27	0	0.8	1.79
2.37	0.58	0	0	0	0	0
2.18	0	3.4	0	0	0	0
4.17	0	5.03	0	0	0	0
3.1	0	4.91	0	0	0	0
0	0	0	0	0	0	0
3.7	0.28	6.38	0	0	0	0
0	0	6.92	0	0	0.37	0.78
2.5	0.52	0	0	0	0	0
0	0	3.22	0	0	0	0
3.51	0.27	0	0.27	0.28	0	0.94
0	0	10.55	0.26	0	0.35	0.68
2.86	0.3	7.17	0	0	0	0
0	0.46	12.3	0	0	0.54	1.01
0.35	0	12.8	0	0.26	0.68	0.82
0	0.35	14.9	0	0.28	0.33	1.13
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0.96	0	1.97	0	0	0	0
0	0	0	0	0	0	0
2.84	0	5.5	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	10.38	0	0	0.39	0.87
3.92	0.39	0	0	0	0	0
0	0.79	0	0.28	0.28	0.65	0.98
0	0.27	6.58	0.27	0	0.37	0.74
0	0.26	0	0	0	0	0
0	0	3.77	0	0	0	0
4.06	0.59	0	0	0	0	0
0	0.65	0	0	0	0.42	1.59
3.34	0.3	6.83	0	0	0	0
3.78	0	6.71	0	0	0	0
1.06	0	1.86	0	0	0	0
0	0	7.24	0.36	0	0.26	0.91
0	0.76	0	0.41	0	0.28	1.01
0	0.31	14.7	0	0.3	1.17	0.9
1.73	0.26	3.04	0	0	0	0
0	0	5.15	0	0	0	0
3.7	0	6.95	0	0	0	0
0	0	1	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	1	0	0	0	0
1	1	1	0	0	0	0
1	0	0	0	0	0	0
2	2	2	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
5	3	5	0	0	0	0



TraesCS1A02G007344	Tha1_2-13	T544-58	T544-55	JX679682	JX679676
N	N	N	N	N	N
2	0	0	0	0	0
4.52	0	0.83	0	15.05	0
0	0	0	0	23.35	0
3.54	2.92	0.42	0	11.5	0
5.72	0	0	0	32.3	0
3.77	0	0	0	15.15	4.29
3.17	2.6	0	0	18.8	0
4.84	0	0.76	0	17.3	0
4.26	2.64	0.33	0	11.65	0
0	0	0	0	0	0
0	0	0	0	13.65	0
0	0	0	0	36.95	0
3.69	3.38	0	0	10.2	4.03
2.92	0	0	10.08	19.2	0
0	3.48	0	0	10.85	3.4
7.81	0	0	0	23.8	0
0	0	0	0	14.65	0
14.8	0	6.63	0	30.75	0
9.53	0	1.64	0	33.65	0
8.88	0	0.92	0	30	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	41.45	0
4.29	2.76	0	0	17	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
5.13	0	5.09	0	36.6	0
4.89	0	2.68	0	17.05	4.48
7.58	0	0	0	26.35	7.07
6.91	0	0	0	33.05	0
0	0	0	10.35	11.3	4.16
3.62	0	0	10	18.3	0
0	0	0	0	13.35	5.4
9.07	0	6.45	0	28.3	0
4.17	0	3.5	0	23.05	0
0	0	0	0	16.25	0
1.14	0	0	0	44.2	0
6.84	0	0	0	35.7	0
8.83	0	0	0	27.1	6.8
9.42	0	7.57	0	29.3	0
2.5	1.39	0	0	17.9	0
4.89	0	0	0	25.05	0
5.58	0	0.46	0	15.8	0
1	1	1	0	0	1
0	0	0	0	1	0
0	1	0	0	0	1
0	1	0	0	0	1
0	0	0	0	0	0
0	1	0	0	0	0
0	1	0	1	1	1
0	0	2	0	1	0
1	4	6	2	1	3
0	0	0	1	1	0
0	0	0	0	1	0
0	0	0	1	1	0
2	9	9	5	7	7



TraesCS1B02G011000	TraesCS1B02G010900	TraesCS1B02G010800	TraesCS1B02G010700	TraesCS1B02G010400
N	N	N	N	N
0	1	0	1	0
41.65	0	0.41	0	4.33
30.3	0	0.67	0	3.87
29.8	0.79	0.84	3.26	0.33
0	0	0	0	0
36.1	0.27	0.92	3.35	3.63
23.8	0.72	0.56	2.42	0.34
40	0.35	1.55	0	4.39
28.95	1.11	1.09	3.18	0.46
39.5	0.9	1.45	3.2	0
29.2	0.27	1.65	2.94	4.32
0	0	0	0	0
26.45	0.62	0.48	2.8	0.29
24.45	0	0.47	0	2.61
27.35	1.24	0.95	2.97	0.47
0	0	0	0	0
32.4	0	0.58	3.67	3.82
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
43.35	0	2.21	0	0
37.4	0	2.73	0	0
11.3	0	0.27	1.11	1.31
39.95	0	2.4	0	0
18.2	0.64	0.87	3.25	0.41
40.6	0	1.83	0	0
39.7	0	2.24	0	0
42.35	0	1.89	0	0
0	0	0	0	0
36.1	0.28	1.22	3.28	4.07
0	0	0	0	0
0	0	0	0	0
27.15	0	1.02	0	5.08
23	0	0.48	0	3.33
43.25	0	0.66	0	4.61
0	0	0	0	0
27	0	0.52	0	3.29
37.9	0.3	1.4	4.32	4.87
12.35	0	0.16	1.5	1.46
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
24.9	0.99	0.27	1.91	0.3
18.9	0	0.68	0	0
37.6	0.32	1.14	0	3.61
0	0	3	2	0
0	1	0	0	0
0	0	0	0	1
0	0	0	0	1
0	0	0	0	0
0	0	0	0	0
1	1	1	1	1
1	1	1	0	1
0	1	2	2	0
2	2	6	7	3
1	2	0	0	0
0	0	0	0	0
1	2	0	0	0
5	9	13	12	7



Zollen_3-9	Zollen_3-7	Zollen_2-11	TreesCS1D02G001000
N	N	N	N
1	0	0	0
0	0	0	0.91
0.29	0	0.42	18.75
0	24.2	0	0.91
0	0	0	21.65
0	0	0	11.5
0	17.4	0	14.45
0.54	0	0.29	0.55
0	21.55	0	0.97
6.02	27.8	0	0
0.31	0	0.28	1.07
0	0	0	28.45
0	19.85	0	7.44
0.26	0	0	14.25
0	19.35	0	8.45
0	0	0	17.75
0	0	0	0.7
0	0	0	1.83
0	0	0	1.56
0	0	0	21.65
9.07	18.6	0	0
12.05	15.7	0	0
0	0	0.28	33.6
11.85	19.75	0	0
0	12.1	0.46	15.3
13.05	16.45	0	0
12.4	16.55	0	0
12.7	18.15	0	0
0	0	0	28.9
0.38	0	0.32	10.8
0	0	0	20.5
0	0	0	29.6
0.33	0	0	8.7
0.27	0	0	16.05
0.26	0	0	12.2
0	0	0	25.1
0.38	0	0.3	18.6
0	0	0	1
0	0	0	32.65
0	0	0	25.95
0	0	0	21.45
0	0	0	1.23
0	20.65	0	14.6
4.08	8.25	0	19.25
0.26	0	0.35	1.41
1	0	0	0
0	0	0	0
1	0	0	0
0	0	0	0
0	0	0	0
1	0	0	1
1	1	1	1
0	0	0	1
3	2	1	2
0	1	0	0
0	0	1	0
0	1	0	0
7	5	3	5



