



POLITÉCNICA



E.T.S. DE INGENIERÍA AGRÓNOMICA,
ALIMENTARIA Y DE BIOSISTEMAS

UNIVERSIDAD POLITÉCNICA DE MADRID
ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA AGRONÓMICA,
ALIMENTARIA Y DE BIOSISTEMAS
GRADO EN BIOTECNOLOGÍA

DEPARTAMENTO DE BIOTECNOLOGÍA-BIOLOGÍA VEGETAL

***Desarrollo de marcadores moleculares tipo KASP
("Kompetitive Allele Specific PCR") útiles para la
selección en programas de mejora***

TRABAJO FIN DE GRADO

Autor: Sandra Hernández Caballero

Tutor: Patricia Giraldo Carbajo

Junio de 2021



POLITÉCNICA



E.T.S. DE INGENIERÍA AGRONÓMICA,
ALIMENTARIA Y DE BIOSISTEMAS

**DESARROLLO DE MARCADORES MOLECULARES TIPO
KASP ("KOMPETITIVE ALLELE SPECIFIC PCR") ÚTILES
PARA LA SELECCIÓN EN PROGRAMAS DE MEJORA**

**Memoria presentada por Sandra Hernández Caballero para la obtención del
título de Graduado en Biotecnología por la Universidad Politécnica de Madrid**

Fdo: Sandra Hernández Caballero

VºBº Tutor


Dra Patricia Giraldo Carbajo
Prof Titular de Universidad
Dpto de Biotecnología-Biología Vegetal
ETSIAAB-Universidad Politécnica de Madrid

Madrid, 24, junio, 2021

AGRADECIMIENTOS

Quiero dar las gracias en primer lugar a los compañeros de Genética, me he sentido como en casa. En especial a mi tutora Patri, por haberme acogido. Por su gran paciencia, con los problemas y los imprevistos que surgían. Por haberme resuelto las dudas cuantas veces fueran necesarias. Por tu cariño y sobre todo tu optimismo en las situaciones difíciles. También, agradecer a Mati su continua sonrisa, su felicidad y su bondad, transmites mucho cariño y me alegra haberme cruzado contigo y haber compartido momentos. Gracias a Albita, nunca imaginé que iba a ser una persona tan importante en esta última etapa del curso, gracias por ser un gran apoyo, por comprenderme y por darme tan buenos consejos. A Paqui, principalmente por su paciencia en el laboratorio, por enseñarme nuevas canciones, por siempre saber que decir en los momentos en los que necesitaba un empujón, y por tus mil y un historias que ojalá poder escuchar más. A Mario por hacerme compañía y por las charlas en el laboratorio, y por estar dispuesto a ayudarme. A Paco por las horas de charlas en tu despacho, por los consejos y por tener todos los días una sonrisa que transmite calma y felicidad. Gracias a Laura y Elena por tener tiempo para charlar conmigo, preocuparse y darme buenos consejos. Y gracias a las tres mellizas, Ale, Alex e Iván por estar conmigo día tras día ayudándome en todo lo necesario. Considero que todas y cada una de las personas que conforman este laboratorio son personas como digo yo que llenan el corazón, llenan el alma, por transmitir cariño, tratar con amor y es un placer haber compartido estos meses de muchas emociones con ustedes, considero que valen mucho y ha sido un placer tener la oportunidad de conocerles.

Gracias a mis niñas de Tenerife, por tener una amistad que aguanta distancia y tiempo, gracias a Pou, María y Nicole por estar desde y sé que será para siempre. Por los viajes, las fiestas, las sonrisas y los lloros, gracias por todos los momentos compartidos y los que compartiremos, por tener ilusión de verme.

Gracias a mis “nuevos” amigos que he encontrado en Madrid en estos años de carrera, a Paqx por nuestros largos caminos y charlas que me ayudaban a desconectar y relajarme. A Rodri y Sara por siempre tener tiempo para mí, para escucharme, y para sacarme una sonrisa. Gracias a Lucía (Megape) por enseñarme que hay tiempo para estudiar y beber una cerveza. Gracias a mis compañeras de piso de Cea Bermúdez, en especial a Carla por tener un corazón enorme, ser comprensiva y nunca rendirte cuando peor lo estaba pasando, por ayudarme a salir y poder encontrarme. No me imagino Madrid sin ustedes, en esta etapa siempre estarán sus nombres como un gran apoyo moral para mí.

Por último, agradecer a mi familia, por hacerme sentir que todo va a ir bien. Mención especial a papá y a mamá, por darme la libertad que necesitaba, por ver en mi esa fuerza que yo no he visto, por haberme levantado siempre que me caía. Gracias a ustedes puedo terminar una etapa que en un primer momento no imaginé.

Muchas gracias a todas aquellas personas que me han acompañado y de las que guardo un pequeño recuerdo en mi corazón.

ÍNDICE DE CONTENIDO

ÍNDICE DE FIGURAS.....	vi
ÍNDICE DE TABLAS.....	viii
LISTA DE ABREVIATURAS.....	ix
SUMMARY.....	xi
CAPÍTULO 1. INTRODUCCIÓN.....	1
1.1. El trigo.....	1
1.2. Las proteínas del endospermo del trigo.....	2
1.3. Base genética y estructura genética de las HMW-GS.....	3
1.4. Análisis de la variabilidad alélica de HMW-GS.....	4
1.5. Genotipado con KASP.....	6
1.6. Objetivos.....	7
CAPÍTULO 2. MATERIALES Y MÉTODOS.....	8
2.1. Material vegetal.....	8
2.2. Electroforesis SDS-PAGE.....	8
2.3. Extracción de ADN.....	8
2.4. Análisis de secuencias de HMW-GS.....	9
2.4.1. Análisis del locus <i>Glu-D1</i>	9
2.5. Análisis de marcadores en termociclador Roche LC96.....	11
2.5.1. KASP.....	11
2.5.2. Curva de disociación.....	12
CAPÍTULO 3. RESULTADOS Y DISCUSIÓN.....	13
3.1. Material vegetal.....	13
3.2. Extracción de proteínas.....	14
3.3. Extracción de ADN.....	15
3.4. Diseño KASP.....	15
3.5. Genotipado con marcadores moleculares KASP.....	21
3.5.1. KASP diseñados para identificar SNPs.....	21
3.5.1.1. Marcadores para el gen <i>Glu-A1-1</i>	21
3.5.1.2. Marcadores para el gen <i>Glu-B1-1</i>	22
3.5.1.3. Marcadores para el gen <i>Glu-D1</i>	25
3.5.2. KASP diseñados para identificar InDels.	26
3.5.2.1. Marcadores para e gen <i>Glu-A1-1</i>	26
3.5.2.2. Marcadores para el gen <i>Glu-B1-1</i>	27
3.6. Curva de disociación.....	28
CAPÍTULO 4: CONCLUSIONES.....	30
CAPÍTULO 5: BIBLIOGRAFÍA.	31
ANEXO	34

ÍNDICE DE TABLAS

Tabla 1. Variabilidad alélica de las gluteninas de alto peso molecular (HMW-GS).....	5
Tabla 2. Secuencia y Tm de los cebadores ppDx y qqDx.....	10
Tabla 3. Condiciones de la PCR de gradiente para determinar la T ^a óptima de los cebadores ppDx y qqDx.....	10
Tabla 4. Condiciones para la amplificación del gen <i>Glu-D1-1</i> con los cebadores ppDx y qqDx.....	10
Tabla 5. Componentes de la reacción de KASP.....	11
Tabla 6. Condiciones para reacción KASP en LC96.....	12
Tabla 7. Componentes de la reacción para la curva de disociación en LC96.....	12
Tabla 8. Condiciones para reacción de curva de disociación en termociclador LC96.....	12
Tabla 9. Variedades seleccionadas de trigo blando según su composición en HMW-GS.....	13
Tabla 10. Variedades seleccionadas de trigo duro según su composición en HMW-GS.....	13
Tabla 11. Números de accesoión de la base de datos del NCBI (National Center for Biotechnology Information) de las secuencias de los alelos utilizados para diseñar los marcadores moleculares tipo KASP. n.d.: se refiere a una entrada directa, no existe artículo de referencia.....	16
Tabla 12. Marcadores KASP: tipo, cromosoma (Crom), secuencias diana para el diseño de los cebadores específicos y descripción del polimorfismo que detecta alelo HEX y el alelo FAM	20

INDICE DE FIGURAS

Figura 1. La evolución y las relaciones del genoma entre trigo duro y trigo blando y en relación con los cultivares silvestres.....	1
Figura 2. Esquema de los cromosomas de distintos grupos de homeología mostrando los principales loci que codifican gluteninas en trigo blando.....	4
Figura 3. Dibujo esquemático de funcionamiento de KASP.....	6
Figura 4. Gel de electroforesis de la extracción de proteínas de trigo blando.....	14
Figura 5. Gel de electroforesis de la extracción de proteínas de trigo duro y trigo blando.....	14
Figura 6. Muestras de DNA de trigo blando y duro, analizadas en un gel de agarosa al 1%....	15
Figura 7. Identificación de polimorfismos para diferentes subunidades (Ax-2*, Ax-2·, Ax-null, Ax-1) del gen <i>Glu-A1-1</i> alineados con el software Geneious v8.....	17
Figura 8. Amplificación de variedades de trigo con cebadores ppDx y qqDx.....	18
Figura 9. Análisis BLAST de la secuencia de la variedad 'Marius' (subunidad Dx-4) con la base de datos del NCBI.....	19
Figura 10. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM.	21
Figura 11. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM.....	22
Figura 12. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM del marcador GluB1x_7oe del gen <i>Glu-B1-1</i>	23
Figura 13. Distribución de las variedades de trigo según fluorescencia de los fluoróforos HEX o FAM	23
Figura 14. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM del marcador GluB1x14_20 del gen <i>Glu-B1-1</i>	24
Figura 15. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM	25
Figura 16. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM del marcador GluD1y_12 del gen <i>Glu-D1-2</i>	26

Figura 17. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM	27
Figura 18. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM del marcador GluB1x_7 del gen <i>Glu-B1-1</i>	27
Figura 19. Cinéticas de amplificación y disociación de marcadores KASP.....	29
Figura suplementaria 1. Alineamiento de las secuencias de las subunidades Bx-13, Bx-7 y Bx-7*	34
Figura suplementaria 2. Alineamiento entre las secuencias de las subunidades Dx-2, Dx-3 y Dx-5 del locus <i>Glu-D1</i> para el diseño de primer mediante el software Primer3 para la secuenciación de la subunidad Dx-4.....	35

LISTA DE ABREVIATURAS

ADN: Ácido Desoxirribonucleico

dNTPs: desoxinucleótidos trifosfato

CDS ("Coding Sequence"): región codificante

g: gramos

HMW-GS ("High Molecular Weight Glutenin Subunit"): Subunidad de glutenina de alto peso molecular

InDels: Inserción o Delección

KASP ("Kompetitive Allele Specific PCR"): PCR específica de alelo competitivo

LMW-GS ("Low Molecular Weight Glutenin Subunit"): Subunidad de glutenina de bajo peso molecular

mA: mili Amperio

min: minuto

ml: mililitro

mm: milímetro

mQ: mili Q

N: nitrógeno

NCBI ("National Center for Biotechnology Information"): Centro Nacional de Información Biotecnológica

ng: nanogramo

PCR ("Polymerase Chain Reaction"): Reacción en Cadena de la Polimerasa

pb: pares de bases

rpm: revoluciones por minuto

s: segundos

SDS-PAGE ("Sodium Dodecyl Sulfate Polyacrylamide gel electrophoresis"): electroforesis en gel de poliacrilamida con dodecilsulfato sódico

SNP ("Single Nucleotide Polymorphism"): polimorfismo de nucleótido único

UTR ("Untranslated Region"): región no traducida

T^a: temperatura

μL: microlitros

°C: grados centígrados

SUMMARY

Wheat is one of the three most important crops for human feed. The content and composition of endosperm seed storage proteins are the main responsible for wheat quality. These proteins are mainly glutenins and gliadins.

In wheat (especially in bread wheat, *Triticum aestivum*), glutenins are the most important determinants of end-use quality. Up to now, sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) of grain proteins is the most widely method used for the identification of glutenins, which can be classified as high molecular weight glutenins (HMW-GS) and low molecular weight glutenins (LMW-GS). HMW-GS are highly diverse and are encoded by *Glu 1-1* and *Glu-1-2* genes, located at *Glu-A1*, *Glu-B1* and *Glu-D1 loci*, for which several alleles that influence the quality of dough have been identified and characterized.

The interpretation of HWM glutenin subunits by the SDS-PAGE method is not always accurate due to the small differences between the electrophoretic mobility of protein bands. Moreover, this method is time-consuming and not suitable for high-throughput analysing. The solution to this problem would be to classify the HMW-GS allelic variants by using molecular DNA markers.

The objective of this work was to develop and validate a set of molecular markers for the main HMW-GS that could assist in the characterization of this key determinants of wheat quality.

A collection of 30 selected varieties, 22 of bread wheat and 8 of durum wheat, representing a great allelic diversity was chosen, including at least two varieties carrying the same allele. The genotype of the varieties was verified by SDS-PAGE from flour extracted from the seeds prior to the DNA extraction. The sequences of the different alleles of the *Glu-A1*, *Glu-B1* and *Glu-D1* loci were obtained from the databases and bibliography and were further analysed with the Geneious software for the identification of polymorphic regions suitable for the design of KASP molecular markers. In the present work, thirteen KASP molecular markers were designed, ten markers designed from SNPs polymorphisms (Single Nucleotide Polymorphism) and three from InDels polymorphisms (Insertion/Deletion).

The KASP markers designed from SNPs worked as expected. The markers for the *Glu-A1* and *Glu-D1* loci discriminated among the different subunits properly, while for the *Glu-B1* gene

there were some misclassifications. KASP markers designed from Indels did not work as expected, suggesting that this type of polymorphisms is not the best option for KASP design, at least in these complex loci.

In this work, some KASP markers for HMW-GS related to good and bad end-use quality have been developed in wheat, and they are ready to be tested in a large set of germplasm in order to evaluate their potential use in high-throughput screening in breeding programs.

