



# **COLEGIO DE POSTGRADUADOS**

**INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS**

**CAMPUS MONTECILLO**

**POSTGRADO EN RECURSOS GENÉTICOS Y PRODUCTIVIDAD**

**GENÉTICA**

## **VALIDACIÓN DEL QTL *HAP-6A-13* ASOCIADO A PESO DE GRANO EN TRIGOS HARINEROS DE PRIMAVERA**

**GUILLERMO GARCÍA BARRIOS**

**T E S I S**

**PRESENTADA COMO REQUISITO PARCIAL**

**PARA OBTENER EL GRADO DE:**

**MAESTRO EN CIENCIAS**

**MONTECILLO, TEXCOCO, ESTADO DE MÉXICO**

**2021**

La presente tesis titulada: Validación del QTL *Hap-6A-13* asociado a peso de grano en trigos harineros de primavera, realizada por el alumno: Guillermo García Barios, bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRO EN CIENCIAS  
RECURSOS GENÉTICOS Y PRODUCTIVIDAD

GENÉTICA

CONSEJO PARTICULAR



CONSEJERO

---

DR. SERAFIN CRUZ IZQUIERDO



ASESORA

---

DRA. SUSANNE DREISIGACKER



ASESOR

---

DR. RICARDO LOBATO ORTIZ

Montecillo, Texcoco, Estado de México, julio de 2021

# VALIDACIÓN DEL QTL *HAP-6A-13* ASOCIADO A PESO DE GRANO EN TRIGOS HARINEROS DE PRIMAVERA

Guillermo García Barrios, M.C.

Colegio de Postgraduados, 2021

## RESUMEN

El peso del grano es un componente estable del rendimiento de grano en el trigo (*Triticum aestivum*). Los marcadores moleculares ligados con el peso del grano pueden usarse en la selección asistida por marcadores para acelerar las ganancias genéticas para este carácter. Este estudio se realizó para validar el QTL *HAP-6A-13*, el cual se reportó asociado a peso de grano en trigos harineros de primavera en CIMMYT. Para realizar la validación se diseñaron dos marcadores de competencia alelo específica (KASP - kompetitive allele specific PCR) para los SNP T/C y A/G que forman el QTL. La asociación de los marcadores KASP con el peso de grano y rendimiento se verificó mediante estudios de asociación del genoma completo (GWAS - genome-wide association study) en cuatro colecciones de trigo. Los dos marcadores KASP mostraron asociación significativa con el peso de grano y rendimiento. Los análisis basados en haplotipos mostraron mejor resultados GWAS que los análisis con los marcadores SNP individuales. Adicionalmente se generaron y evaluaron diez familias endogámicas heterogéneas (HIF - heterogeneous inbred family) para peso de mil granos, rendimiento y otros caracteres morfométricos del grano. Las evaluaciones fenotípicas se realizaron en el noreste y centro de México. En contraste con los resultados GWAS las evaluaciones de HIF mostraron efectos inconsistentes del QTL para los caracteres rendimiento, peso, ancho y grosor de grano. Los genotipos con el haplotipo TG fueron superiores solo en longitud de grano en los dos ambientes, pero con poca significancia, lo que demuestra la dificultad de validar un QTL de efecto menor en trigo.

**Palabras clave:** *Triticum aestivum*, validación de QTL, peso de mil granos, haplotipos SNP, familias endogámicas heterogéneas.

# VALIDATION OF QTL *HAP-6A-13* ASSOCIATED WITH GRAIN WEIGHT IN SPRING BREAD WHEAT

Guillermo Garcia Barrios, M.C.

Colegio de Postgraduados, 2021

## ABSTRACT

Grain weight is a main component of grain yield in wheat (*Triticum aestivum*). Molecular markers that are linked to grain weight might be used in marker-assisted selection to accelerate genetic gains for this trait. The current study was carried out to validate the QTL *HAP-6A-13*, which was reported to be associated to grain weight in CIMMYT spring bread wheat. For validation two competitive allele specific PCR (KASP) markers were designed for the SNPs T/C and A/G comprising the QTL. The association of the markers KASP with thousand kernel weight and grain yield was verified by genome-wide association study (GWAS) in four independent datasets. The two KASP markers showed significant association with thousand kernel weight and grain yield. Haplotype-based analyzes showed better GWAS results than analyses with single SNP markers. In addition, ten heterogeneous inbred families (HIF) were generated and evaluated for the thousand kernel weight, grain yield and other morphometric traits of the grain. Phenotypic evaluations were conducted in the northeastern and central Mexico. In contrast to the GWAS datasets, the evaluations of HIF showed inconsistent effects of the QTL for grain yield, thousand grain weight, width, and thickness. The genotypes carrying the TG haplotype were only superior for grain length in the two environments, but with low significance, which show, the difficulty of validating a minor QTL in wheat.