Supplementary Table 5.5. The number of alpha- and gamma-gliadin/secalin amplicons, and epitopes distribution per line in the set of genotypes used in the study. The mean and range are indicated. Mean: average for all the lines classified in each specie category. Range: the minimum and maximum value found for the lines of each specie. Rye YES: bread wheat with rye translocation; Rye NO: bread wheat without rye translocation.

	Bread wheat				Durum wheat		Tritordeum	
	Rye YES		Rye NO		Mean	Range	Mean	Range
	Mean	Range	Mean	Range				
<u>Alpha-gliadin amplicon</u>								
Number of amplicons	31.0	(22-40)	34.2	(23-45)	33.0	(32-34)	31.8	(31-32)
Number putative genes	24.5	(21-30)	26.4	(21-32)	21.0	(20-22)	19.8	(19-20)
Number pseudogenes	6.5	(1-12)	7.8	(1-13)	12.0	(12-12)	12.0	(12-12)
Ratio pseudogenes/p. genes	0.26	(0.05-0.44)	0.28	(0.05-0.48)	0.57	(0.55-0.60)	0.61	(0.60-0.63)
Total number of epitopes	49.9	(35-77)	57.5	(36-80)	46.5	(46-47)	42.6	(42-43)
<u>Gamma-gliadin/secalin amplicon</u>								
Number of amplicons	15.0	(12-20)	17.3	(13-22)	8.0	(8-8)	8.4	(8-10)
Number putative genes	11.5	(8-16)	14.3	(10-18)	5.0	(5-5)	5.0	(5-5)
Number pseudogenes	3.5	(3-4)	3.0	(1-5)	3.0	(3-3)	3.4	(3-5)
Ratio pseudogenes/p. genes	0.32	(0.21-0.50)	0.21	(0.06-0.38)	0.60	(0.60-0.60)	0.68	(0.60-1.00)
Total number of epitopes	73.0	(51-125)	123.0	(90-156)	54.0	(54-54)	58.8	(54-78)



Supplementary Table 5.6. Abundance of CD epitopes in alpha-gliadins' putative gene amplicons.

Genotype	Specie	Rye	p31-43	DQ2.5_glia_α1a	DQ2.5_glia_α1b	DQ2.5_glia_α2	DQ2.5_glia_α3
ACA201	BW	N	28.99	50.08	15.79	42.66	31.38
ACA202	BW	N	28.14	49.66	15.59	42.95	31.26
ACA321	BW	N	16.66	29.55	10.89	26.62	19.4
ACA903B	BW	Y	17.59	28.97	8.3	22.55	20.17
Antequera	BW	N	17.35	31.35	10.79	27.49	19.97
ANZA	BW	N	18.42	30.17	8.78	23.46	20.78
ARAGON03	BW	N	27.68	37.54	5.92	16.05	34.55
ArturNick	BW	N	28.62	36.85	5.52	14.63	35.22
Aucan	HT	N	31.19	22.18	0	0.72	23.27
Baquette9	BW	N	36.45	50.38	8.95	23.09	45.91
BIOINTA1001	BW	Y	16.83	30.49	9.69	26.04	19.14
BIOINTA1004	BW	N	17.48	31.69	9.99	27.17	20.32
BIOINTA1005	BW	N	26.85	36.42	5.76	15.07	34.49
BIOINTA2004	BW	Y	35.42	48.15	7.48	20.12	45.45
BuckBaqueano	BW	Y	26.68	50.59	17.15	46.46	31.34
BuckBrasil	BW	N	26.71	37.1	6.47	16.48	33.83
BuckHuanchen	BW	Y	15.38	28.38	9	24.72	17.25
BuckNorteno	BW	Y	16.99	30.59	9.44	25.73	19.09
BuckPuelche	BW	Y	13.25	32.13	11.59	29.61	20.5
Bulel	HT	N	34.06	33.61	0	0	32.28
CATEDRAL	DW	N	29.05	31.4	0	0.31	30
DMarioAtlax	BW	N	27.51	37	5.38	14.71	34.76
DonPedro	DW	N	33.21	33.04	0	0.29	31.56
Escacena	BW	N	17.77	30.63	8.92	24.26	19.29
HT435	HT	N	34.25	33.72	0	0	32.39
HT444	HT	N	33.08	32.85	0	0	31.4



Genotype	Specie	Rye	p31-43	DQ2.5_glia_α1a	DQ2.5_glia_α1b	DQ2.5_glia_α2	DQ2.5_glia_α3
HT460	HT	N	33.42	23.81	0	0.77	25.08
INIACondor	BW	Y	26.39	35.2	5.17	14	33.45
Jerezano	BW	N	17.64	28.87	8.5	22.43	19.69
KleinDonEnrique	BW	Y	17.31	32.03	10.26	27.97	19.62
KleinGavilan	BW	Y	16.63	30.23	9.45	25.75	18.7
KleinLeon	BW	N	26.4	36.6	6.14	16.21	33.41
KleinRayo	BW	N	26.78	35.08	4.47	12.51	33.51
KleinRendidor	BW	N	27.64	49.49	15.71	43.26	30.72
KleinTigre	BW	Y	25.38	36.12	4.77	13.27	34.46
MarcosJuarez	BW	N	31.53	42.98	7.31	18.84	39.47
OlaetaArtillero	BW	N	26.69	50.48	18.79	48.15	32.21
proINTAGaicho	BW	N	17.4	29.47	9.07	24.23	20.01
proINTAGranar	BW	Y	13.16	31.03	11.1	28.08	19.83
proINTAGuazu	BW	Y	26.34	36.07	5.72	15.25	33.1
proINTAOasis	BW	Y	17.46	30.03	8.8	24.31	19.01
RelmoSiriri	BW	N	27.19	34.64	3.66	11.27	32.98
THA85	BW	N	27.63	31.39	0	3.83	32.48
XEIXA	BW	N	17.51	31.27	10.83	27.32	19.87



Supplementary Table 5.7. Abundance of CD epitopes in gamma-gliadins/secalins' putative gene amplicons.