CAPÍTULO 1. INTRODUCCIÓN

1.1. El trigo

El trigo es un cereal fundamental para la alimentación humana, siendo uno de los tres más importantes en cuanto a producción media después del arroz y del maíz. La producción mundial de trigo alcanzó 765,7 millones de toneladas en 2019 (FAO). El trigo presenta una amplia diversidad genética que le aporta la capacidad de adaptarse a distintos climas. Esta diversidad se puede apreciar en el desarrollo de, aproximadamente, 2.500 variedades distintas (Shewry et al., 2003)

Existen diferentes tipos de trigo cultivado, los más importantes son el trigo duro o semolero (*Triticum turgidum* ssp. *durum* L), que se usa para la producción de pasta, cuscús y algún tipo de pan; y el trigo blando o panadero (*Triticum aestivum* ssp. *vulgare* L) mucho más cultivado y, empleado para la fabricación de pan, galletas, bizcochos.

El trigo blando es un hexaploide ($2n=42$) que presenta tres genomios A, B y D que son homeólogos entre ellos y están formados por 7 pares de cromosomas homólogos. Mientras que el trigo duro se diferencia por ser un tetraploide ($2n=28$) con sólo genomios A y B.

El origen de la poliploidización tuvo lugar hace miles de años. El trigo no deriva de un solo progenitor. El genoma de *T.turgidum* es el producto de la hibridación de *T.urartu* donador del genomio A y una segunda especie con el genomio B. El donador del genomio B es desconocido, pero se piensa que está estrechamente relacionado con *Aegilops speltoides*. Más tarde se cruzaría *T.turgidum* con un donante del genomio D, *T.tauschii* dando lugar a los distintos trigos hexaploides existentes hoy en día (Figura 1) (Shewry, 2009).

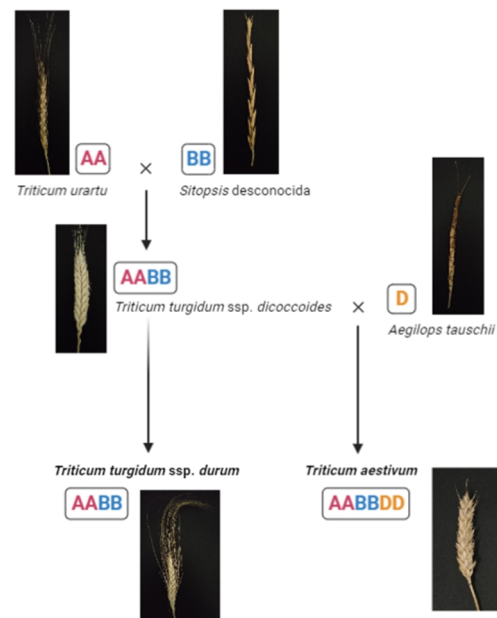


Figura 1. La evolución y las relaciones del genoma entre trigo duro y trigo blando y en relación con los cultivares silvestres.

1.2. Las proteínas del endospermo del trigo

La adaptabilidad y el alto rendimiento del trigo no son suficientes para explicar su importancia a nivel mundial. En comparación con otros cultivos, su principal ventaja son las características únicas de la masa elaborada con harina de trigo, obtenida a partir de la molienda del grano, con la que se puede elaborar panes, galletas, pastas y otros productos. Estas características de la masa dependen de la estructura y de la interacción de las proteínas de reserva del grano, que forman una red tridimensional denominada gluten (Shewry, 2009).

La calidad harino-panadera es un concepto amplio, debido a que hay una gran variedad de productos que se obtienen a partir de la harina de trigo en la industria panadera y, a que en su obtención y evaluación no sólo está implicada esta, sino también el consumidor cuyos criterios de calidad pueden ser diferentes. Uno de los parámetros que determinan la calidad son las interacciones de las proteínas de reserva del endospermo para formar la red de gluten que va a proporcionar a la masa unas características viscoelásticas (fuerza, elasticidad y extensibilidad) únicas (Guo et al., 2018).

El gluten, formado por prolaminas puede fácilmente representar el 80% de la proteína total de la harina de trigo. Estas proteínas se denominan prolaminas por su alto contenido en los aminoácidos de prolina y glutamina y dentro de ellas se diferencian las gluteninas y las gliadinas (Espinosa et al., 2015).

Las gliadinas son proteínas monoméricas y según su movilidad electroforética en geles de poliacrilamida se pueden clasificar en cuatro grupos, α -, β -, γ - y ω -gliadinas. Las α -, β -, y γ -gliadinas son ricas en residuos de cisteína y metionina, se pueden estabilizar por puentes disulfuro y puentes de hidrógeno; mientras que las ω -gliadinas, son más importantes por su implicación en la celiaquía, presentan ausencia de cisteína en su estructura, pero están compuestas por secuencias repetitivas con elevadas proporciones de glutamina, prolina y alguna fenilalanina, estabilizándose por puentes de hidrógeno ya que no pueden formar puentes disulfuro (Espí, 2013).

Las gluteninas son proteínas poliméricas que están implicadas en las propiedades viscoelásticas de la masa. Están unidas por puentes disulfuro inter- e intra-catenarios. Los polímeros de gluteninas son una mezcla heterogénea de polímeros con pesos moleculares variables. Las gluteninas están entre las macromoléculas más grandes encontradas en la naturaleza (Wrigley y Morris, 1996). En presencia de un agente reductor, se pueden clasificar según su movilidad electroforética en geles de electroforesis SDS-PAGE en gluteninas de alto

peso molecular (HMW-GS) y gluteninas de bajo peso molecular (LMW-GS). Las HMW-GS son las principales responsables de la calidad de la harina, explicando hasta un 60% de la variabilidad de este carácter.

La secuencia codificante de las gluteninas está compuesta por un dominio central repetitivo y dos secuencias únicas terminales. El tamaño del dominio central repetitivo es altamente variable lo que explica por qué las proteínas se pueden distinguir fácilmente por diferencias en movilidad electroforética en geles de electroforesis SDS-PAGE.

1.3. Base genética y estructura genética de las gluteninas

La naturaleza poliploide del trigo blando tiene importantes consecuencias para la composición y análisis de las proteínas del gluten. Los tres genomios contienen genes que codifican para proteínas equivalentes y que están localizados en posiciones similares en los cromosomas homeólogos.

Las HMW-GS son las más estudiadas a nivel genético, ya que son las prolaminas menos complejas y se han realizado numerosos estudios que relacionan la variedad alélica para los genes de HMW-GS con la calidad panadera. Estas proteínas están codificadas por genes situados en los loci complejos denominados *Glu-1*, que se encuentran en los brazos largos de los cromosomas homeólogos del grupo 1, calificándolos como *Glu-A1*, *Glu-B1*, y *Glu-D1* según el genomio en el que se encuentran (Figura 2) (Payne et al., 1981). Cada locus complejo está constituido por dos genes estrechamente ligados, el *Glu-1-1* y el *Glu-1-2*, que codifican para dos proteínas diferentes, tipo-x y tipo-y, respectivamente, que se denominan subunidades de HMW.

En el trigo no siempre se expresan todos los genes, de tal forma que el locus *Glu-A1* puede codificar para una o para ninguna proteína y los loci *Glu-B1* y *Glu-D1* codifican para 1 o 2 subunidades. De esta forma una variedad de trigo blando puede contener hasta 5 subunidades HMW-GS distintas y una de trigo duro hasta 3 subunidades distintas (Waines y Payne, 1987; Li et al., 2021).

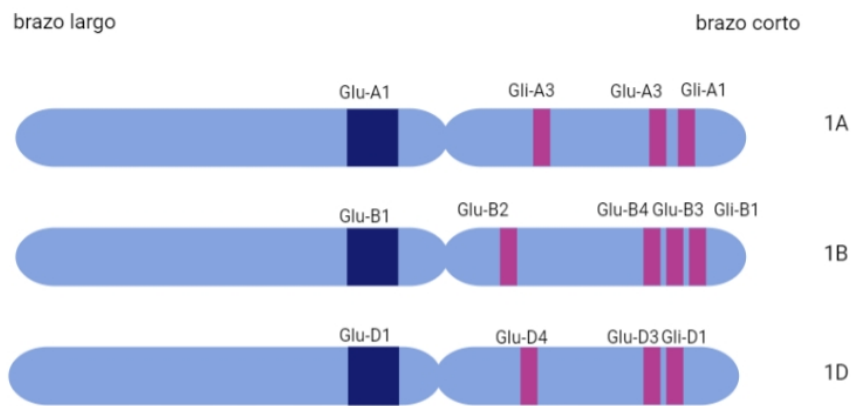


Figura 2. Esquema de los cromosomas de distintos grupos de homeología mostrando los principales loci que codifican gluteninas en trigo blando

Las LMW-GS incluyen muchos componentes y no se conoce muy bien su relación con la calidad en trigo blando, aunque son fundamentales para la calidad semolera en trigo duro. Estas proteínas están codificadas por los loci *Glu-A3*, *Glu-B3* y *Glu-D3* localizados en los brazos cortos del grupo 1 de homeología de cromosomas (Figura 2). Estos loci presentan gran variación alélica y la organización genética de los genes LMW-GS no se conoce bien, aunque se estima que existen más de 40 genes implicados. Estas proteínas son más difíciles de identificar debido a su complejidad, heterogeneidad y similitud entre ellas (Dreisigacker et al., 2020). Además de estos genes, los loci de prolaminas se caracterizan por tener una elevada presencia de pseudogenes, lo que dificulta el análisis genético.

1.4. Análisis de la variabilidad alélica de HMW-GS

La variabilidad de las proteínas de reserva del endospermo del trigo, y en concreto de las HMW-GS, es muy importante a la hora de determinar la calidad de la harina, La variación alélica de las HMW-GS se describe en base a las combinaciones de subunidades de tipo x e y que se expresan y su análisis mediante electroforesis SDS-PAGE. En la Tabla 1, se muestra la principal variabilidad descrita para estos loci, en la que se aprecia que el locus *Glu-B1* es el más polimórfico. El catálogo completo de variantes de HMW-GS se encuentra recopilado en el "Catalogue of Gene Symbols for Wheat" (<https://wheat.pw.usda.gov/GG3/wgc>).

La combinación de los diferentes alelos de HMW-GS determina la calidad. Una buena calidad panadera se atribuye a subunidades como Ax-1 y Ax-2 del *Glu-A1*, las combinaciones Bx-7+By-8, Bx-7+ By-9, Bx-13 + By-16 y la Bx-17+ By-18 del *Glu-B1* y la combinación de subunidades Dx-5+ Dy-10 del *Glu-D1*. Por el contrario, una peor calidad está relacionada con

la subunidad Ax-null del *Glu-A1*, las combinaciones de subunidades Bx-6 + By-8, Bx-20 + By-20 y Bx-22 del *Glu-B1* y la combinación de subunidades Dx-2+ Dy-12 del *Glu-D1*. (Gálová et al., 2002; Filip, 2018; Sharma et al, 2020).

Tabla 1. Variabilidad alélica de las gluteninas de alto peso molecular (HMW-GS).

Locus	Subunidades		Alelos
	x	y	
<i>Glu-A1</i>	1	-	<i>a</i>
	2*	-	<i>b</i>
	null	-	<i>c</i>
	2..	-	<i>y</i>
<i>Glu-B1</i>	7	-	<i>a</i>
	7	8	<i>b</i>
	6	8	<i>d</i>
	20	20	<i>e</i>
	13	16	<i>f</i>
	14	15	<i>h</i>
	7oe	8	<i>al</i>
	-	18	<i>am</i>
	17	18	<i>i</i>
	7	8*	-
	7*	9	-
	7*	8*	-
<i>Glu-D1</i>	2	12	<i>a</i>
	4	12	<i>c</i>
	5	10	<i>d</i>

Para los programas de mejora, los geles de electroforesis en SDS-PAGE se siguen utilizando para identificar tanto las HMW-GS como las LMW-GS y seleccionar aquellas variedades que portan subunidades relacionadas con una buena calidad. El método SDS-PAGE requiere muy poca harina para poder evaluar la proteína del grano, pudiéndose partir tan sólo de medio grano para la muestra. La interpretación de las HWM-GS por el método SDS-PAGE no siempre es acertado debido a que la diferencia entre bandas (diferente movilidad electroforética) a veces no es visible y algunas de estas proteínas pueden confundirse, por ejemplo, las subunidades Ax-2* vs Ax-2.., Bx-7 vs Bx-7* o By-8 vs By-8* (Giraldo et al 2010; Espí et al, 2012; Ravel et al., 2020). Esto es importante ya que se han clasificado como iguales para en su composición para HMW-GS variedades que realmente portan subunidades diferentes que incluso pueden tener una influencia distinta en la calidad (Espí et al, 2012).

La solución a este inconveniente sería poder clasificar las HMW-GS mediante marcadores moleculares (Liu et al., 2008). Identificar las gluteninas con marcadores moleculares

requeriría poca cantidad de ADN que puede ser fácilmente extraído de harina o de las hojas de las plántulas.

Desde la década de 1990, se han realizado muchos intentos de desarrollar marcadores moleculares a partir de secuencias de gluteninas, generalmente basado en la presencia / ausencia o tamaño de polimorfismos, según lo revisado por Gale (2005). Jin et al., (2011), Liu et al., (2010) y más recientemente Ibba et al., (2018) vieron que la identificación de subunidades de HMW-GS o LMW-GS mediante PCR concordaba con los resultados de SDS-PAGE.

El paso previo en el desarrollo de marcadores es la identificación de polimorfismos. Se han descrito polimorfismos de tipo SNPs y de tipo inserciones/delecciones (InDels) tanto en HWM-GS como en LMW-GS, que pueden permitir la identificación de los diferentes alelos. Sin embargo, los pocos SNPs disponibles para distinguir entre diferentes HMW-GS y la existencia de dominios repetitivos limita la elección del sitio para el desarrollo de marcadores.