**Key words:** *Triticum aestivum*, QTL validation, thousand kernel weight, SNP haplotypes, heterogeneous inbred families.

## **AGRADECIMIENTOS**

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada para los estudios de maestría.

Al Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) y sus investigadores que fueron una inspiración durante mi formación.

Al Colegio de Postgraduados.

Al Campus Montecillo.

Al Posgrado en Recursos Genéticos y Productividad.

## CONTENIDO

<b>RESUMEN</b> .....	<b>iii</b>
<b>ABSTRACT</b> .....	<b>iv</b>
<b>AGRADECIMIENTOS</b> .....	<b>v</b>
<b>LISTA DE CUADROS</b> .....	<b>viii</b>
<b>LISTA DE FIGURAS</b> .....	<b>ix</b>
<b>I. INTRODUCCIÓN</b> .....	<b>1</b>
<b>II. OBJETIVOS E HIPÓTESIS</b> .....	<b>2</b>
<b>Objetivo general</b> .....	<b>2</b>
<b>Objetivos específicos</b> .....	<b>2</b>
<b>Hipótesis</b> .....	<b>2</b>
<b>III. REVISIÓN DE LITERATURA</b> .....	<b>2</b>
<b>Mejoramiento de caracteres cuantitativos</b> .....	<b>2</b>
<b>Peso de grano: un carácter complejo</b> .....	<b>3</b>
<b>Caracteres morfométricos del grano</b> .....	<b>6</b>
<b>Estudio de asociación del genoma completo</b> .....	<b>7</b>
<b>Prueba de asociación</b> .....	<b>7</b>
<b>Estructura de la población</b> .....	<b>9</b>
<b>Validación de QTL obtenido mediante GWAS para su uso en mejoramiento genético</b> .....	<b>9</b>
<b>Replicación estadística</b> .....	<b>10</b>
<b>Validación en líneas casi isogénicas</b> .....	<b>10</b>
<b>líneas endogámicas retrocruzadas</b> .....	<b>10</b>
<b>Familias endogámicas heterogéneas</b> .....	<b>11</b>
<b>IV. MATERIALES Y MÉTODOS</b> .....	<b>11</b>
<b>Diseño de marcadores KASP</b> .....	<b>11</b>
<b>Genotipado con los marcadores KASP</b> .....	<b>12</b>
<b>Validación de marcadores KASP mediante GWAS</b> .....	<b>13</b>
<b>Validación del efecto del QTL en familias endogámicas heterogéneas</b> .....	<b>14</b>
<b>V. RESULTADOS Y DISCUSIÓN</b> .....	<b>16</b>
<b>Diseño de marcadores moleculares</b> .....	<b>16</b>

<b>Estudio de GWAS .....</b>	<b>18</b>
<b>Validación en familias endogámicas heterogéneas.....</b>	<b>21</b>
<b>VI. CONCLUSIONES.....</b>	<b>24</b>
<b>VII. LITERATURA CITADA .....</b>	<b>25</b>
<b>ANEXOS.....</b>	<b>32</b>

## LISTA DE CUADROS

Cuadro 1. Primer KASP.....	12
Cuadro 2. Colecciones de trigo utilizadas en la validación de marcadores KASP.....	13
Cuadro 3. Pedigree de las familias endogámicas heterogéneas .....	15
Cuadro 4. Prueba de X2 para comparar genotipado GBS y KASP .....	18
Cuadro 5. Resumen de resultados de GWAS con marcadores individuales y haplotipos para validar la asociación de los marcadores KASP con el peso de mil granos en cuatro colecciones de trigo.....	19
Cuadro 6. Resumen de resultados de GWAS con marcadores individuales y haplotipos para validar la asociación de los marcadores KASP con rendimiento en WAMI .....	20
Cuadro 7. Comparación de medias de rendimiento, peso, longitud, ancho y grosor de grano en diez HIF portadoras de los haplotipos CA y TG del QTL HAP-6A-13 evaluadas en los ambientes Batán y Obregón.....	22



## LISTA DE FIGURAS

Figura 1. Construcción de familias endogámicas heterogéneas (HIF)..	14
Figura 2. Gráficos de dispersión para ensayos KASP mostrando el agrupamiento de genotipos de trigo en los ejes X (FAM) e Y (HEX)..	17
Figura 3 Gráficos de Manhattan de los resultados de GWAS para peso de mil granos en el ensayo de rendimiento EYT2013-14	20

## I. INTRODUCCIÓN

Los caracteres peso, longitud, ancho y grosor del grano son caracteres que se heredan de una manera relativamente estable, y además exhiben valores de heredabilidad mayores que el rendimiento general (Dong et al., 2014; Kuchel et al., 2007). Dado que estos caracteres están regulados por múltiples genes, el aprovechamiento de la variación genética es un enfoque prometedor para mejorar el rendimiento del trigo (Chen et al., 2020; Liu et al., 2020).