Genotype	Specie	Rye	DQ2.5_glia_γ1	DQ2.5_glia_γ2a	DQ2.5_glia_γ2b	DQ2.5_glia_γ3	DQ2.5_glia_γ4a	DQ2.5_glia_γ4b	DQ2.5_glia_γ4c	DQ2.5_glia_γ4d	DQ2.5_glia_γ4e	DQ2.5_glia_γ5	DQ2.5_sec1	DQ2.5_sec2
ACA201	BW	N	92.08	27.12	57.52	57.52	15.05	21.3	220.65	11.62	25.48	52.97	4.33	0
ACA202	BW	N	98.27	20.69	53.65	53.65	23.77	27.16	206.21	9.78	28.14	61.49	4.39	0
ACA321	BW	N	90.13	23.84	65.5	65.5	11.5	14.37	206.67	9.33	15.68	37.46	3.25	0
ACA903B	BW	Y	88.81	46.35	32.3	32.3	32.3	22.01	204.53	11	44.15	87.81	0	0.36
Antequera	BW	N	92.05	27.29	51.25	51.25	15.15	18.42	218.3	7.92	20.57	46.65	7.92	0
ANZA	BW	N	93.67	21.5	60	60	18.8	21.79	200.28	6.34	22.48	47.59	2.94	0
ARAGON03	BW	N	89.37	25.59	57.3	57.3	17.59	17.65	215.21	9.42	23.15	56.37	4.39	0
ArturNick	BW	N	89.08	26.19	62.15	62.15	11.65	16.11	209.45	8.01	17.23	43.93	3.1	0
Aucan	HT	N	75.38	10.98	67.3	67.3	0	4.46	169.82	0	0	2.9	0	0
Baquette9	BW	N	95.06	42.23	42.85	42.85	13.93	25.86	246.35	10.7	20.62	49.44	4.49	0
BIOINTA1001	BW	Y	85.47	28.16	36.95	36.95	36.95	28.72	168.65	6.92	45.04	89.24	0	0.27
BIOINTA1004	BW	N	92.56	31.56	56.5	56.5	10.2	16.34	228.05	7.7	15.49	36.7	7.7	0
BIOINTA1005	BW	N	96.79	23.5	59.67	59.67	19.2	21.18	212.84	5.83	24.44	47.17	2.61	0
BIOINTA2004	BW	Y	85.5	21.65	57.55	57.55	10.85	19.78	194.12	7.35	15.35	37.89	7.35	0
BuckBaqueano	BW	Y	87.19	59.15	23.8	23.8	23.8	18.05	237.42	10.55	37.04	80.92	0.57	0.3
BuckBrasil	BW	N	95.6	38.24	47.05	47.05	14.65	22.61	238.26	10.99	21.22	45.54	3.82	0
BuckHuanchen	BW	Y	73.88	44.52	30.75	30.75	30.75	17.65	199.25	12.3	42.79	97.67	0	0
BuckNorteno	BW	Y	77.3	45.88	33.65	33.65	33.65	18.53	207.41	12.8	45.97	104.06	0	0.38
BuckPuelche	BW	Y	79.48	45.76	30	30	30	21.65	205.25	14.9	41.71	98.91	0	0
Bulel	HT	N	72.14	14.25	61.95	61.95	0	2.21	168.72	0	0	4.42	0	0



CATEDRAL	DW	N	65.61	17.27	53.1	53.1	0	2.73	161	0	0	5.46	0	0
DMarioAtlax	BW	N	98.6	10.82	52.75	52.75	41.73	38.56	173.36	3.28	43.27	81.74	1.31	0
DonPedro	DW	N	70.97	15.41	59.7	59.7	0	2.4	168.62	0	0	4.8	0	0
Escacena	BW	N	91.24	31.37	47.3	47.3	17.46	24.47	212.08	8.67	21.7	50.92	3.44	0
HT435	HT	N	67.2	13.31	57.05	57.05	0	1.83	157.86	0	0	3.66	0	0
HT444	HT	N	65.56	13.23	56.25	56.25	0	2.24	153.66	0	0	4.48	0	0
HT460	HT	N	68.54	11.49	60.5	60.5	0	1.89	156.61	0	0	3.78	0	0
INIACondor	BW	Y	85.84	35.36	36.6	36.6	36.6	29.16	195.73	10.38	46.25	103.91	0	0.26
Jerezano	BW	N	87.56	22.08	53.15	53.15	17.37	20.01	207.73	8.55	22.68	56.07	8.55	0
KleinDonEnrique	BW	Y	87.9	52.35	26.35	26.35	26.35	20.8	236.54	7.07	38.44	82.73	7.32	0.3
KleinGavilan	BW	Y	89.47	33.57	33.05	33.05	33.05	29.87	193.12	6.58	40.07	82.94	0	0.27
KleinLeon	BW	N	99.68	33.94	55.6	55.6	11.3	28.99	250.68	9.24	20.36	40.7	9.24	0
KleinRayo	BW	N	96.02	22.84	57.1	57.1	18.3	24.84	210.22	7.1	25.7	51.29	3.33	0
KleinRendidor	BW	N	98.56	24.43	56.6	56.6	13.35	21.8	227.91	10.01	20.28	51.33	10.01	0
KleinTigre	BW	Y	78.83	45.18	28.3	28.3	28.3	25.68	222.6	11.15	41.93	106	11.15	0.58
MarcosJuarez	BW	N	91.13	20.89	50.05	50.05	23.35	25.75	202.18	10.12	27.55	66.12	3.29	0
OlaetaArtillero	BW	N	94.31	26.11	54.15	54.15	16.25	19.19	215.03	11.58	21.76	48.06	5.34	0
proINTAGaicho	BW	N	96.28	6.67	56.55	56.55	44.2	35.33	161.48	3.32	45.95	85.14	1.46	0
proINTAGranar	BW	Y	89.9	34.36	35.7	35.7	35.7	26.21	192.49	7.24	41.81	79.63	0	0.26
proINTAGuazu	BW	Y	86.24	49.29	27.1	27.1	27.1	21.73	228.13	6.8	38.78	83.23	7.11	0.28
proINTAOasis	BW	Y	76.26	52.65	29.3	29.3	29.3	17.35	215.85	14.7	44	104.68	0	0.39
RelmoSiriri	BW	N	94.35	18.15	63.45	63.45	17.9	19.63	198.53	4.73	21.5	45.45	1.96	0
THA85	BW	N	90.61	25.65	52.2	52.2	25.05	19.93	190.68	5.15	30.44	61.31	0	0
XEIXA	BW	N	89.46	28.45	53.66	53.66	16.15	17.09	215.98	10.56	23.04	56.6	3.61	0



Supplementary Table 5.8. Abundance of alpha-gliadins' amplicons 'types' for each line. Each 'type' was established by the number of epitopes matched in the putative gene amplicon peptide sequences.