1.5. Genotipado con KASP

Uno de los métodos más utilizados hoy en día para la identificación de polimorfismos de tipo SNP en mejora vegetal es el ensayo KASP ("Kompetitive Allele Specific PCR"), este método es utilizado tanto para la identificación de SNPs como de InDels. Recientemente se han publicado los primeros estudios en los que aplican esta tecnología para el genotipado de variedades de trigo para distintos caracteres (Rasheed et al, 2016), incluyendo genes relacionados con calidad (Dreisigacker et al 2020; Ravel et al 2020).

En este método se utilizan tres cebadores no marcados y que son específicos del ensayo: dos cebadores específicos de alelo y un cebador común. Cada uno de los cebadores específicos de alelo alberga una secuencia de cola única que se corresponde con un "casette" FRET (transferencia de energía resonante de fluorescencia) universal; uno etiquetado con el fluoróforo FAM y el otro con HEX. Se parte de un ADN de doble cadena, en un primer paso se produce la desnaturalización y el anillamiento de los cebadores (1^{er} ciclo de PCR). El cebador específico se une a la secuencia que presenta el polimorfismo a detectar y se produce la elongación, se une así la secuencia de la cola única a la hebra recién sintetizada. En el segundo ciclo de PCR se genera la secuencia complementaria a la cola específica del alelo. Y en el tercer ciclo de PCR se une el "casette" que contiene el fluoróforo al ADN y emite

fluorescencia. El ensayo KASP logra una discriminación mediante la unión de los cebadores específicos a los diferentes alelos.

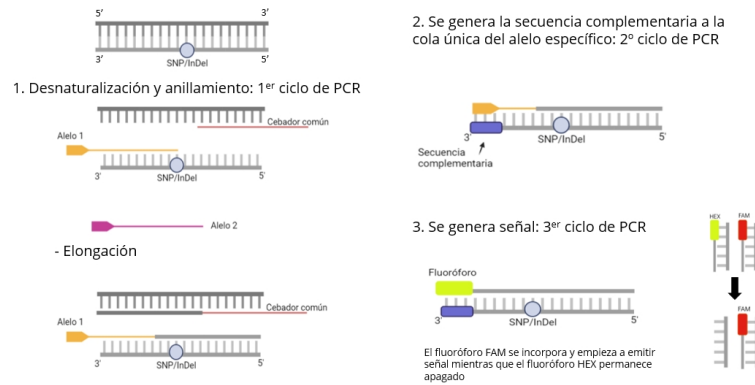


Figura 3. Dibujo esquemático de funcionamiento de KASP

1.6. Objetivos

El **objetivo principal** de este trabajo es desarrollar marcadores moleculares basados en ensayos KASP para la identificación de diferentes HMW-GS que puedan ser utilizados para la selección de variedades en función de su calidad.

Para ello, se abordarán los siguientes objetivos parciales:

1. Diseño y comprobación de un set de variedades que reflejen la variabilidad más importante para HMW-GS.
2. Análisis de secuencias de HMW-GS, obtenidas a partir de la bibliografía y bases de datos, y búsqueda de polimorfismos.
3. Diseño de secuencias óptimas para análisis por marcadores moleculares
4. Puesta a punto de reacciones de genotipado KASP

CAPÍTULO 2: MATERIALES Y MÉTODO

2. 1. Material vegetal

En este trabajo se han utilizado 30 variedades, donde se incluyen 22 variedades de trigo blando y 8 de trigo duro, intentando escoger las variedades necesarias para representar la mayor diversidad alélica, e incluyendo al menos dos variedades con la misma subunidad. Fueron seleccionadas a través de la información disponible en el "Catalogue for gene symbol of wheat" (<https://shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp>)

Se pusieron a germinar las semillas, en total 10-15 semillas por placa. Cinco de esas semillas fueron cortadas a la mitad, el embrión se puso a germinar con el resto de las semillas y la otra mitad fue utilizada para hacer harina para poder realizar una electroforesis SDS- PAGE y comprobar el genotipo de las variedades.

Una vez germinadas (después de una semana), se recogieron entre 0,5-1 g de hoja verde, se guardaron en tubos Falcon de 50 ml y se congelaron a -80°C para preservar el material vegetal hasta la futura extracción de DNA.

2. 2. Electroforesis SDS-PAGE

Para la extracción de HMW-GS se realizó un protocolo estándar de extracción de proteínas del endospermo (Singh et al, 1991). Las muestras se analizaron mediante electroforesis SDS-PAGE (Payne et al. 1980) utilizando geles cortos, con separadores de 1 mm con un gel separador al 12% y un gel concentrador al 3% Los geles se corrieron durante 3 horas a 50 mA. Posteriormente, para la tinción se utilizó una solución colorante de azul de Coomassie al 6%,

2.3. Extracción de ADN

Se realizó una maxi extracción de ADN según el siguiente protocolo:

- Recoger 0,5-1 g de hoja joven y machacar en mortero con N₂ líquido
- Añadir 7,5 ml de CTAB Buffer previamente calentado a 65°C
- Incubar tubos a 65°C durante 30 min
- Centrifugar 5 min a 4000 rpm. Recoger sobrenadante.
- Añadir 7,5 ml de Cloroformo-Isoamilico (24:1)
- Centrifugar 20 min a 4000 rpm. Recoger sobrenadante.
- Añadir 7,5 ml de Cloroformo-Isoamilico (24:1)

- Centrifugar 20 min a 4000 rpm. Recoger sobrenadante
- Añadir 1 volumen de Isopropanol (-20°C) Mezclar y dejar 30 min en reposo a Tª ambiente
- Centrifugar 10 min a 4000 rpm. Eliminar sobrenadante
- Añadir 5 ml de Washing Buffer, invertir tubos hasta que se desprenda el pellet. Dejar en nevera mínimo 2 horas
- Centrifugar 5 min a 4000 rpm. Eliminar sobrenadante. Dejar secar pellet.
- Añadir 1 ml de TE
- Tratar muestras con 1 µl de RNAsa a 50 µg/ml de concentración final 1 hora 37°C

Para comprobar la calidad de la extracción se visualizó el ADN en geles de agarosa al 1% y se cuantificó mediante nanodrop. Se prepararon diluciones de trabajo a 50-100ng/µl.

2.4. Análisis de secuencias de HMW-GS

Las secuencias de los diferentes alelos de los loci *Glu-A1*, *Glu-B1* y *Glu-D1* se obtuvieron a partir de la bibliografía (Giraldo et al., 2010; Espí et al., 2012; Ravel et al., 2016; Rasheed et al., 2020). Los códigos de acceso se introdujeron en la plataforma NCBI ("National Center for Biotechnology Information", <https://www.ncbi.nlm.nih.gov/>) y se descargaron las secuencias en formato FASTA. Además, se hizo una búsqueda en la base de datos para completar con mayor número de secuencias cada locus.

Se seleccionaron las secuencias que tuvieran región codificante completa (CDS), regiones 5' no codificantes (5' UTR), región promotora y 3'UTR. En esos casos en los que varias accesiones correspondían con el mismo alelo se realizó un alineamiento utilizando el software Geneious v8 y se escogió la que más completa estuviera o la que se considerara más fiable tras alinear las secuencias del mismo tipo.

El análisis con el software Geneious v8 permitió identificar regiones polimórficas que permitiesen el diseño de marcadores de tipo KASP. Se acotaron entre 50-100 pb alrededor del polimorfismo y se envió esa información a LCG Genomics (<https://www.lcggroup.com/>) para el diseño de los cebadores.

2.4.1. Análisis del locus *Glu-D1*

Se buscaron y alinearon las secuencias disponibles para el gen *Glu-D1-1*, correspondientes a las subunidades Dx-2, Dx-3 y Dx-5, y se diseñaron cebadores con el software Primer3 en las regiones comunes, con el objetivo de amplificar ese gen en variedades portadoras de

esos alelos y poder identificar polimorfismos. Se obtuvieron dos parejas de cebadores (ppDx y qqDx) (Tabla 2) que se optimizaron realizando una PCR de gradiente de temperatura.

Tabla 2. Secuencia y Tm de los cebadores ppDx y qqDx

PRIMER 1		Tm
qqDxFw	GGAAGAAACACTTGGAGATCATAGA	60.02
qqDxRv	GAATTGATCTCTAGTTTGTGGTGCT	60.08
PRIMER 2		
ppDxFw	AATATATAAAACATTCCGGCAACAA	59.90
ppDxRv	ATGATAAGATTGTGAAGTTGGCTA	60.26

La PCR de gradiente de temperatura se realizó en un BioRad thermal cycler, utilizando una temperatura de anillamiento que variaba entre 50-60°C (Tabla 3):

Tabla 3. Condiciones de la PCR de gradiente para determinar la Tª óptima de los cebadores ppDx y qqDx

Paso	Descripción	Tª	Tiempo	Numero de ciclos
1	Activación	95°C	4 min	1 ciclo
2	Desnaturalización	95°C	1 min	35 ciclos
	Anillamiento	50°-60°C	1min	
	Elongación	72°C	1 min	
3	Elongación	72°C	10 min	1 ciclo
4	-	12°C	∞	

Una vez que se determinó que la temperatura óptima eran 58°C, los cebadores se utilizaron para amplificar el DNA de las variedades en las siguientes condiciones (Tabla 4):

Tabla 4. Condiciones para la amplificación del gen *Glu-D1-1* con los cebadores ppDx y qqDx

Paso	Descripción	Tª	Tiempo	Numero de ciclos
1	Activación	95°C	4 min	1 ciclo
2	Desnaturalización	95°C	1 min	35 ciclos
	Anillamiento	58°C	1 min	
	Elongación	72°C	1 min	
3	Elongación	72°C	10 min	1 ciclo

El resultado de la amplificación se visualizó en geles de agarosa al 1% y se purificó con columnas de Sephadex G-50 Superfine para enviar a secuenciar siguiendo el siguiente protocolo:

- Se llena la placa de carga con el Sephadex. Una vez llenos los pocillos, se encaja la placa de filtro en la placa de carga y se invierte. Se golpea suavemente para asegurar que todo el Sephadex cae en el interior de los pocillos de la placa de filtro.
- La placa de carga se retira y se añade 325 µl de agua mQ en cada pocillo.

- Se deja polimerizar el Sephadex 1 hora a Tª ambiente.
- Pasada la hora, se encaja la placa ELISA debajo de la placa de filtro y se fija con cinta para que no se mueva.
- Se centrifuga 2 min a 2000 rpm para secar las columnas de Sephadex de la placa de carga y se desecha el agua que cae en la placa de ELISA.
- Se toma la placa de PCR donde se vayan a recoger las muestras purificadas y se encaja debajo de la placa de carga.
- Se cargan las muestras de PCR a purificar (23-25 µl) intentando que caiga en el centro de la columna de Sephadex, y se añaden otros 25 µl de agua de PCR.
- Se centrifuga 2 min a 2000 rpm, el producto de PCR pasa a través de la columna y cae en el pocillo de la placa de PCR ya purificado.
- La placa con el producto de PCR purificado se envía a secuenciar a Macrogen Inc (<https://dna.macrogen-europe.com/eng/>).

Una vez recibidas las secuencias se analizaron con el software Geneious.

2.5. Análisis de marcadores en termociclador Roche LC96

2.5.1. KASP

Los marcadores KASP recibidos se analizaron en un termociclador Roche LC96. Para poder realizar este paso es necesario el ADN, el *assay mix* (contiene los dos cebadores específicos y el cebador común) y el master mix (contiene los fluoróforos, dNTPs y la Taq polimerasa). El volumen final de la reacción es de 10 µl (Tabla 5) que se disponen en placas ópticas de 96 pocillos (LightCycler 480 Multiwell Plate 96, white).

Tabla 5. Componentes de la reacción de KASP

	Placa de 96 pocillos (µl) Volumen necesario (µl) / muestra
DNA	2
KASP Master mix	5
KASP Assay mix	0.14
Agua	3
Volumen final	10

Las reacciones de amplificación se llevaron a cabo utilizando el programa recomendado por LGC genomics aumentando en 6 ciclos el paso 3 (de 26 a 32) para aumentar el rendimiento, y añadiendo 20 ciclos al final para medir la fluorescencia a tiempo final (Tabla 6).

Tabla 6. Condiciones para reacción KASP en LC96

Paso	Descripción	Tª	Tiempo	Numero de ciclos
1	Activación	94°C	15 min	1 ciclo
2	Desnaturalización	94°C	20 s	10 ciclos
	Annealing/elongación	61-55°C	60 s (-1 °C / ciclo)	
3	Desnaturalización	94°C	20 s	32 ciclos
	Annealing/elongación	55°C	60 s	
4	Elongación (medir fluorescencia)	39°C	10 s	20 ciclos
		37°C	30 s	

2.5.2. Curvas de disociación

Para realizar las pruebas de las de curvas de disociación en el

Tabla 7. Componentes de la reacción para la curva de disociación en LC96

termociclador Roche LC96, se utilizó el Kasp *assay mix* (contiene cebadores específicos y el cebador común) en combinación con el *Fast Start mix* (contiene SYBR Green, Taq polimerasa y dNTPs) en un volumen final de reacción de 20 µl

	Placa de 96 pocillos (µl) Volumen necesario para cada pocillo
DNA	2
Mix Fast Start	10
KASP assay	0,28
Agua	7,72
Volumen final	20

(Tabla 7), Se utilizaron placas ópticas de 96 pocillos (LightCycler 480 Multiwell Plate 96, white). El programa que se utilizó tenía las siguientes condiciones (Tabla 8):

Tabla 8. Condiciones para reacción de curva de disociación en termociclador LC96

Paso	Descripción	Tª	Tiempo	Numero de ciclos
1	Pre-incubación	95°C	600 s	1 ciclo
2	Amplificación	95°C	10 s	45 ciclos
		55°C	10s	
		72°C	10 s	
3	Melting	95°C	10s	1 ciclo
		65°C	65s	
		97°C	1s	

CAPÍTULO 3. RESULTADOS Y DISCUSIÓN

3.1. Material vegetal

Las variedades escogidas según la diversidad alélica y los datos bibliográficos utilizados, se muestran en las siguientes tablas:

Tabla 9. Variedades seleccionadas de trigo blando según su composición en HMW-GS

	Glu-A1	Glu-B1	Glu-D1
NEEPAWA	2*	7*+9	5+10
GLENLEA	2*	7OE+8	5+10
HOPE	1	6+8	5+10
CAPPELLE DESPREZ	null	7	2+12
BERDUN	1	17+18	2+12
GABO	2*	17+18	2+12
OPATA	2*	13+16	2+12
HALBERD	1	20x+20y	5+10
GORDEJUELA	1	18	2+12
SAPPO	2*	14+15	2+12
CSO	null	7+8	2+12
RIBEIRO	2..	6+8	2+12
CHAMPLEIN	null	7+8	4+12
GAZUL	null/2..	7+8	5+10
CRACLIN	null	7+8	2+12
NORSTAR	1	7*+8*	5+10
MARIUS	null	7*+9	4+12
AGL-385	2..	13+16	2+12
Shan 229	1	20x+20y	5+10
FIEL	null	7+8*	5+10
ROYO	2..	18	2+12
AGL-367	2*	7	2+12

Tabla 10. Variedades seleccionadas de trigo duro según su composición en HMW-GS

	Glu-A1	Glu-B1
LANGDON	null	6+8
ALAGA	1	13+16
SENATORE CAPPELLI	null/1	20
BIDI 17	null	20
AMILCAR	null	7+8
ANTALIS	2*	17+18
DAURUR	1	6+8
SVEVO	null	7+8

Se intentó cubrir la variabilidad alélica más común, incluyendo alelos relacionados con buena y con mala calidad. Aunque el objetivo principal de este trabajo era desarrollar marcadores para trigo blando, se incluyeron algunas variedades de trigo duro para comprobar la posibilidad de aplicarlos también en esta especie. La nomenclatura de subunidades de HMW-GS es la misma en ambas especies, y se ha realizado en base a la movilidad electroforética en geles SDS-PAGE, pero no se sabe a ciencia cierta si a nivel

molecular son idénticas. El uso de marcadores moleculares puede ayudar a aclarar este punto.