Genotype	Specie	Rye	alpha_0	Alpha_1	Alpha_2	Alpha_3	Alpha_4	Alpha_5	Alpha_7
ACA201	BW	No	38.3	6.23	16.98	12.55	12.47	0.38	1.47
ACA202	BW	No	38.34	6.16	16.14	13.23	12.24	0.43	1.46
ACA321	BW	No	50.32	3.88	9.94	7.15	6.27	0	2.31
ACA903B	BW	Yes	49.99	4.06	10.66	7.73	4.74	0	1.78
Antequera	BW	No	49.36	3.94	10.71	7.44	7.44	0.29	1.53
ANZA	BW	No	49.46	4.38	11.11	7.65	5.01	0.27	1.75
ARAGON03	BW	No	36.26	1.21	27.41	5.07	4.2	0	0.86
ArturNick	BW	No	36.97	1.23	27.74	4.92	2.86	0	1.33
Aucan	HT	No	48.69	1.09	21.46	0.72	0	0	0
Baquette9	BW	No	28.54	1.36	36.24	6.75	5.83	0	1.56
BIOINTA1001	BW	Yes	50.08	3.84	10.3	7.6	7.51	0.3	0.94
BIOINTA1004	BW	No	52.93	4.32	10.19	8.5	7.05	0.32	1.31
BIOINTA1005	BW	No	39.39	1.01	27.11	4.96	2.94	0	1.41
BIOINTA2004	BW	Yes	28.94	1.28	35.51	6.91	3.98	0	1.75
BuckBaqueano	BW	Yes	37.88	6	15.28	13.86	13.25	0.5	1.7
BuckBrasil	BW	No	39.68	0.98	27.09	4.65	4.25	0	1.11
BuckHuanchen	BW	Yes	52.02	3.75	8.91	7.53	7.38	0	0.81
BuckNorteno	BW	Yes	51.04	4.13	10.17	7.75	7.37	0.27	0.9
BuckPuelche	BW	Yes	51.49	4.04	10.07	8.29	7.59	0.28	1.86
Bulel	HT	No	40.48	1.33	32.28	0	0	0	0
CATEDRAL	DW	No	36.5	1.4	29.69	0.31	0	0	0



DMarioAtlax	BW	No	37.4	1.04	27.67	5	3.28	0	1.05
DonPedro	DW	No	40.4	1.48	31.27	0.29	0	0	0
Escacena	BW	No	48.22	4.19	11.1	7.17	7.15	0.27	0.75
HT435	HT	No	40.53	1.33	32.39	0	0	0	0
HT444	HT	No	41.34	1.45	31.4	0	0	0	0
HT460	HT	No	47.24	1.27	23.04	0.77	0	0	0
INIACondor	BW	Yes	39.49	1.04	26.37	4.85	2.79	0	1.19
Jerezano	BW	No	49.36	4.25	10.69	7.08	4.93	0.27	1.65
KleinDonEnrique	BW	Yes	51.46	4.27	10.05	8.36	8.14	0.3	0.91
KleinGavilan	BW	Yes	50.41	4.08	9.85	7.71	7.45	0.28	0.86
KleinLeon	BW	No	38.79	0.95	26.53	4.93	4.14	0	1
KleinRayo	BW	No	39.36	0.98	27.04	4.53	2.55	0	0.96
KleinRendidor	BW	No	37.79	6.2	15.74	13.2	12.57	0.42	1.36
KleinTigre	BW	Yes	34.38	1.09	27.62	4.74	2.75	0	1.01
MarcosJuarez	BW	No	33.2	1.24	31.45	5.5	4.75	0	1.28
OlaetaArtillero	BW	No	37.32	5.68	15.44	13.47	12.59	0.4	2.9
proINTAGaicho	BW	No	50.41	4.13	10.18	7.82	5.33	0.28	1.73
proINTAGranar	BW	Yes	48.66	3.76	10.29	7.58	7.44	0.26	1.7
proINTAGuazu	BW	Yes	38.38	1.03	26.54	4.67	4	0	0.86
proINTAOasis	BW	Yes	50.57	3.92	10.6	7.56	7.1	0	0.85
RelmoSiriri	BW	No	38.91	1.06	27.03	4.42	2.72	0	0.47
THA85	BW	No	39.85	1.09	27.56	3.83	0	0	0
XEIXA	BW	No	49.18	3.94	10.84	7.21	7.46	0.27	1.55



Supplementary Table 5.9. Abundance of gamma-gliadins' amplicons 'types' for each line. Each 'type' was established by the number of epitopes matched in the putative gene amplicon peptide sequences.

Genotype	Specie	Rye	Gamma_0	Gamma_3	Gamma_4	Gamma_5	Gamma_6	Gamma_7	Gamma_9	Gamma_10	Gamma_11	Gamma_12	Gamma_13	Gamma_14
ACA201	BW	No	0	0.33	0	57.79	4.63	22.79	3.42	0.41	0.46	5.26	0.41	0
ACA202	BW	No	0	0.42	3.54	58.44	2.8	28.13	0	0	0.32	4.47	0.67	0
ACA321	BW	No	0	0	0	68.38	0	13.81	4.9	0	0.52	3.66	0.84	0
ACA903B	BW	Yes	2.48	0.72	0	32.65	11.65	32.3	0	0	1.69	10.16	0	0
Antequera	BW	No	0	0.58	2.69	49.97	9.4	23.07	0	0	0.47	4.95	0.92	0
ANZA	BW	No	0	0	0	61.23	4.8	20.8	2.6	0	0	3.68	0.56	0
ARAGON03	BW	No	0	0.29	2.94	56.48	0.31	21.95	3.64	0	0.31	5.54	1.55	0
ArturNick	BW	No	0	0	0	65.37	0.33	14.07	4.56	0	0.44	5.14	1.09	0
Aucan	HT	No	0	0	0	67.3	3.62	3.01	0	0	0	0	1.45	0
Baquette9	BW	No	0	0.56	3.56	47.39	8.88	26.22	0	0	0.4	6.57	1.65	0
BIOINTA1001	BW	Yes	1.5	0.27	0	35.37	5.06	36.95	0	0	0.43	7.66	0	0
BIOINTA1004	BW	No	0	0.52	0	56.24	9.88	16.77	3.38	0	0.27	5.02	0.48	0
BIOINTA1005	BW	No	0	0	0	57.94	10.42	24.19	0	0	0	3.77	0.47	0
BIOINTA2004	BW	Yes	9	0.27	0	58.66	0	17.64	3.48	0	0.32	4.18	0.95	0
BuckBaqueano	BW	Yes	2.78	0.3	0	28.3	21.85	24.37	0	0	0.92	12.05	0	0.27
BuckBrasil	BW	No	0	0.3	3.48	50.01	8.42	26.24	0	0	0.42	6.15	0.58	0
BuckHuanchen	BW	Yes	2.1	0.46	0	29.95	0.68	30.75	7.46	0	1.77	10.27	0	0
BuckNorteno	BW	Yes	3.33	0.38	0	30.95	0.38	33.65	8.06	0	0.82	11.5	0	0
BuckPuelche	BW	Yes	3.35	0.35	0	36.55	0.87	30	6.57	0	1.01	10.7	0	0
Bulel	HT	No	0	0	0.36	61.95	7.62	0	0	0	0	0	2.21	0



CATEDRAL	DW	No	0	0	0.7	53.1	9.08	0	0	0	0	0	2.73	0
DMarioAtlas	BW	No	0	0.28	1.24	47.83	1.98	45.18	0	0	0.27	1.55	0.27	0
DonPedro	DW	No	0	0	0.66	59.7	8.21	0	0	0	0	0	2.4	0
Escacena	BW	No	0	0.46	0	53.94	8.81	19.97	2.76	0	0.33	4.37	0.87	0
HT435	HT	No	0	0	0.5	57.05	7.82	0	0	0	0	0	1.83	0
HT444	HT	No	0	0	0.56	56.25	6.51	0	0	0	0	0	2.24	0
HT460	HT	No	0	0	0.33	60.5	5.82	0	0	0	0	0	1.89	0
INIACondor	BW	Yes	2.06	0.26	0	39.28	0.31	36.6	5.09	0	0.72	8.65	0	0.28
Jerezano	BW	No	0	0.71	3.58	50.82	0	25.6	2.68	0	0.39	5.24	1.22	0
KleinDonEnrique	BW	Yes	2.38	1.09	0	20.5	21.1	33.67	0	0	0.64	11.45	0	0
KleinGavilan	BW	Yes	2.21	0.54	0	36.18	12.95	33.05	0	0	0.77	6.25	0	0
KleinLeon	BW	No	0	0.26	0	51.68	10.13	31.13	0	0.66	0	4.8	1.02	0
KleinRayo	BW	No	0	0	0	57.78	8.81	24.08	0	0.42	0.32	4.13	0.48	0
KleinRendidor	BW	No	0	0.59	4.32	59.51	2.92	23.63	0	0	0.39	6.54	0.66	0
KleinTigre	BW	Yes	2.45	1.23	0	25.1	0	39.45	6.77	0	1.83	11.8	0	0
MarcosJuarez	BW	No	0	0.6	3.4	55.77	0	26.34	3.5	0	0.41	4.09	0.52	0
OlaetaArtillero	BW	No	0	0	4.95	57.15	3.8	21.97	0	0	0.53	4.98	1.4	0
proINTAGaicho	BW	No	0	0	0.79	47.92	0	45.66	0.83	0	0	1.75	0.16	0
proINTAGranar	BW	Yes	1.84	0.26	0	33.19	14.9	35.7	0	0	0.35	5.76	0	0
proINTAGuazu	BW	Yes	2.7	1.04	0	21.45	18.45	34.21	0.3	0	1.27	10.03	0	0.38
proINTAOasis	BW	Yes	3.77	0.7	0	31.66	0.29	29.3	7.86	0	1.4	12.9	0	0.4
RelmoSiriri	BW	No	0	0.26	0	64.92	4.37	19.81	1.39	0	0.39	3.21	0.27	0
THA85	BW	No	0	0	0.26	51.55	7.68	25.05	0	0	0.52	4.87	0.68	0
XEIXA	BW	No	0	0.35	3.94	56.5	0.49	19.8	2.98	0.13	0.85	6.26	1.14	0



Supplementary Table 5.10. The $\log_2(\text{FC})$ of the abundance for all the CD epitopes and p31-41 peptide variants found in the alpha-gliadins amplicons for each line. The mean of the $\log_2(\text{FC})$ for all the lines was calculated for each variant. The variants with the mean higher than zero are indicated in bold. FC: fold-change.

Original epitope	DQ2.5_α1a	DQ2.5_α1a	DQ2.5_α1a	DQ2.5_α1a
Original seq	PFPPQQLPY	PFPPQQLPY	PFPPQQLPY	PFPPQQLPY
Variant seq	PFPPQQLPY	PFPPQQLPY	PFPPQQLPY	PFPPQQLPY
Position of	5	7	3	8
Mean $\log_2(\text{FC})$	-25.55	-25.77	-16.99	-18.86
ACA201	-28.90	-28.90	-28.90	-28.90
ACA202	-28.89	-28.89	-28.89	-28.89
ACA321	-28.14	-28.14	-28.14	-28.14
ACA903B	-28.11	-28.11	-28.11	-28.11
Antequera	-28.22	-28.22	-28.22	-28.22
ANZA	-28.17	-28.17	-28.17	-28.17
ARAGON03	-28.48	-28.48	-6.55	-4.07
ArturNick	-28.46	-28.46	-6.42	-4.00
Aucan	-27.72	-27.72	-5.41	-2.13
Baquette9	-28.91	-28.91	-6.78	-4.16
BIOINTA1001	-28.18	-28.18	-28.18	-28.18
BIOINTA1004	-28.24	-28.24	-28.24	-28.24
BIOINTA1005	-28.44	-28.44	-6.70	-4.06
BIOINTA2004	-28.84	-28.84	-6.84	-4.08
BuckBaqueano	-28.91	-28.91	-28.91	-28.91
BuckBrasil	-28.47	-28.47	-6.54	-4.16
BuckHuanchen	-28.08	-28.08	-28.08	-28.08
BuckNorteno	-28.19	-28.19	-28.19	-28.19
BuckPuelche	-28.26	-28.26	-28.26	-28.26
Bulel	-3.63	-5.99	-2.77	-28.32
CATEDRAL	-3.39	-5.24	-2.61	-28.23
DMarioAtlas	-28.46	-28.46	-6.61	-3.99
DonPedro	-3.40	-5.31	-2.72	-28.30
Escacena	-28.19	-28.19	-28.19	-28.19
HT435	-3.69	-5.63	-2.89	-28.33
HT444	-3.51	-5.36	-2.81	-28.29
HT460	-27.83	-27.83	-5.54	-2.11
INIACondor	-28.39	-28.39	-6.61	-4.02
Jerezano	-28.10	-28.10	-28.10	-28.10
KleinDonEnrique	-28.25	-28.25	-28.25	-28.25
KleinGavilan	-28.17	-28.17	-28.17	-28.17
KleinLeon	-28.45	-28.45	-6.75	-4.05
KleinRayo	-28.39	-28.39	-6.65	-4.01
KleinRendidor	-28.88	-28.88	-28.88	-28.88
KleinTigre	-28.43	-28.43	-6.46	-4.06
MarcosJuauez	-28.68	-28.68	-6.78	-4.10
OlaetaArtillero	-28.91	-28.91	-28.91	-28.91
proINTAGaicho	-28.13	-28.13	-28.13	-28.13
proINTAGranar	-28.21	-28.21	-28.21	-28.21
proINTAGuazu	-28.43	-28.43	-6.57	-4.10
proINTAOasis	-28.16	-28.16	-28.16	-28.16
RelmoSiriri	-28.37	-28.37	-6.51	-4.02
THA85	-28.23	-28.23	-6.41	-3.85
XEIXA	-28.22	-28.22	-28.22	-28.22