3.2. Extracción de proteínas

Para comprobar que las variedades seleccionadas de trigo presentaban realmente los alelos que se indican en las Tablas 9 y 10, se realizaron extracciones de HMW- GS y se examinaron con geles SDS-PAGE. En las Figuras 4 y 5 se muestra el resultado de la electroforesis. Las subunidades se identificaron según su movilidad electroforética, confirmándose la preasignación de alelos y pudiéndose corregir indeterminaciones en aquellas variedades que presentaban en el catálogo la posibilidad de tener distintas subunidades en el mismo locus, como el caso de 'Gazul' (Ax-null o Ax-2·) o 'Senatore cappelli' (Ax-null o Ax-1).

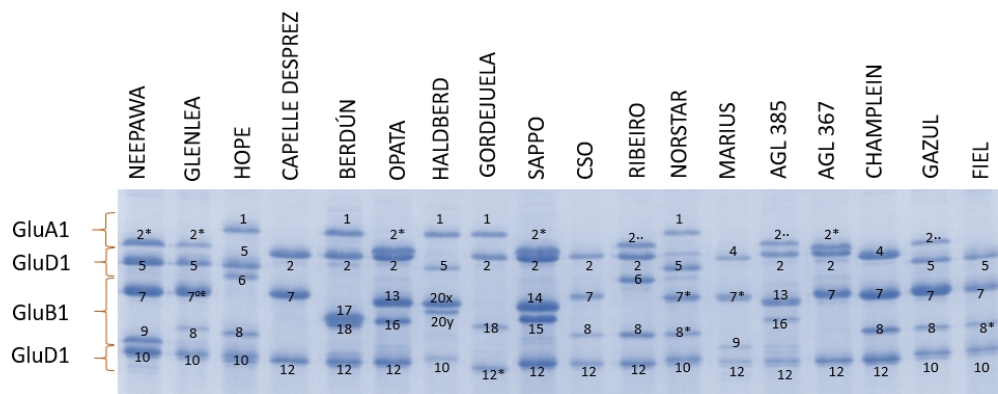


Figura 4. Gel de electroforesis de la extracción de proteínas de trigo blando

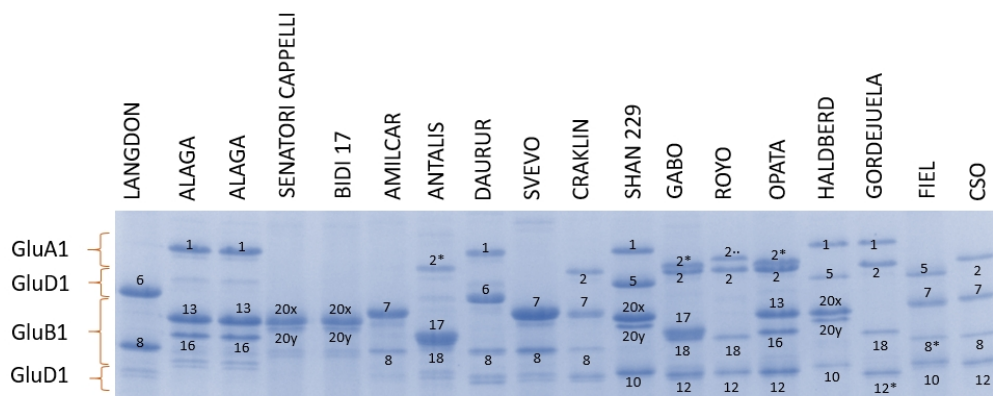


Figura 5. Gel de electroforesis de la extracción de proteínas de trigo duro y trigo blando

Se puede observar en los geles de electroforesis que los alelos correspondían con los que se indican en las Tablas 1 y 2. Por otro lado, se pudo confirmar que la variedad 'Gazul'

presentaba la subunidad Ax-2 en el locus *Glu-A1* y que en 'Senatore capelli' no aparecía ninguna banda en el locus *Glu-A1*, por lo que le correspondía la subunidad Ax-null.

La presencia de distintos genotipos en la misma variedad no es infrecuente ya que pueden existir distintos biotipos bajo el mismo nombre (Metakovsky et al., 2019), por este motivo es imprescindible realizar siempre una comprobación del material, además de para evitar posibles errores debido a un mal manejo del mismo.

3.3. Extracción de ADN

Posteriormente, se realizó una maxi extracción de ADN. La calidad se comprobó mediante electroforesis en geles de agarosa (Figura 6) y las muestras se cuantificaron con nanodrop.

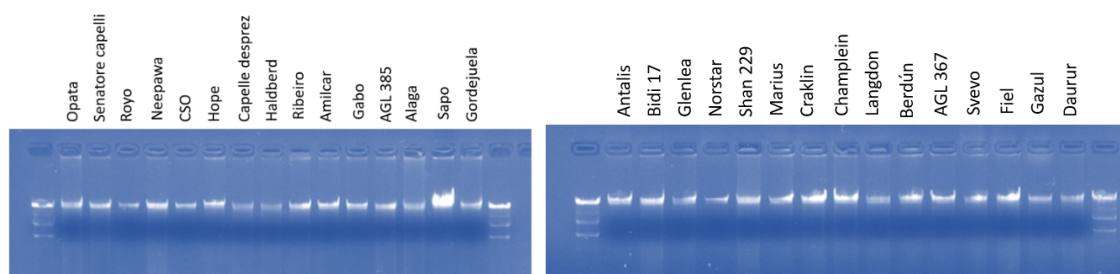


Figura 6. Muestras de ADN de trigo blando y duro, analizadas en un gel de agarosa al 1%

3.4. Diseño KASP

A partir de la bibliografía (Giraldo et al., 2010; Espí et al., 2012; Ravel et al., 2020; Rasheed et al., 2016) se seleccionaron como referencia para cada alelo y para el diseño de los marcadores tipo KASP las secuencias que se muestran en la Tabla 11.

La secuencia parcial del gen *Glu-B1-1* correspondiente a la subunidad Bx-7* y no disponible en la base de datos se obtuvo de Espí et al (2012).

Los genes que codifican para HMW-GS, presentan una región codificante de alrededor de 2400 pb y no presentan intrones. Hay que destacar que el número de secuencias de HMW-GS en la base de datos del NCBI no es muy elevado, muchas son secuencias parciales y además, la mayoría son de mala calidad y/o contienen errores, bien en la propia secuencia bien en la asignación. No es infrecuente que secuencias correspondientes a un genotipo se asignen a otro debido a la alta homología que presentan estos genes.

Como se puede ver en la Tabla 11, para algunas subunidades no se pudieron localizar secuencias codificantes completas y para otras se disponían de una gran cantidad de secuencia flanqueante.

Tabla 11. Números de accesoión de la base de datos del NCBI (National Center for Biotechnology Information) de las secuencias de los alelos utilizados para identificar polimorfismos. n.d.: se refiere a una entrada directa, no existe artículo de referencia

	Gen	Número de accesoión	pb	Referencia
Glu-A1-x-null	GluA1-1	AF145590.1	2537	De Bustos et al., 2000
Glu-A1-x-2*	GluA1-1	M22208.2	6837	Anderson, OD y Greene, FC., 1989)
Glu-A1-x-2..	GluA1-1	DQ533690.1	4648	Gobaa et al., 2007
Glu-A1-x-1	GluA1-1	X61009.1	2885	Halford et al., 1992
Glu-D1-x-2	GluD1-1	DQ208971.1	1252	n.d.
Glu-D1-x-5	GluD1-1	X12928.5	8463	Anderson et al., 1989
Glu-D1-y-10	GluD1-2	DQ537337.1	152010	Anderson, OD, 2009
Glu-D1-y-12	GluD1-2	BK006459.1	1977	Gu et al., 2006
Glu-B1-x-6	GluB1-1	KM116481.1	747	Ravel et al., 2014
Glu-B1-x-7	GluB1-1	DQ537336.1	206063	Gu et al., 2006
Glu-B1-x-13	GluB1-1	KX454510.1	2391	n.d.
Glu-B1-x-14	GluB1-1	AY367771.1	4021	n.d.
Glu-B1-x-17	GluB1-1	AB263219.1	2244	n.d.

Para el desarrollo de marcadores moleculares es necesario identificar aquellos polimorfismos (SNPs/InDels) que permitan diferenciar unas subunidades de otras. Se realizaron los alineamientos de las secuencias correspondientes a cada gen y se buscaron las regiones más adecuadas para el diseño de los marcadores.

En la Figura 7 se puede observar uno de los alineamientos que se realizaron con las secuencias del gen *Glu-A1-1*, correspondientes a las subunidades Ax-1, Ax-null, Ax-2* y Ax-2.. Se puede apreciar que las secuencias tienen un alto grado de homología pero que pueden identificarse algunos polimorfismos, son esos SNPs e InDels los que se utilizarán para la discriminación e identificación de los diferentes alelos que presentan las distintas variedades. Las regiones que están rodeadas son las que se han utilizados para el diseño de los cebadores específicos (marcados con HEX o FAM).

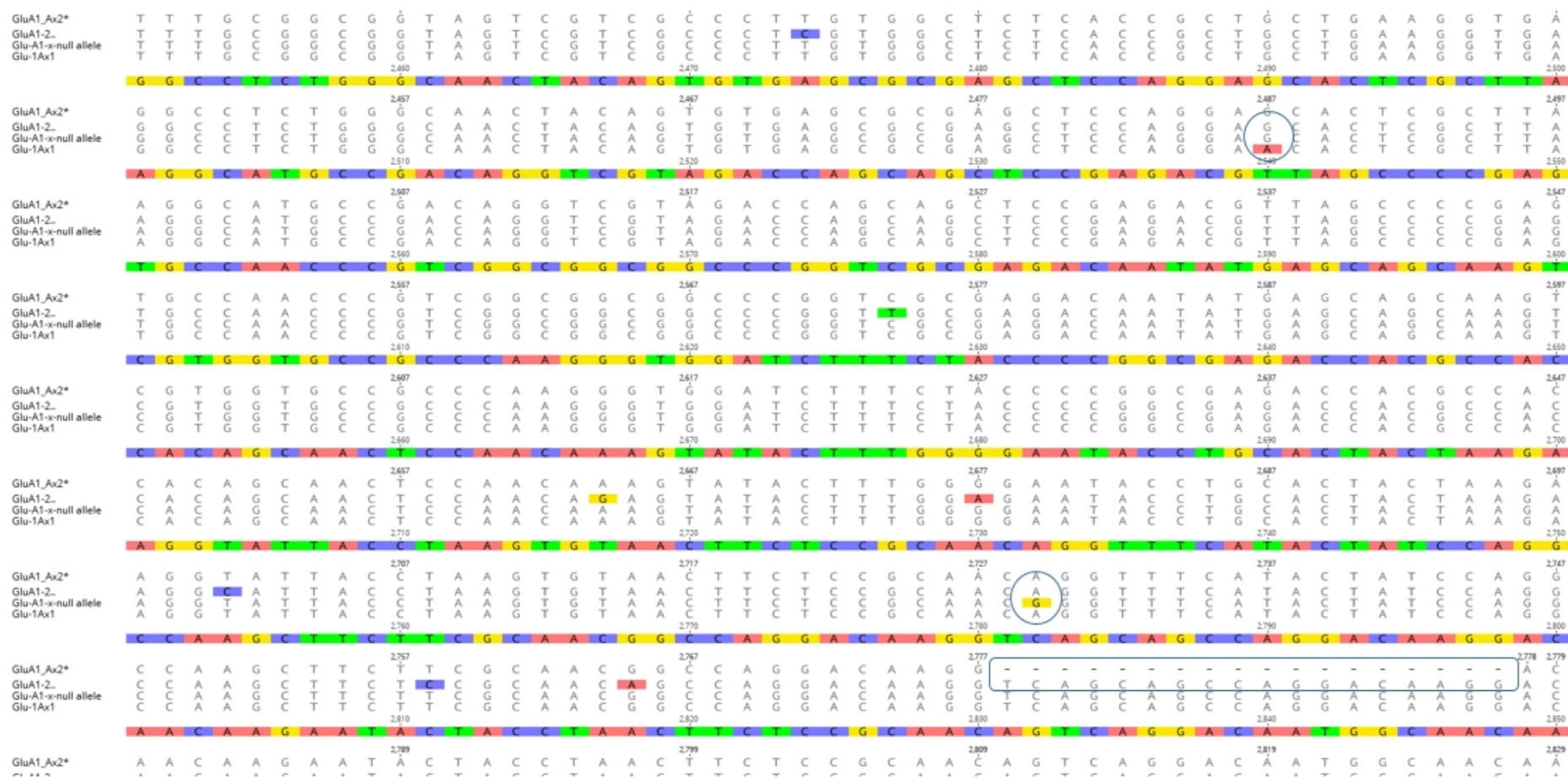


Figura 7. Identificación de polimorfismos para diferentes subunidades (Ax-2*, Ax-2-, Ax-null, Ax-1) del gen *GluA1-1* alineados con el software Geneious v8

Esta metodología se siguió también para los genes *GluD1-1*, *GluD1-2* y *GluB1-1*. En la Tabla 12 se muestran los diferentes marcadores que se han desarrollado con la secuencia diana y los polimorfismos asociados a cada subunidad. Se diseñaron un total de 13 marcadores, 10 de ellos para polimorfismos de tipo SNP y 3 para pequeños InDels.