DQ2.5_α3		DQ2.5_α3		DQ2.5_α2		DQ2.5_α2		DQ2.5_α1a		DQ2.5_α1a	
FRPQQPYPQ	FRPQQPYPQ	FRPQQPYPQ	FRPQQPYPQ	PQQLPYPQ	PQQLPYPQ	PQQLPYPQ	PQQLPYPQ	PFQPQLPY	PFQPQLPY	PFQPQLPY	PFQPQLPY
FRPQQLYFQ	FRPQQLYFQ	FRPQQLYFQ	FRPQQLYFQ	PQQLPYLQ	PQQLPYLQ	PQQLPYLQ	PQQLPYLQ	PFQPQLAY	PFQPQLAY	PFQPQLAY	PFQPQLAY
6	6	2	2	8	8	8	8	8	8	8	3
-18.51	-18.51	1.81	1.81	-12.56	-12.56	2.35	2.35	-20.06	-20.06	-19.57	-19.57
-28.23	-28.23	1.91	1.91	-28.67	-28.67	-0.88	-0.88	-28.90	-28.90	-28.90	-28.90
-28.22	-28.22	1.92	1.92	-28.68	-28.68	-0.95	-0.95	-28.89	-28.89	-28.89	-28.89
-27.53	-27.53	2.33	2.33	-27.99	-27.99	-0.95	-0.95	-28.14	-28.14	-28.14	-28.14
-27.59	-27.59	2.28	2.28	-27.75	-27.75	-0.62	-0.62	-28.11	-28.11	-28.11	-28.11
-27.57	-27.57	2.30	2.30	-28.03	-28.03	-0.91	-0.91	-28.22	-28.22	-28.22	-28.22
-27.63	-27.63	2.24	2.24	-27.81	-27.81	-0.60	-0.60	-28.17	-28.17	-28.17	-28.17
-3.95	-3.95	1.34	1.34	-1.37	-1.37	0.69	0.69	-4.25	-4.25	-5.53	-5.53
-3.93	-3.93	1.32	1.32	-1.28	-1.28	0.86	0.86	-4.17	-4.17	-5.53	-5.53
-2.20	-2.20	1.85	1.85	1.66	1.66	5.10	5.10	-27.72	-27.72	-5.28	-5.28
-4.03	-4.03	1.11	1.11	-1.99	-1.99	0.56	0.56	-4.31	-4.31	-5.81	-5.81
-27.51	-27.51	2.36	2.36	-27.96	-27.96	-0.88	-0.88	-28.18	-28.18	-28.18	-28.18
-27.60	-27.60	2.33	2.33	-28.02	-28.02	-0.90	-0.90	-28.24	-28.24	-28.24	-28.24
-3.98	-3.98	1.37	1.37	-1.34	-1.34	0.77	0.77	-4.20	-4.20	-5.79	-5.79
-4.00	-4.00	1.09	1.09	-1.26	-1.26	0.72	0.72	-4.14	-4.14	-5.81	-5.81
-28.22	-28.22	1.90	1.90	-28.79	-28.79	-1.13	-1.13	-28.91	-28.91	-28.91	-28.91
-4.02	-4.02	1.42	1.42	-1.92	-1.92	0.62	0.62	-4.22	-4.22	-6.00	-6.00
-27.36	-27.36	2.50	2.50	-27.88	-27.88	-0.97	-0.97	-28.08	-28.08	-28.08	-28.08
-27.51	-27.51	2.36	2.36	-27.94	-27.94	-0.85	-0.85	-28.19	-28.19	-28.19	-28.19
-27.61	-27.61	2.32	2.32	-28.14	-28.14	-1.07	-1.07	-28.26	-28.26	-28.26	-28.26
-28.27	-28.27	1.36	1.36	24.68	24.68	28.35	28.35	-28.32	-28.32	-28.32	-28.32
-28.16	-28.16	1.44	1.44	2.93	2.93	6.69	6.69	-28.23	-28.23	-28.23	-28.23
-3.90	-3.90	1.34	1.34	-1.26	-1.26	0.83	0.83	-4.11	-4.11	-5.81	-5.81
-28.23	-28.23	1.38	1.38	3.14	3.14	6.85	6.85	-28.30	-28.30	-28.30	-28.30
-27.52	-27.52	2.33	2.33	-27.85	-27.85	-0.67	-0.67	-28.19	-28.19	-28.19	-28.19
-28.27	-28.27	1.37	1.37	24.76	24.76	28.35	28.35	-28.33	-28.33	-28.33	-28.33
-28.23	-28.23	1.40	1.40	24.66	24.66	28.33	28.33	-28.29	-28.29	-28.29	-28.29
-2.18	-2.18	1.73	1.73	1.63	1.63	5.11	5.11	-27.83	-27.83	-4.97	-4.97
-3.95	-3.95	1.40	1.40	-1.19	-1.19	0.82	0.82	-4.11	-4.11	-5.69	-5.69
-27.55	-27.55	2.30	2.30	-27.74	-27.74	-0.59	-0.59	-28.10	-28.10	-28.10	-28.10
-27.55	-27.55	2.35	2.35	-28.06	-28.06	-0.97	-0.97	-28.25	-28.25	-28.25	-28.25
-27.48	-27.48	2.39	2.39	-27.94	-27.94	-0.89	-0.89	-28.17	-28.17	-28.17	-28.17
-3.92	-3.92	1.42	1.42	-1.40	-1.40	0.62	0.62	-4.13	-4.13	-5.91	-5.91
-3.95	-3.95	1.40	1.40	-1.01	-1.01	1.03	1.03	-4.07	-4.07	-5.80	-5.80
-28.19	-28.19	1.94	1.94	-28.69	-28.69	-0.98	-0.98	-28.88	-28.88	-28.88	-28.88
-4.00	-4.00	1.32	1.32	-0.71	-0.71	0.98	0.98	-4.04	-4.04	-5.73	-5.73
-3.97	-3.97	1.23	1.23	-1.37	-1.37	0.64	0.64	-4.27	-4.27	-5.66	-5.66
-28.26	-28.26	1.86	1.86	-28.84	-28.84	-1.19	-1.19	-28.91	-28.91	-28.91	-28.91
-27.58	-27.58	2.30	2.30	-27.85	-27.85	-0.76	-0.76	-28.13	-28.13	-28.13	-28.13
-27.56	-27.56	2.30	2.30	-28.06	-28.06	-1.00	-1.00	-28.21	-28.21	-28.21	-28.21
-3.98	-3.98	1.42	1.42	-1.29	-1.29	0.71	0.71	-4.12	-4.12	-5.79	-5.79
-27.50	-27.50	2.37	2.37	-27.86	-27.86	-0.74	-0.74	-28.16	-28.16	-28.16	-28.16
-3.95	-3.95	1.43	1.43	-0.85	-0.85	1.19	1.19	-4.09	-4.09	-5.67	-5.67
-3.90	-3.90	1.42	1.42	0.83	0.83	2.76	2.76	-3.85	-3.85	-5.45	-5.45
-27.57	-27.57	2.30	2.30	-28.03	-28.03	-0.89	-0.89	-28.22	-28.22	-28.22	-28.22



DQ2.5_α3 FRPQQPYPQ FQPQQPYPQ	DQ2.5_α3 FRPQQPYPQ F*TPQQPYPQ	DQ2.5_α3 FRPQQPYPQ FRPQQPYPR	DQ2.5_α3 FRPQQPYPQ FSFQQPYPQ	DQ2.5_α3 FRPQQPYPQ FRPQQPYPQ	DQ2.5_α3 FRPQQPYPQ FRPQQPYPQ
2	2	9	2	3	5
-15.31	-19.70	-18.49	-16.18	-25.11	-25.31
-2.33	-28.23	-28.23	-28.23	-28.23	-28.23
-2.34	-28.22	-28.22	-28.22	-28.22	-28.22
-2.32	-27.53	-27.53	-27.53	-27.53	-27.53
-2.31	-27.59	-27.59	-27.59	-27.59	-27.59
-2.34	-27.57	-27.57	-27.57	-27.57	-27.57
-2.25	-27.63	-27.63	-27.63	-27.63	-27.63
-28.36	-4.13	-3.85	-4.03	-28.36	-28.36
-28.39	-4.10	-3.75	-3.91	-28.39	-28.39
-27.79	-27.79	-2.96	-3.21	-27.79	-27.79
-28.77	-4.18	-3.75	-3.64	-28.77	-28.77
-2.32	-27.51	-27.51	-27.51	-27.51	-27.51
-2.23	-27.60	-27.60	-27.60	-27.60	-27.60
-28.36	-4.12	-3.79	-3.61	-28.36	-28.36
-28.76	-4.05	-3.69	-3.63	-28.76	-28.76
-2.38	-28.22	-28.22	-28.22	-28.22	-28.22
-28.33	-4.09	-3.84	-3.56	-28.33	-28.33
-2.20	-27.36	-27.36	-27.36	-27.36	-27.36
-2.21	-27.51	-27.51	-27.51	-27.51	-27.51
-2.34	-27.61	-27.61	-27.61	-27.61	-27.61
-28.27	-28.27	-28.27	-3.34	-2.71	-4.60
-28.16	-28.16	-28.16	-3.32	-2.55	-4.42
-28.37	-4.02	-3.91	-3.63	-28.37	-28.37
-28.23	-28.23	-28.23	-3.35	-2.66	-4.41
-2.20	-27.52	-27.52	-27.52	-27.52	-27.52
-28.27	-28.27	-28.27	-3.26	-2.83	-4.61
-28.23	-28.23	-28.23	-3.25	-2.74	-4.44
-27.90	-27.90	-3.08	-3.45	-27.90	-27.90
-28.32	-4.04	-3.79	-3.75	-28.32	-28.32
-2.21	-27.55	-27.55	-27.55	-27.55	-27.55
-2.20	-27.55	-27.55	-27.55	-27.55	-27.55
-2.20	-27.48	-27.48	-27.48	-27.48	-27.48
-28.32	-4.00	-3.93	-3.57	-28.32	-28.32
-28.32	-4.00	-3.83	-3.73	-28.32	-28.32
-2.31	-28.19	-28.19	-28.19	-28.19	-28.19
-28.36	-3.97	-3.77	-28.36	-28.36	-28.36
-28.56	-4.15	-3.75	-3.55	-28.56	-28.56
-2.50	-28.26	-28.26	-28.26	-28.26	-28.26
-2.28	-27.58	-27.58	-27.58	-27.58	-27.58
-2.40	-27.56	-27.56	-27.56	-27.56	-27.56
-28.30	-4.00	-3.87	-3.58	-28.30	-28.30
-2.28	-27.50	-27.50	-27.50	-27.50	-27.50
-28.30	-4.01	-3.76	-3.70	-28.30	-28.30
-28.27	-3.90	-3.52	-3.30	-28.27	-28.27
-2.33	-27.57	-27.57	-27.57	-27.57	-27.57



p31-41		p31-41		p31-41		p31-41		DQ2.5_α3	
LGQQQFFPQQPY	LGQQQFFPQQPY	LGQQQFFPQQPY	IGQQQFFPQQPY	IGQQQFFPQQPY	IGQQQFFPQQPY	IGQQQFFPQQPY	FRPQQPYPQ	FRPQQPYPQ	FRPQQPYPQ
LGQQQFFPQQPY	IGQQQFFPQQPY	LGQQQFFPQQPY	IGQQQFFPQQPY	IGQQQFFPQQPY	IGQQQFFPQQPY	PGQQQFFPQQPY	FRPQQSYPO	FRPQQSYPO	FLPQQPYPQ
8	6	6	1	6	1	2	6	6	2
-16.62	-25.02	-25.02	1.03	-25.02	1.03	1.03	-6.93	-6.93	-5.29
-28.11	-28.11	-28.11	0.86	-28.11	0.86	0.86	-1.33	-1.33	-2.21
-28.07	-28.07	-28.07	0.93	-28.07	0.93	0.93	-1.35	-1.35	-2.21
-27.31	-27.31	-27.31	1.92	-27.31	1.92	1.92	-1.63	-1.63	-2.42
-27.39	-27.39	-27.39	1.80	-27.39	1.80	1.80	-2.09	-2.09	-2.35
-27.37	-27.37	-27.37	1.84	-27.37	1.84	1.84	-1.42	-1.42	-2.50
-27.46	-27.46	-27.46	1.78	-27.46	1.78	1.78	-2.05	-2.05	-2.40
-3.82	-28.04	-28.04	0.48	-28.04	0.48	0.48	-3.04	-3.04	-3.91
-3.88	-28.09	-28.09	0.41	-28.09	0.41	0.41	-3.62	-3.62	-3.90
-3.77	-28.22	-28.22	0.29	-28.22	0.29	0.29	-27.79	-27.79	-3.03
-28.44	-28.44	-28.44	-0.34	-28.44	-0.34	-0.34	-2.98	-2.98	-3.77
-27.33	-27.33	-27.33	1.88	-27.33	1.88	1.88	-1.35	-1.35	-2.40
-27.38	-27.38	-27.38	1.93	-27.38	1.93	1.93	-1.53	-1.53	-2.43
-4.02	-28.00	-28.00	0.59	-28.00	0.59	0.59	-3.55	-3.55	-3.92
-3.81	-28.40	-28.40	-0.40	-28.40	-0.40	-0.40	-3.51	-3.51	-3.74
-27.99	-27.99	-27.99	1.06	-27.99	1.06	1.06	-1.24	-1.24	-2.26
-27.99	-27.99	-27.99	0.66	-27.99	0.66	0.66	-2.99	-2.99	-3.96
-27.20	-27.20	-27.20	2.10	-27.20	2.10	2.10	-1.22	-1.22	-2.27
-27.34	-27.34	-27.34	1.93	-27.34	1.93	1.93	-1.37	-1.37	-2.43
-26.98	-26.98	-26.98	2.41	-26.98	2.41	2.41	-1.43	-1.43	-27.61
-2.79	-3.76	-3.76	0.14	-3.76	0.14	0.14	-28.27	-28.27	-3.45
-2.50	-3.63	-3.63	0.52	-3.63	0.52	0.52	-28.16	-28.16	-28.16
-3.89	-28.04	-28.04	0.46	-28.04	0.46	0.46	-3.41	-3.41	-3.90
-2.73	-3.69	-3.69	0.16	-3.69	0.16	0.16	-28.23	-28.23	-3.52
-27.40	-27.40	-27.40	1.82	-27.40	1.82	1.82	-1.43	-1.43	-2.43
-2.92	-3.82	-3.82	0.15	-3.82	0.15	0.15	-28.27	-28.27	-3.45
-2.82	-3.70	-3.70	0.22	-3.70	0.22	0.22	-28.23	-28.23	-3.50
-3.81	-28.32	-28.32	0.09	-28.32	0.09	0.09	-27.90	-27.90	-3.16
-3.83	-27.98	-27.98	0.60	-27.98	0.60	0.60	-3.58	-3.58	-3.92
-27.39	-27.39	-27.39	1.84	-27.39	1.84	1.84	-2.00	-2.00	-2.37
-27.37	-27.37	-27.37	1.95	-27.37	1.95	1.95	-1.27	-1.27	-2.34
-27.31	-27.31	-27.31	1.98	-27.31	1.98	1.98	-1.33	-1.33	-2.38
-3.89	-27.98	-27.98	0.62	-27.98	0.62	0.62	-3.01	-3.01	-3.91
-3.86	-28.00	-28.00	0.59	-28.00	0.59	0.59	-3.72	-3.72	-3.95
-28.04	-28.04	-28.04	1.01	-28.04	1.01	1.01	-1.29	-1.29	-2.21
-2.73	-27.92	-27.92	0.68	-27.92	0.68	0.68	-3.65	-3.65	-28.36
-3.86	-28.23	-28.23	0.07	-28.23	0.07	0.07	-3.05	-3.05	-3.82
-27.99	-27.99	-27.99	1.06	-27.99	1.06	1.06	-1.36	-1.36	-2.33
-27.37	-27.37	-27.37	1.84	-27.37	1.84	1.84	-1.91	-1.91	-2.34
-26.97	-26.97	-26.97	2.33	-26.97	2.33	2.33	-1.41	-1.41	-27.56
-3.81	-27.97	-27.97	0.62	-27.97	0.62	0.62	-3.05	-3.05	-3.86
-27.38	-27.38	-27.38	1.84	-27.38	1.84	1.84	-1.42	-1.42	-2.38
-3.74	-28.02	-28.02	0.53	-28.02	0.53	0.53	-3.60	-3.60	-3.77
-3.71	-28.04	-28.04	0.37	-28.04	0.37	0.37	-28.27	-28.27	-3.73
-27.38	-27.38	-27.38	1.83	-27.38	1.83	1.83	-1.41	-1.41	-2.38