No se encontraron en las bases de datos o en la bibliografía, secuencias de DNA para todos los alelos que tenían las variedades seleccionadas. Para completar las secuencias del gen *GluD1-1* se intentó secuenciar la subunidad Dx-4 mediante el diseño de cebadores conservados en todas las variedades. Se utilizó el software Geneious para alinear las secuencias encontradas en las bases de datos de las subunidades Dx-2, Dx-3 y Dx-5. En este alineamiento se identificó una región en la zona 5' común a las tres subunidades (Figura S1, ver Anexo) en la que se diseñaron los cebadores con los que se amplificaron todas las variedades, incluyendo 'Marius' y 'Champlein' portadoras de la subunidad Dx-4.

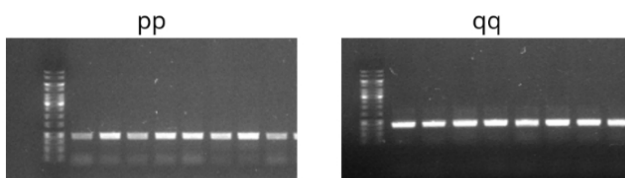


Figura 8. Amplificación de variedades de trigo con cebadores ppDx y qqDx

Las secuencias de los productos amplificados (Figura 8) se secuenciaron y se alinearon, pero no se identificó ningún polimorfismo presente en las variedades de 'Marius'

y 'Champlein' que permitieran distinguirlos del resto de variedades que presentan alelos diferentes.

Cuando se realizó un blast con las secuencias, se obtuvo un resultado inesperado. Además de presentar una alta homología con el gen *Glu-D1-1* (95,3%) (Figura 9) presentaban un nivel similar con secuencias correspondientes al gen *Glu-A1-1*. Es posible que a pesar de diseñar los cebadores con las secuencias correspondientes al genotipo D, debido a la alta homología, estos mismo cebadores sean capaces de amplificar el genotipo A de forma incluso más eficiente. Cuando se realizó la amplificación con muestras de trigo duro (genotipo AB) se obtuvo el mismo resultado lo que apoya la hipótesis de que debido a la alta homología de secuencias no se estaba amplificando realmente el genotipo deseado. No fue posible identificar otras regiones adecuadas para el diseño de cebadores con la información de secuencia disponible.

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
✓	Triticum aestivum clones BAC 1031P08; BAC 754K10; BAC 1344C16, complete sequence	Triticum aestivum	924	1370	100%	0.0	99.80%	292102	DQ537335.1
✓	Triticum aestivum high molecular weight glutenin subunit (HMW-Glu-1Ax) gene, promoter region	Triticum aestivum	918	918	100%	0.0	99.80%	2002	KC820627.1
✓	Triticum aestivum cultivar Xiaoyan 81 high molecular weight glutenin subunit 1Ax1 (Glu-1Ax1) gene, promoter region	Triticum aestivum	907	907	100%	0.0	99.20%	4526	KJ531446.1
✓	Triticum aestivum high molecular weight glutenin subunit (GluA1) gene, GluA1-2, allele, promoter region...	Triticum aestivum	907	907	100%	0.0	99.20%	4648	DQ533690.1
✓	Triticum aestivum HMW glutenin subunit Ax2* gene, complete cds	Triticum aestivum	896	896	100%	0.0	98.81%	6837	M22208.2
✓	Triticum aestivum Glu-1D-1d gene for high molecular weight glutenin subunit 5	Triticum aestivum	800	800	100%	0.0	95.43%	8463	X12928.5
✓	Triticum aestivum clone BAC 897M20, complete sequence	Triticum aestivum	800	1280	100%	0.0	95.43%	152010	DQ537337.1
✓	Triticum aestivum high molecular weight glutenin subunit (HMW-Glu-1Dx2) gene, promoter region	Triticum aestivum	795	795	100%	0.0	95.23%	2437	KC820631.1
✓	Triticum aestivum high molecular weight glutenin subunit 1Dx2 gene, promoter region and 5' UTR	Triticum aestivum	795	795	100%	0.0	95.23%	1252	DQ208971.1
✓	Triticum aestivum partial HMW-GS1Dx2 gene for high molecular weight glutenin subunit and promoter region	Triticum aestivum	784	784	100%	0.0	94.83%	1041	AJ577815.1
✓	Triticum aestivum Glu-1D-1 gene, promoter region	Triticum aestivum	784	784	100%	0.0	94.83%	1249	AJ301618.1

Figura 9. Análisis BLAST de la secuencia de la variedad 'Marius' (subunidad Dx-4) con la base de datos del NCBI

Las especies poliploides suponen un reto para la identificación de polimorfismos de tipo SNP debido a la presencia de varios genomas (Kaur et al, 2012). En el caso de los alopoloides, como el trigo, la presencia de los genomas homeólogos dificulta la identificación de secuencias específicas de genomio, y el correcto mapeo de las secuencias amplificadas. En el caso de los genes de las gluteninas es especialmente complicado, ya que además de la homeología se une la presencia de varias secuencias con alta homología en el mismo locus complejo. Por este motivo, el resultado obtenido en el análisis de la secuencia de la subunidad Dx-4 no es del todo sorprendente.

Aún a día de hoy, la región de los loci de gluteninas en el genoma de referencia del trigo no está bien anotada, lo que dificulta la asignación de secuencias a loci concretos. Una de las formas de asignar una secuencia a un genomio, es usar en la PCR líneas de sustitución en las que un par de cromosomas homólogos de trigo se han sustituido por un par de otra especie, por ejemplo, centeno (Scchapova et al, 1982). En el laboratorio se disponía de algunas de estas líneas de sustitución, pero no las de los cromosomas 1A, 1B y 1D por lo que no se pudo comprobar de esta forma.

Tabla 12. Marcadores KASP: tipo, cromosoma (Crom), secuencias diana para el diseño de los cebadores específicos y descripción del polimorfismo que detecta alelo HEX y el alelo FAM

KASP	Tipo	Crom	Diana	Alelo HEX		Alelo FAM	
GluA1x_null	SNP	A1	CCACGCCACCACAGCAACTCCAACAAAGTATACTTTGGGGAATACCTGCACTACTAAGAAGGTAT TACCTAAGTGTAACTTCTCGCAAC [G/A] GGTTCATACTATCCAGGCCAAGCTTCTTCGCAAC GGCCAGGACAAGGTCAGCAGCCAGGACAAGGACAACAAGAATACTACCTAACCTTC	G	null	A	otro
GluA1x_1	SNP	A1	TTTTGCGGCGGTAGTCGTCGCCCTTGTGGCTCTCACCGCTGCTGAAGGTGAGGCTCTGGGCAAC TACAGTGTGAGCGGAGCTCCAGGA [A/G] CACTCGCTTAAGGCATGCCGACAGGTCGTAGACCA GCAGCTCCGAGACGTTAGCCCCGAGTGCCAACCCGTCGGCGGGCGGCCGGTCCG	A	1	G	otro
GluA1x_2	InDel	A1	TGGCAACAATTAGGACAAGGGCAACAAGGGTACTACCCAACCTTCTCCGAGCAGTCAGGACAAGGGCAA CAGCCAGGACAATGGCTGCAA [CCAGGACAATGGCTGCAA/] TCAGGGTACTACCTAACTTCTCCGAGCA GTTAGGACAAGGGCAACAGCCAAGACAATGGCTGCAACCAAGACAAGGGCAACAAGGATAC	-	2..	CCAGGACAATGGCTGCAA	otro
GluA1x_2star	InDel	A1	TTCTCCACAGCAGTCAGGACAATTGCAACAACAGCACAAGGGCAGCAACCAGGACAAGAGCAAC AAGGTCAACAGCCAGGACAAGGGCA [ACAAGGTCAACAGCCAGGACAAGGGCA/] GCAACCGGGA CAAGGGCAGCCAGGGTACTACCCAACCTTCTCCGAGCAGTCAGGACAAGAGCAACAGCTAGAACA ATGGCAACAGTCAGG	-	2*	ACAAGGTCAACAGCCAGGACAAGGGCA	otro
GluB1x_6	SNP	B1	TTAGCAGACCGTCCAAAAATCTGTTTTGCAAGCACCAATGCTCCTTACTTATCCAGCTT [C/T] TTTTGTGTTGGCAAACTGCCCTTTTCCAACCGATTTTGTCTTCTCACGCTTTCTTCATA	T	6	C	otro
GluB1x_13	SNP	B1	ACGAGACAATACGAGCAGCAACCTGTGGTGCCTGCCGTCGAAGCCGGATCCTTCTACCCGAGC [G/A] AGACTACGCCCTTCCGAGCAACTCCAACAATGATATTTTGGGGAATACCTGCACTACTAA	A	13	G	otro
GluB1x_7oe	SNP	B1	ACAATAAATTATATGTGTTTTAGTTTTAGTTCTCATATCCAAATATACATGTTTCATACAACCA AATCTCATTAAATATATRTAAAATATCCRGCAACAACCTTGTGGGTACATCTAGTTACAGTG GAATATTAGTGATGGCGTGA [C/G] CAAGCGATAAGGCCAACGAGAGAAGAAGTGCCTGCTCTAT GGAGGCCAGGAAAAGACAATGGACATGCAAAGAGGTAGGGGCAGGGGAAGAAACACTTGGAGATCA TAGAAGAACATAAGAGGTTAAACATAGGAGGGCATAATGGACAAT	G	otro	C	7OE
GluB1x7_17	SNP	B1	CTCATCATACCCACAACACCGAGCACCCCAATCTACAGATCAATTCAGTACAGTTTAC [C/T] GAGATGGCTAAGCGCCTGGTCTCTTTGCGGCAGTAGTCGTCGCCCTCATGGCTCTCACC	C	otro	T	7/17
GluB1x14_20	SNP	B1	CTGACAGTTCACCGAGATGGCTAAGCGCCTGGTCTCTTTGCGGCAGTAGTCGTCGCCCTC [A/G] TGGCTCTACCGCCGCTGAAGGTGAGGCCTCTGGACAACATAATGTGAGC	A	14/20	G	otro
GluB1x_7	InDel	B1	TCACAGCAGTCAGGACAAGGGCATAATCAGGACAAGGGCAACAAGGGTACTACCCAACCTTCTCTGT GGCAACCAGGACAAGGGCAACAA [CCAGGACAAGGGCAACAA/] GGCTACGCCAGCCATACCATGTT AGCGCGGAGTACCAGGCGGCCCTAAAGGTGGCAAGGGCAGCAGCTCGCGGCACAGCTGCCG	-	7*	CCAGGACAAGGGCAACAA	otro
GluD1x_2	SNP	D1	CCGCGAGACCACGCCACCAGCAGCAACTCCAACAACGTATATTTTGGGGAATACCTGCACTACTA AAAAGGTATTACCCAAGTGAACCT [C/G] TCCGAGCAGGTTTCTACTATCCAGGCCAAGCTT CTCCGCAACGGCCAGGACAAGGTGAGCAGCCAGGACAAGGGCAACAATCAGGACA	G	5	C	2
GluD1x_5	SNP	D1	TAGATTTGCAAAAAGCAATGGCTAACAGACATATTTGCAAAACCCCAAGAAGGATAATCACTT TCTTAGATAAAAAAGAACAAGCA [A/G] TATACAAACATCCACACTCTGCAACAATAACATC AGAAGTATTACGCCGATACGTGGCTTTAGCAGACTGTCCAAAAATCTGTTT	G	2	A	5
GluD1y_12	SNP	D1	ATAGAGATCAATTCAGTACAGTCCACCGAGATGGCTAAGCGGCTGGTCTCTTTGCGGCAGTAG TCATCGCCCTCGTGGCTCTCACCAC [C/T] GCTGAAGGTGAGGCTCTAGGCAACTACAGTGTGA GCGCGAGCTCCAGGAGAGCTCGCTTGGGATGCCGCGAGGTCGTGGACCAACAG	C	12	T	10

3.5. Genotipado con marcadores moleculares KASP

Una vez obtenidos los cebadores, las reacciones de amplificación se optimizaron en el termociclador Roche LC96. Los resultados se muestran representados en forma de gráfica según la fluorescencia FAM o HEX a tiempo final.

3.5.1. KASP diseñados para identificar SNPs:

3.5.1.1. Marcadores para el gen *Glu-A1-1*

Se diseñaron 2 marcadores basados en SNPs para este gen, para la subunidad Ax-null y para la Ax-1.

GluA1x_null: este marcador funcionó según lo esperado para todas las variedades. Se puede distinguir una agrupación correspondiente a todas las variedades tanto de trigo duro como de trigo blando que presentaban la subunidad A1x-null para este gen.

GluA1x_1: en la Figura 10.B, se puede ver como este marcador distingue perfectamente las variedades (de trigo duro y trigo blando) que presentan la subunidad A1x-1 de las que no lo presentan.