p31-41 LGQQQFFPPQQPY LGQQQFFPPQQPY	p31-41 LGQQQFFPPQQPY LGKQQFFPPQQPY	p31-41 LGQQQFFPPQQPY LGQ*QFFPPQQPY	p31-41 LGQQQFFPPQQPY LGQQQFFRFPQQPY
6	6	3	4
-19.14	-24.99	-24.11	-25.23
-28.11	-28.11	-28.11	-28.11
-28.07	-28.07	-28.07	-28.07
-27.31	-27.31	-27.31	-27.31
-27.39	-27.39	-27.39	-27.39
-27.37	-27.37	-27.37	-27.37
-27.46	-27.46	-27.46	-27.46
-2.71	-28.04	-28.04	-28.04
-2.80	-28.09	-28.09	-28.09
-28.22	-28.22	-6.36	-28.22
-2.65	-28.44	-28.44	-28.44
-27.33	-27.33	-27.33	-27.33
-27.38	-27.38	-27.38	-27.38
-2.64	-28.00	-28.00	-28.00
-2.58	-28.40	-28.40	-28.40
-27.99	-27.99	-27.99	-27.99
-2.61	-27.99	-27.99	-27.99
-27.20	-27.20	-27.20	-27.20
-27.34	-27.34	-27.34	-27.34
-26.98	-26.98	-26.98	-26.98
-28.34	-3.65	-5.05	-5.71
-28.11	-3.28	-4.33	-5.44
-2.67	-28.04	-28.04	-28.04
-28.31	-3.40	-4.56	-5.51
-27.40	-27.40	-27.40	-27.40
-28.35	-3.71	-4.88	-5.72
-28.30	-3.52	-4.54	-5.67
-28.32	-28.32	-5.64	-28.32
-2.62	-27.98	-27.98	-27.98
-27.39	-27.39	-27.39	-27.39
-27.37	-27.37	-27.37	-27.37
-27.31	-27.31	-27.31	-27.31
-2.59	-27.98	-27.98	-27.98
-2.61	-28.00	-28.00	-28.00
-28.04	-28.04	-28.04	-28.04
-2.56	-27.92	-27.92	-27.92
-2.62	-28.23	-28.23	-28.23
-27.99	-27.99	-27.99	-27.99
-27.37	-27.37	-27.37	-27.37
-26.97	-26.97	-26.97	-26.97
-2.60	-27.97	-27.97	-27.97
-27.38	-27.38	-27.38	-27.38
-2.69	-28.02	-28.02	-28.02
-2.56	-28.04	-28.04	-28.04
-27.38	-27.38	-27.38	-27.38



p31-41
LGQQQPFPPQQPY
LGQQQTFPPQQPY

1

-16.18
-4.45
-4.45
-3.99
-4.01
-4.13
-4.17
-28.04
-28.09
-28.22
-28.44
-4.18
-4.43
-28.00
-28.40
-4.52
-27.99
-4.16
-4.30
-3.95
-28.34
-28.11
-28.04
-28.31
-4.01
-28.35
-28.30
-28.32
-27.98
-4.00
-4.30
-4.22
-27.98
-28.00
-4.50
-27.92
-28.23
-4.55
-4.36
-3.89
-27.97
-4.33
-28.02
-28.04
-3.99





CHAPTER VI

Conclusions



CONCLUSIONS

- I) Gliadin silencing by RNAi showed a strong efficiency despite the high temperature and nitrogen availability environmental conditions, providing these lines could be further tested in field trials.
- II) During grain filling, the strong silencing of gliadins by RNAi leads to stress-like transcriptomic changes.
- III) The silencing of wheat gliadins provides evidence for the coregulation of prolamins, other proteins (α -amylase/trypsin inhibitors, avenin-like proteins) and starch in the RNAi lines, and highlights the possible implication of new transcription factors in the regulation of the synthesis of these components in the grain.
- IV) Globulins, which are mostly participating in grain protein compensation of the RNAi lines, are not subject to the network of transcription factors that in cascade regulate the synthesis of prolamins.
- V) There is an active sink-source communication as a consequence of gliadins silencing to supply the metabolic components to readjust the protein composition by increasing the non-gluten proteins, mainly globulins.
- VI) The developed amplicon-based pipeline is an efficient throughput tool for the screening of celiac disease immunogenic epitopes in cereals, allowing the identification of low-immunogenic varieties as potential candidates for precision breeding programs.