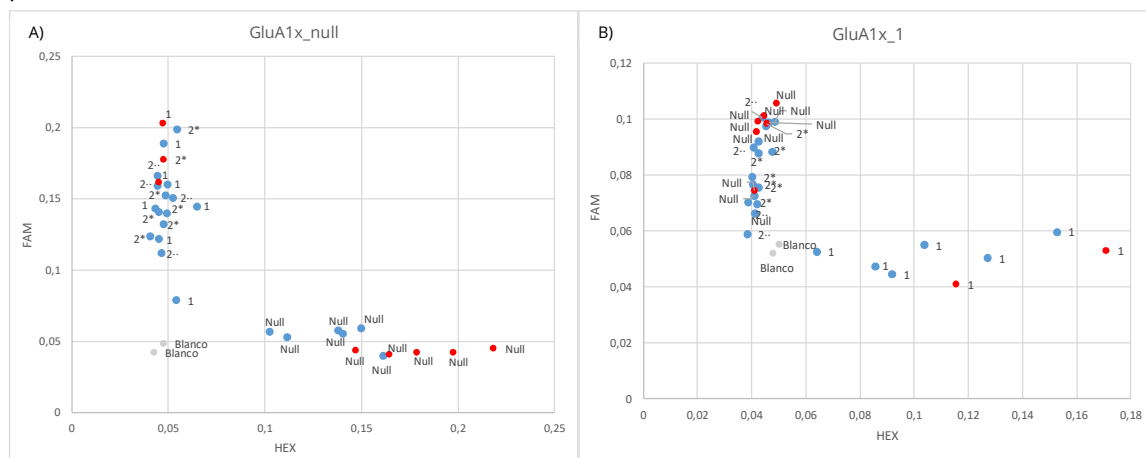


Figura 10. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM. A) Marcador GluA1x_null del gen *Glu-A1-1*. B) Marcador GluA1x_1 del gen *Glu-A1-1*

En un estudio previo realizado en trigo blando (Ravel et al, 2020), los marcadores que fueron diseñados para la identificación de las subunidades del gen *Glu-A1-1* también funcionaron adecuadamente, los resultados que se observaban en los geles SDS-PAGE correspondían con los que se obtenían cuando se utilizan los marcadores moleculares. Aunque había alguna excepción, en algunas variedades lo que se observaba en el perfil electroforético no correspondía con lo que se visualizaba cuando usaban los KASP, presentando una tasa de error del 1%. Los marcadores moleculares utilizados en este trabajo no presentan tasa de

la presencia de la subunidad Bx-7 respecto de otras subunidades, luego podría funcionar como un marcador específico de esta subunidad, pero en trigo blando no diferencia entre la subunidad Bx-7 y la subunidad Bx-7OE, además de que no agrupa dos variedades que presentan la subunidad Bx-7, 'Fiel' y 'Champlein'.

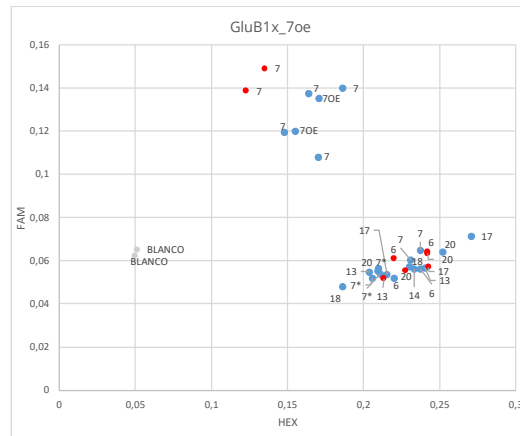


Figura 12. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM del marcador GluB1x_7oe del gen *Glu-B1-1*

La variante Bx7OE es característica de trigos bandos, se debe a la duplicación de gen GluB1-1 en esas variedades lo que conlleva la sobreexpresión de la subunidad Bx-7 lo que tiene un efecto positivo en la calidad (Han et al, 2014). Este marcador podría servir para diferenciar la Bx-7 en trigos duros, ya que no presentan la variante Bx-7OE, pero habría que analizar más muestras para confirmarlo.

GluB1x_7-17: en la Figura 13 se puede apreciar que el marcador es capaz de identificar las variedades con subunidades Bx-7 o Bx-17. Aunque sería necesario optimizarlo para que realice un mejor agrupamiento. En la Figura 13.A se puede ver como agrupa juntas las subunidades Bx-7OE y Bx-7*, igual que en el caso anterior, en trigo blando no se diferencian esas dos variantes. En el caso del trigo duro (Figura 13.B) la discriminación es incluso mejor.

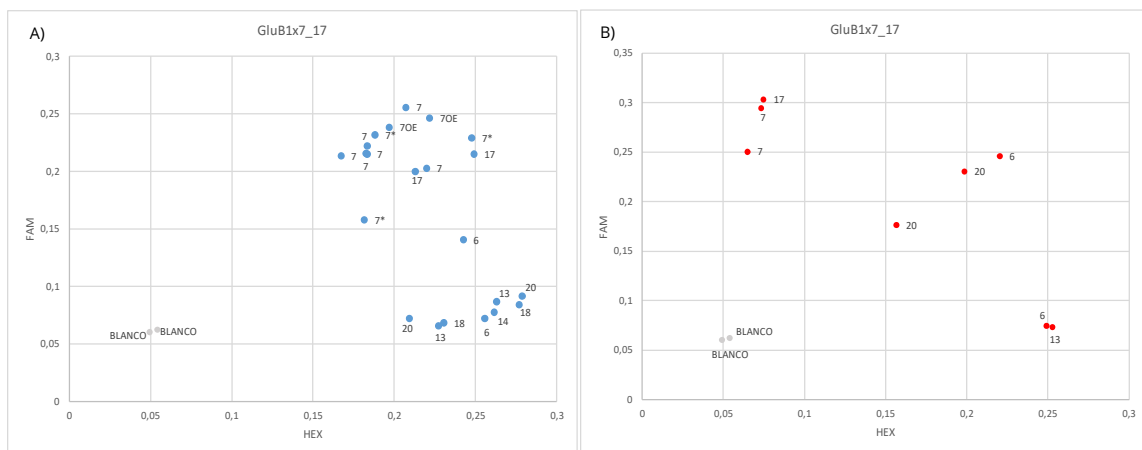


Figura 13. Distribución de las variedades de trigo según fluorescencia de los fluoróforos HEX o FAM A) Marcador GluB1x_7oe del gen *Glu-B1-1* en las variedades de trigo blando. B) Marcador GluB1x7_17 gen *Glu-B1-1* en las variedades de trigo duro

Ravel y colaboradores (2020) diseñaron marcadores moleculares para estas subunidades (Bx-7OE, Bx-7 y Bx-17). El marcador para detectar la subunidad Bx-7OE presentaba una baja tasa de error, y el marcador para la discriminación de las subunidades Bx-7 y Bx-17 presentaba una eficacia del 98%. En este caso, como se puede observar en las figuras anteriores, el marcador para la subunidad Bx-7OE no realiza una discriminación tan buena, y es que, aunque la región polimórfica sea similar (la diana que reconocen los marcadores moleculares es la misma que en el artículo de Ravel (2020)), el diseño del cebador no es el mismo. Mientras que el marcador GluB1x_7-17 parece funcionar de manera similar. Aunque ambos marcadores necesitan optimizarse para mejorar la agrupación de las diferentes subunidades.

GluB1x14_20: en este caso el marcador que se diseñó para distinguir la presencia la subunidad Bx-14 o Bx-20 del resto no funcionó según lo esperado, como se ve en la Figura 14 es capaz de distinguir la subunidad Bx-20 en trigo duro y trigo blando, de las cuatro variedades con las que se trabajan identificó todas, pero no es capaz de distinguir la presencia de la subunidad Bx-14.

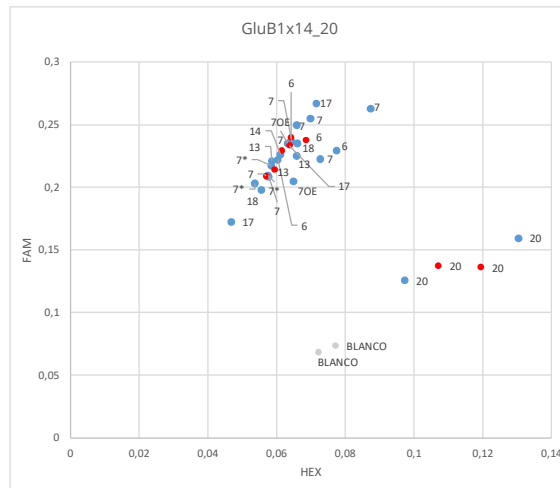


Figura 14. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM del marcador GluB1x14_20 del gen *Glu-B1-1*

En otros artículos (Ravel et al., 2020) los marcadores que se diseñaron para la detección de las subunidades Bx-14 o Bx-20 funcionaron de forma específica, detectando ambas subunidades en las variedades que las portaban. En este caso, como se observa en la Figura

10, el marcador solo es capaz de distinguir la subunidad Bx-20. Aun así, podría ser útil ya que la presencia de esta subunidad está relacionada con una mala calidad panadera.

3.5.1.3. Marcadores para el gen *Glu-D1*

Se diseñaron tres marcadores basados en SNPs para este gen, para identificar las subunidades Dx-5, Dx-2 y Dy-12

GluD1x_5: este marcador discriminó muy bien entre la subunidad Dx-5 y la subunidad Dx-2. Además, se puede ver en la Figura 15.A, que con las variedades que presenta la subunidad Dx-5 también se agrupaban las variedades de trigo duro y con las variedades que presentan la subunidad Dx-2 se agrupaban las que tienen la subunidad Dx-4.

GluD1x_2: se puede apreciar que este marcador también distingue correctamente las variedades que tienen la subunidad Dx-5 de las que tienen la subunidad Dx-2, excepto la variedad 'Gabo' que se agrupaba con los blancos (alelo nulo). Además, en este caso el trigo duro no se agrupa con ninguna subunidad, sino que se situaba con los blancos, a excepción de 'Langdon' que se agrupaba con las variedades que portan la subunidad Dx-5.

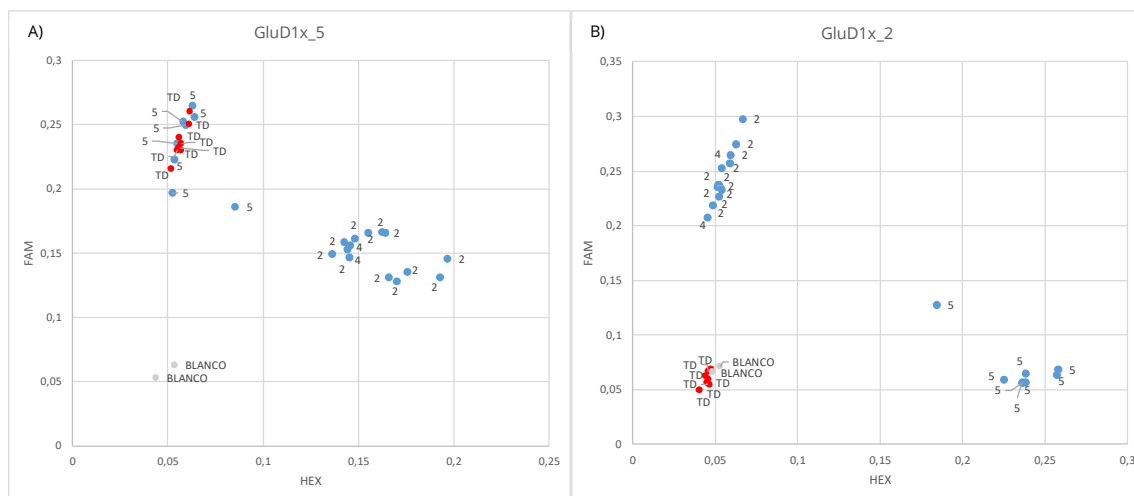


Figura 15. Genotipos de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM A) Marcador GluD1x_5 del gen *Glu-D1-1* B) Marcador GluD1x_2 del gen *Glu-D1-1*

GluD1y_12: en la Figura 16, se observa que el marcador funcionó de la forma adecuada, siendo capaz de hacer una distinción entre variedades de trigo con subunidades Dy-10 o Dy-12. Se puede apreciar la tendencia de los trigos duros a agruparse con las variedades que portan la subunidad Dy-12, aunque no deberían presentar amplificación.

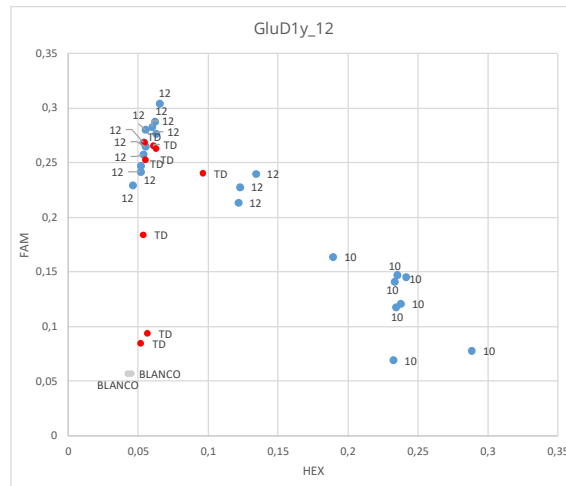


Figura 16. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM del marcador GluD1y_12 del gen *Glu-D1-2*

En artículos previos (Ravel et al., 2020) los marcadores diseñados para el locus *Glu-D1* detectaron las subunidades y se correspondían con las bandas electroforéticas. En las figuras anteriores, se puede visualizar como los marcadores también distinguen correctamente los alelos para este locus, aunque los marcadores GluD1x_5 y GluD1y_12 son más robustos y menos específicos que el marcador GluD1x_2, la posible razón por la que los marcadores GluD1x_5 y GluD1y_12 detecten el trigo duro es porque el trigo es una especie poliploide, y la diana de estos marcadores puede ser una secuencia que se encuentra repetida en los genomas A o B del trigo duro.

3.5.2. KASP diseñados para identificar InDels:

Se diseñaron 3 marcadores basados en InDels, dos para el gen *Glu-A1-1* y uno para el gen *Glu-B1-1*.

3.5.2.1. Gen *Glu-A1-1*

GluA1x_2: el marcador diseñado para identificar esta subunidad no funciona correctamente, como se puede ver en la Figura 17.A todas las variedades, indistintamente de que porten o no la subunidad Ax-2^o o no, se encontraban agrupadas juntos.

GluA1x_2star: en la Figura 17.B se puede apreciar como los cebadores diseñados para esta subunidad tampoco se comportaron según lo esperado, todas las variedades se encontraban agrupadas indistintamente de que porten o no la subunidad Ax-2*.

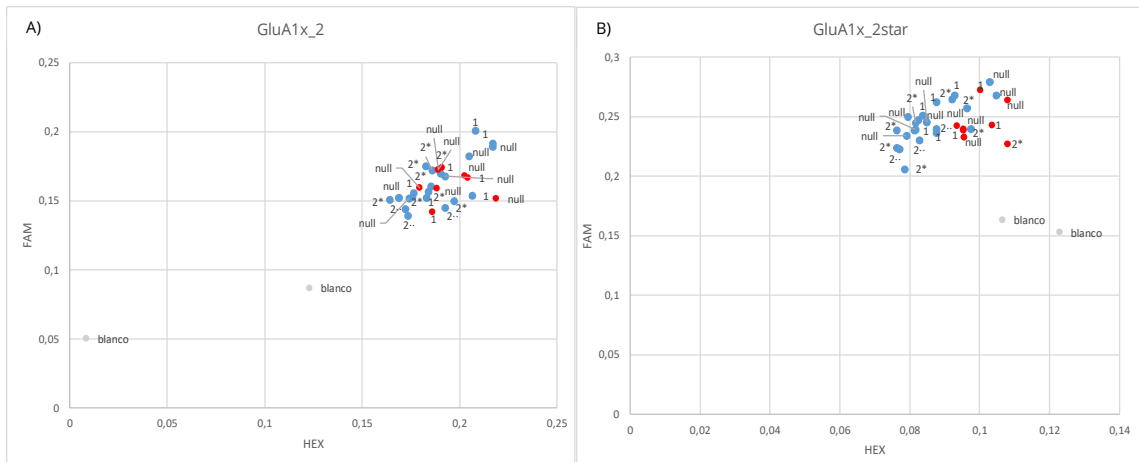


Figura 17. Distribución de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM A) Marcador GluA1x_2 del gen *Glu-A1-1* B) Marcador GluA1x_2star del gen *Glu-A1-1*

3.5.2.2. Gen *Glu-B1-1*

GluB1x_7: como ocurría con los cebadores diseñados para la identificación de las subunidades Ax-2* o Ax-2·, este marcador molecular tampoco funcionó correctamente. En la Figura 18, se puede ver que no era capaz de identificar la subunidad Bx-7* en ninguna variedad, ni en trigo duro ni trigo blando, y que todas las variedades volvían a estar agrupadas de forma inespecífica.

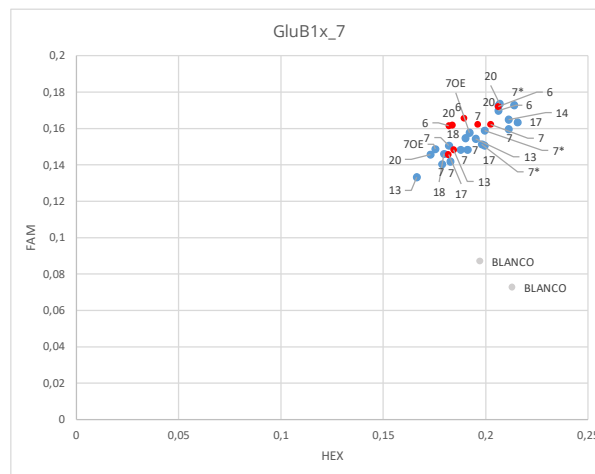


Figura 18. Genotipos de las variedades de trigo blando (azul) y trigo duro (rojo) según fluorescencia de los fluoróforos HEX o FAM del marcador GluB1x_7 del gen *Glu-B1-1*

Los marcadores moleculares diseñados a partir de pequeñas inserciones o deleciones no funcionaron de la manera esperada. Aunque los marcadores KASP están pensados para identificar SNPs, también se pueden usar en el caso de InDels. En otros artículos (Rasheed et al., 2016; Ravel et al., 2020) se diseñaron marcadores KASP cuya diana eran InDels y, aparentemente, hacían una buena discriminación alélica.

Hay que tener en cuenta que estos InDels, se corresponden con motivos repetidos en la secuencia de las gluteninas, lo que probablemente está limitando el diseño de los cebadores además de que pueden estar dándose amplificaciones inespecíficas (Figura S2, ver Anexo). Para intentar entender el diferente comportamiento de los dos tipos de marcadores, se realizaron unas pruebas de amplificación seguida de desnaturalización.

3.6. Curva de disociación

La temperatura de desnaturalización del ADN puede variar mucho dependiendo de la secuencia, la longitud y el contenido de CG. Por eso, esta técnica es útil para analizar la especificidad de los productos obtenidos de una amplificación por PCR a través de sus perfiles de desnaturalización.

Se emplearon los cebadores de los ensayos KASP de InDels y se realizaron amplificaciones con SYBR Green para obtener posteriormente las curvas de disociación. En la Figura 19 se muestran las cinéticas de amplificación (a la izquierda del panel) y las curvas de disociación de los productos amplificados (a la derecha del panel) para un marcador basado en SNPs y un marcador basado en InDels.

En la 19.A y B, se muestra el resultado para el marcador molecular GluA1x_1 que generó una curva de disociación en la que se observa claramente un único pico, indicando la presencia de un único producto amplificado.

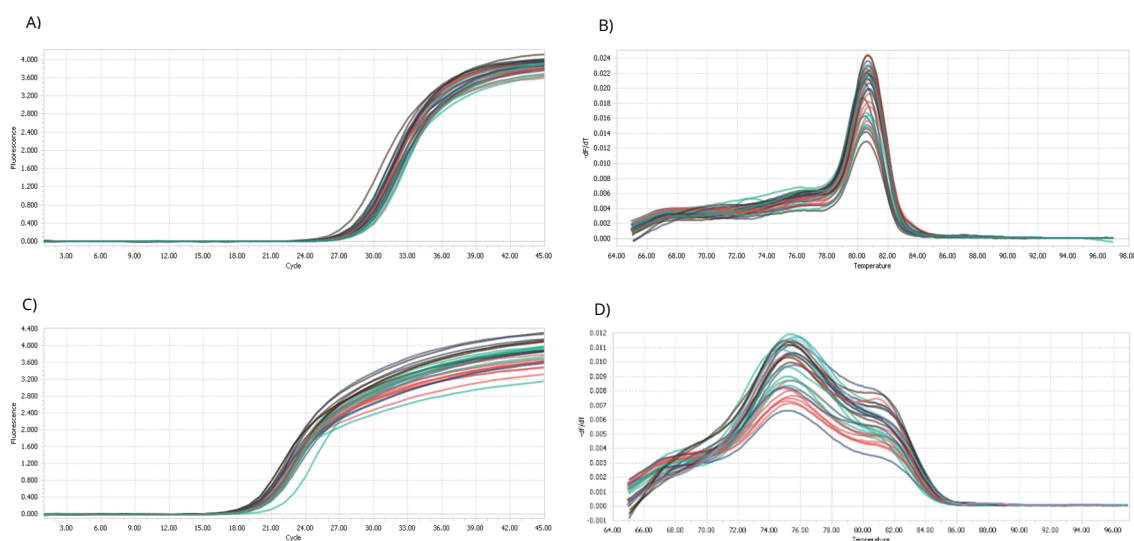


Figura 19. Cinéticas de amplificación y disociación de marcadores KASP A) Curva de amplificación de las 30 variedades de trigo con el marcador GluA1x_1 (*Glu-A1-1*) diseñado B) Curva de disociación de la amplificación con el marcador GluA1x_1. C) Curva de amplificación de las 30 variedades de trigo con el marcador GluA1x_2star (*Glu-A1-1*) D) Curva de disociación de la amplificación con el marcador GluA1x_2star.

En la Figura 19 C y D, se muestra el resultado para uno de los marcadores basados en InDels que generó una curva de disociación totalmente diferente, indicativa de que no se estaba produciendo una amplificación específica, de un único fragmento, en la reacción. Esta prueba se realizó con los tres marcadores diseñados con InDels, y con los tres se visualizaba una curva del mismo tipo, con al menos dos picos sugiriendo que el problema que presentaban esos marcadores se debía en todos los casos a la obtención de varios amplicones, independientemente del genotipo de las variedades. Esto puede ser debido a varias razones. Una de ellas puede ser que al ser el trigo una especie poliploide, el marcador KASP puede estar detectando regiones similares en distintos genomios; la otra posibilidad puede que sea que los marcadores KASP diseñados no sean totalmente específicos del polimorfismo y se amplifiquen otras regiones en el mismo gen.

CAPÍTULO 4: CONCLUSIONES

- Se ha seleccionado y verificado un set de variedades de trigo que se puede considerar una “master set” de variabilidad para HMW-GS útil para la puesta a punto de marcadores moleculares.
- Se han conseguido desarrollar y poner a punto marcadores moleculares para HMW-GS relacionados con alelos de buena y mala calidad en trigo blando, que pueden ser utilizados para seleccionar materiales en programas de mejora.
- Se han analizado por primera vez marcadores KASP relacionados con HMW-GS en trigo duro, obteniéndose unos resultados comparables a trigo blando, aunque es necesario confirmar estos resultados en un mayor número de variedades.
- La información de HMW-GS disponible en las bases de datos no es suficiente, es necesario generar más información de secuencia para poder diseñar nuevos marcadores que identifiquen otras variantes.
- Los KASP diseñados a partir de InDels, no han dado el resultado esperado, generando amplificaciones inespecíficas que posiblemente estén relacionadas con la presencia de motivos repetidos en estos genes.

CAPÍTULO 5: BIBLIOGRAFÍA

- Anderson, O. D., & Greene, F. C. (1989). The characterization and comparative analysis of high-molecular-weight glutenin genes from genomes A and B of a hexaploid bread wheat. *Theoretical and Applied Genetics*, 77(5), 689-700.
- Anderson, O. D. (2009). EST mining for structure and expression of genes in the region of the wheat high-molecular-weight glutenin loci. *Genome*, 52(8), 726-740.
- Anderson, O. D., Greene, F. C., Yip, R. E., Halford, N. G., Shewry, P. R., & Malpica-Romero, J. M. (1989). Nucleotide sequences of the two high-molecular-weight glutenin genes from the D-genome of a hexaploid bread wheat, *Triticum aestivum* L. cv Cheyenne. *Nucleic Acids Research*, 17(1), 461.
- De Bustos, A., Rubio, P., & Jouve, N. (2000). Molecular characterisation of the inactive allele of the gene Glu-A1 and the development of a set of AS-PCR markers for HMW glutenins of wheat. *Theoretical and Applied Genetics*, 100(7), 1085-1094.
- Dreisigacker, S., Xiao, Y., Sehgal, D., Guzman, C., He, Z., Xia, X., & Peña, R. J. (2020). SNP markers for low molecular weight glutenin subunits (LMW-GSs) at the *Glu-A3* and *Glu-B3* loci in bread wheat. *PLoS ONE*, 15(5), 1-16. <https://doi.org/10.1371/journal.pone.0233056>
- Espí Plaza, A. (2013). *Estudio molecular de gluteninas de alto y bajo peso molecular en Triticum aestivum ssp. vulgare L. y su relación con la calidad panadera*. Tesis (Doctoral), E.T.S.I. Agrónomos (UPM). http://oa.upm.es/21603/1/ARACELI_ESPI_PLAZA.pdf
- Espí, A., Giraldo, P., Rodríguez-Quijano, M., & Carrillo, J. M. (2012). A PCR-based method for discriminating between high molecular weight glutenin subunits Bx7 and Bx7* in *Triticum aestivum* L. *Plant breeding*, 131(5), 571-573.
- Espinosa, N. H., Reyes, M. R., Jiménez, F. E. G., Bretón, L. C. N., & Bribiesca, B. L. C. (2015). Importancia de las proteínas de almacenamiento en cereales (prolaminas). *VERTIENTES Revista Especializada En Ciencias de La Salud*, 18(1), 3-7.
- FAOSTAT, <<http://www.fao.org/faostat/es/#data/QC>> Datos de 2019 [Consulta Abril 2021]
- Filip E (2018) Composition of high molecular weight glutenin subunits in polish common wheat cultivars (*Triticum aestivum* L.). *J Food Quality*. <https://doi.org/10.1155/2018/2473420>
- Gale, K. R. (2005). Diagnostic DNA markers for quality traits in wheat. *Journal of Cereal Science*, 41(2), 181-192. <https://doi.org/10.1016/j.jcs.2004.09.002>
- Gálová, Z., Michalík, I., Knoblochová, H., & Gregová, E. (2002). Variation in HMW glutenin subunits of different species of wheat. *Rostlinna Vyroba*, 48(1), 15-19. <https://doi.org/10.17221/4199-pse>
- Giraldo, P., Rodríguez-Quijano, M., Simon, C., Vazquez, JF, Carrillo JM (2010). Allelic variation in HMW glutenins in Spanish wheat landraces and their relationship with bread quality. *Spanish journal of agricultural research*, (4), 1012-1023.
- Gobaa, S., Kleijer, G., & Stamp, P. (2007). 2⁺, a new high molecular weight glutenin subunit coded by Glu-A1: its predicted structure and its impact on bread-making quality. *Plant breeding*, 126(1), 1-4.
- Gu, Y. Q., Salse, J., Coleman-Derr, D., Dupin, A., Crossman, C., Lazo, G. R., Naxin Huo, Harry Belcram, Catherine Ravel, Gilles Charmet, Mathieu Charles, Olin D. Anderson, Chalhoub, B. (2006). Types and rates of sequence evolution at the high-molecular-weight glutenin locus in hexaploid wheat and its ancestral genomes. *Genetics*, 174(3), 1493-1504.

- Guo, X., Sun, X., Zhang, Y., Wang, R., & Yan, X. (2018). Interactions between soy protein hydrolyzates and wheat proteins in noodle making dough. *Food Chemistry*, 245(July 2017), 500-507. <https://doi.org/10.1016/j.foodchem.2017.10.126>
- Halford, N. G., Field, J. M., Blair, H., Urwin, P., Moore, K., Robert, L., ... & Shewry, P. R. (1992). Analysis of HMW glutenin subunits encoded by chromosome 1A of bread wheat (*Triticum aestivum* L.) indicates quantitative effects on grain quality. *Theoretical and Applied Genetics*, 83(3), 373-378.
- Ibba, M. I., Kiszonas, A. M., & Morris, C. F. (2018). Development of haplotype-specific molecular markers for the low-molecular-weight glutenin subunits. *Molecular Breeding*, 38(5). <https://doi.org/10.1007/s11032-018-0827-9>
- Jin, H., Yan, J., Peña, R. J., Xia, X. C., Morgounov, A., Han, L. M., Zhang, Y., & He, Z. H. (2011). Molecular detection of high- and low-molecular-weight glutenin subunit genes in common wheat cultivars from 20 countries using allele-specific markers. *Crop and Pasture Science*, 62(9), 746-754. <https://doi.org/10.1071/CP11134>
- Kaur, S., Francki, M. G., and Forster, J. W. (2012). Identification, characterization and interpretation of single-nucleotide sequence variation in allopolyploid crop species. *Plant Biotechnol. J.* 10, 125-138. doi: 10.1111/j.1467-7652.2011.00644.x
- Li, J., Han, C., Zhen, S., Li, X., & Yan, Y. (2014). Characterization of HMW glutenin subunit Bx7OE and its distribution in common wheat and related species. *Plant Genetic Resources*, 12(2), 191.
- Li, Y., Fu, J., Shen, Q., & Yang, D. (2021). High-molecular-weight glutenin subunits: Genetics, structures, and relation to end use qualities. *International Journal of Molecular Sciences*, 22(1), 1-14. <https://doi.org/10.3390/ijms22010184>
- Liu, L., Ikeda, T. M., Branlard, G., Peña, R. J., Rogers, W. J., Lerner, S. E., Kolman, M. A., Xia, X., Wang, L., Ma, W., Appels, R., Yoshida, H., Wang, A., Yan, Y., & He, Z. (2010). Comparison of low molecular weight glutenin subunits identified by SDS-PAGE, 2-DE, MALDI-TOF-MS and PCR in common wheat. *BMC Plant Biology*, 10. <https://doi.org/10.1186/1471-2229-10-124>
- Liu, S., Chao, S., & Anderson, J. A. (2008). New DNA markers for high molecular weight glutenin subunits in wheat. *Theoretical and Applied Genetics*, 118(1), 177-183. <https://doi.org/10.1007/s00122-008-0886-0>
- López-Fernández, M.; Pascual, L.; Faci, I.; Fernández, M.; Ruiz, M.; Benavente, E.; Giraldo, P. Exploring the End-Use Quality Potential of a Collection of Spanish Bread Wheat Landraces. *Plants* 2021, 10, 620. <https://doi.org/10.3390/plants10040620>
- Metakovsky, E., Melnik, V., Pascual, L., Romanov, G. A., & Wrigley, C. W. (2019). Types, frequencies and value of intra-varietal genotypic non-uniformity in common wheat cultivars: Authentic biotypes and foreign seeds. *Journal of Cereal Science*, 89, 102813.
- Payne, P. I., Law, C. N., & Mudd, E. E. (1980). Control by homoeologous group 1 chromosomes of the high-molecular-weight subunits of glutenin, a major protein of wheat endosperm. *Theoretical and applied genetics*, 58(3-4), 113-120
- Payne, P. I., Holt, L. M., & Law, C. N. (1981). Structural and genetical studies on the high-molecular-weight subunits of wheat glutenin - Part 1: Allelic variation in subunits amongst varieties of wheat (*Triticum aestivum*). *Theoretical and Applied Genetics*, 60(4), 229-236. <https://doi.org/10.1007/BF02342544>
- Rasheed, A., Wen, W., Gao, F., Zhai, S., Jin, H., Liu, J., Guo, Q., Zhang, Y., Dreisigacker, S., Xia, X., & He, Z. (2016). Development and validation of KASP assays for genes underpinning key economic traits in

- bread wheat. *Theoretical and Applied Genetics*, 129(10), 1843–1860. <https://doi.org/10.1007/s00122-016-2743-x>
- Ravel, C., Faye, A., Ben-Sadoun, S., Ranoux, M., Dardevet, M., Dupuits, C., Exbrayat, F., Poncet, C., Sourdille, P., & Branlard, G. (2020). SNP markers for early identification of high molecular weight glutenin subunits (HMW-GSs) in bread wheat. *Theoretical and Applied Genetics*, 133(3), 751–770. <https://doi.org/10.1007/s00122-019-03505-y>
- Ravel, C., Fiquet, S., Boudet, J., Dardevet, M., Vincent, J., Merlino, M., Michard, R & Martre, P. (2014). Conserved cis-regulatory modules in promoters of genes encoding wheat high-molecular-weight glutenin subunits. *Frontiers in plant science*, 5, 621.
- Sharma, A., Garg, S., Sheikh, I., Vyas, P., & Dhaliwal, H. S. (2020). Effect of wheat grain protein composition on end-use quality. *Journal of Food Science and Technology*, 57(8), 2771–2785. <https://doi.org/10.1007/s13197-019-04222-6>
- Shchapova, A. and Kravtsova, L., The Production of Wheat–Rye Substitution Lines by Using the Giemsa Staining Technique, *Cereal Res. Commun.*, 1982, vols. 1–2, pp. 33–39.
- Singh, N. K., Shepherd, K. W., & Cornish, G. B. (1991). A Simplified SDS-PAGE Procedure for Separating. *Journal of cereal science*, 14, 203–208.
- Shewry, P. R. (2009). Wheat. *Journal of Experimental Botany*, 60(6), 1537–1553. <https://doi.org/10.1093/jxb/erp058>
- Shewry, Peter R., Halford, N. G., & Lafiandra, D. (2003). Genetics of Wheat Gluten Proteins. *Advances in Genetics*, 49, 111–184. [https://doi.org/10.1016/S0065-2660\(03\)01003-4](https://doi.org/10.1016/S0065-2660(03)01003-4)
- Waines, J. G., & Payne, P. I. (1987). Electrophoretic analysis of the high-molecular-weight glutenin subunits of *Triticum monococcum*, *T. urartu*, and the A genome of bread wheat (*T. aestivum*). *Theoretical and Applied Genetics*, 74(1), 71–76. <https://doi.org/10.1007/BF00290086>
- Wrigley, C. W., & Morris, C. F. (1996). Breeding cereals for quality improvement. *Cereal Grain Quality*, 321–369. https://doi.org/10.1007/978-94-009-1513-8_11


```

Bx7_DQ119142 GA GATGGCTAA GCGCCTGGT CCTCTTT GCGGCA GTAGT CGTCCGCTCTCGGCTCTCACCGCCGCTGAAGTGAGGCTCTGGCAACTA 90
Bx7*_X13927 ..... 90
Bx13_EF540764 ..... 90
Bx7_DQ119142 CAATGTGAGCAGAGGCTCGAGGCATGCCAACAGGTGGTGGACCGCAACTCCGAGACGTTAGCCCGGGTGC CGCCCATCACCGTCAGC 180
Bx7*_X13927 ..... 180
Bx13_EF540764 ..... 180
Bx7_DQ119142 CCGGGCAGAGCAATAACGAGCAACCTGTGGTGGCTCCAAAGCGGATCGTCTACCCCAAGCAGACTACGCTTCGAGCAACTC 270
Bx7*_X13927 ..... 270
Bx13_EF540764 ..... 270
Bx7_DQ119142 CAACAAATGATATTTGGGGAATACCTGCACACTAAAGAAAGTATTACCCAAGTGTAACTTCTCGCAGCAGGGGTCACTACTACAGCG 360
Bx7*_X13927 ..... 360
Bx13_EF540764 ..... 360
Bx7_DQ119142 CAAGCTTCTCCCAACAGTCAAGCAAGGACAGGACAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 450
Bx7*_X13927 ..... 450
Bx13_EF540764 ..... 450
Bx7_DQ119142 AGACAACAAAGGATACTACCCAACTTCTCGGCAACAGGCAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 540
Bx7*_X13927 ..... 540
Bx13_EF540764 ..... 540
Bx7_DQ119142 CACGAGCAGGACAAAAGCAGGACAGGACAGGCAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 630
Bx7*_X13927 ..... 630
Bx13_EF540764 ..... 630
Bx7_DQ119142 GGAACAAGGGCAACAACCGGCAAGGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 720
Bx7*_X13927 ..... 720
Bx13_EF540764 ..... 720
Bx7_DQ119142 CAACAACCAAGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 810
Bx7*_X13927 ..... 810
Bx13_EF540764 ..... 810
Bx7_DQ119142 TACTACCCAATTTCTCCGCAACAGCCGGCAAGGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 900
Bx7*_X13927 ..... 900
Bx13_EF540764 ..... 900
Bx7_DQ119142 GGAACAATGGCAGCAACCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 990
Bx7*_X13927 ..... 990
Bx13_EF540764 ..... 990
Bx7_DQ119142 CAACAAGGATACTACCCAACCTCTCTGCAACAGCCAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 1080
Bx7*_X13927 ..... 1080
Bx13_EF540764 ..... 1080
Bx7_DQ119142 CAGTCGGAAACAAGGGCAAGCAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 1170
Bx7*_X13927 ..... 1170
Bx13_EF540764 ..... 1170
Bx7_DQ119142 CAAGGGCAACAACCTGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAG 1260
Bx7*_X13927 ..... 1260
Bx13_EF540764 ..... 1260
Bx7_DQ119142 CAAGGGTACTACCCAACCTCTCCGCAACAGTCAGGCAAGGGCAACAACCGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGG 1350
Bx7*_X13927 ..... 1350
Bx13_EF540764 ..... 1350
Bx7_DQ119142 CAGTCAGGCAACAAGGGCAAGCAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 1440
Bx7*_X13927 ..... 1440
Bx13_EF540764 ..... 1440
Bx7_DQ119142 TACCCAACCTCTCTCGCAACAAGTCAAGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 1530
Bx7*_X13927 ..... 1530
Bx13_EF540764 ..... 1530
Bx7_DQ119142 TCTCCACAGCAGCAGGCAACAAGGCAACAATCAGGCAAGGCAACAATCAGGCAAGGCAACAATCAGGCAAGGCAACAATCAGGCAAGGCAACA 1620
Bx7*_X13927 ..... 1620
Bx13_EF540764 ..... 1620
Bx7_DQ119142 CAGCCAGGCCAATTGCAACAACAGCAACAAGGGCAACAACAGGCAACAAGGGCAACAATCAGGCAACAAGGCAACAAGGCAACAAGGCAACA 1710
Bx7*_X13927 ..... 1710
Bx13_EF540764 ..... 1710
Bx7_DQ119142 CAATCAGGCAATGGCAACTAGTGTACTACCCAACCTCTCCGCAACAGCCAGGCAACAATGCAACAACAGGCAACAAGGGCAACAAGGGGTAC 1800
Bx7*_X13927 ..... 1800
Bx13_EF540764 ..... 1800
Bx7_DQ119142 TACCCAACCTCTCCCAACAAGTCAGGCAAGGGCAACAAGGGTACTACCCAACCTCTCCGCAACAGTCAGGCAAGGGCAACAAGGGTAC 1890
Bx7*_X13927 ..... 1890
Bx13_EF540764 ..... 1890
Bx7_DQ119142 TACCCAACCTCTCCGCAACAAGTCAGGCAAGGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 1980
Bx7*_X13927 ..... 1980
Bx13_EF540764 ..... 1980
Bx7_DQ119142 TCTCCGCAAGTCAAGGCAACAACAGGCAACAAGGGCAACAAGGGTACTACCCAACCTCTCCGCAAGCAGTCAGGCAAGGGCAACAAGGGGAA 2070
Bx7*_X13927 ..... 2070
Bx13_EF540764 ..... 2070
Bx7_DQ119142 CAACAAGGATGAGCAACAGGCAACAAGGGCAACAAGGGTACTATCCAACCTCTCACAGCAGTCAGGCAAGGGCAACAAGGGTAC 2160
Bx7*_X13927 ..... 2160
Bx13_EF540764 ..... 2160
Bx7_DQ119142 CAAGGCATCAATCAGGCAACAAGGGCAACAAGGGTACTACCCAACCTCTCTGTGGCAACAGGCAACAAGGGCAACAAGGGCAACAAGGGCAACA 2250
Bx7*_X13927 ..... 2235
Bx13_EF540764 ..... 2250
Bx7_DQ119142 CAAGGCTACGCCAGCCATACCATGTTAGCGGGAGTACCAAGCGGCCCGCCTAAAGGTGGCAAGGGCAAGGGCAAGGGCAAGGGCAAGGGT 2340
Bx7*_X13927 ..... 2322
Bx13_EF540764 ..... 2340
Bx7_DQ119142 CCGGCAATGTGCGGGCTGGAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 2430
Bx7*_X13927 ..... 2412
Bx13_EF540764 ..... 2394
Bx7_DQ119142 GCACCTTAGCTATACAATAAAGC 2453
Bx7*_X13927 ..... 2435
Bx13_EF540764 ..... 2394

```

Figura suplementaria 2. Alineamiento de las secuencias de las subunidades Bx-13, Bx-7 y Bx-7*. Se señalan subrayados el inicio y fin de la secuencia codificante. El polimorfismo que distingue entre las subunidades Bx-7 y Bx-7* se marca en amarillo. En verde se marcan las repeticiones completas de la secuencia del InDel, y en naranja algunas repeticiones casi completas (a excepción del nucleótido 5').