La detección de loci de carácter cuantitativo (QTL - quantitative Trait Loci) amplía las oportunidades para desarrollar marcadores diagnóstico que puedan utilizarse en la selección asistida por marcadores moleculares y en la pirimidación de genes, permitiendo aumentar las ganancias genéticas en caracteres como peso de grano y rendimiento (Sehgal et al., 2019; Tyagi et al., 2014).

Sehgal et al. (2019) reportaron el QTL *HAP-6A-13* con efecto en peso de mil granos en trigo harinero, el QTL se detectó mediante un estudio de asociación del genoma completo (GWAS - genome-wide association study) basado en haplotipos. Los autores utilizaron 4,302 líneas de trigo que comprendieron cinco ensayos de rendimiento del programa de mejoramiento del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) y 8,443 marcadores moleculares del tipo genotipado por secuenciación (GBS - Genotyping by sequencing). El haplotipo favorable del QTL *HAP-6A-13* contribuyó con un incremento de hasta 2.60 gramos en muestras de mil granos. El QTL se localizó en el cromosoma 6A, en una región no codificante. El gen más próximo al QTL es TraesCS6A02G258900 y codifica para una proteína quinasa, se han reportado algunas proteínas quinasa como reguladoras del tamaño del grano en trigo y arroz (Duan et al., 2014; Miao et al., 2017).

Las líneas casi isogénicas (NIL – near isogenic lines) son un conjunto de líneas efectivas para estudiar los efectos fenotípicos atribuibles a un QTL ya que el fondo genético, caracteres morfológicos y fenológicos que comúnmente influyen en las evaluaciones fenotípicas de los caracteres cuantitativos, está estandarizado (Pumphrey et al., 2007).

Las familias endogámicas heterogéneas (HIF – Heterogeneous inbred family), son un tipo especial de NIL, se obtienen al cruzar dos progenitores distintos y realizar autopolinizaciones durante varias generaciones. Seguido se detectan líneas segregantes heterocigóticas para la región de un QTL a verificar, estas líneas identificadas, se autopolinizan una vez más, para obtener una segregación

mendeliana del QTL y finalmente se compara el fenotipo de las líneas segregadas (Kooke et al., 2012; Tuinstra et al., 1997).

## II. OBJETIVOS E HIPÓTESIS

### Objetivo general

Validar el QTL *HAP-6A-13* en ensayos de rendimiento y familias isogénicas de trigos harineros de primavera desarrollados por CIMMYT y valorar la conveniencia de ser integrado en programas de mejoramiento genético.

### Objetivos específicos

- Diseñar marcadores moleculares tipo KASP a partir de secuencias GBS.
- verificar la asociación de los marcadores con el peso de mil granos mediante GWAS en nuevas poblaciones de mejoramiento
- Construir familias isogénicas bajo la metodología HIF.
- Validar el efecto del QTL *HAP-6A-13* al comparar el peso de grano entre líneas isogénicas HIF.

### Hipótesis

El QTL *HAP-6A-13* tiene un efecto lo suficientemente consistente en peso de grano para ser integrado en programas de mejoramiento genético para la piramidación de genes.

## III. REVISIÓN DE LITERATURA

### Mejoramiento de caracteres cuantitativos

Los caracteres cuantitativos pueden ser, en principio, todos los rasgos que pueden medirse en un individuo, tanto si se trata de dimensiones, proporciones anatómicas o funciones fisiológicas de todo tipo. La distribución de frecuencias de los caracteres cuantitativos se aproxima en mayor o menor medida, a la distribución normal (Falconer y Mackay, 2006).

La variación cuantitativa puede ser explicada por el modelo infinitesimal, el cual establece que los caracteres cuantitativos son la suma de un componente genético y otro no genético (ambiental), además supone que la variación genética de los caracteres cuantitativos es causada por un infinito número de QTL con efectos alélicos infinitesimales (Barton et al., 2017; Mackay, 2001).

Shrimpton y Robertson (1988) proponen que los caracteres cuantitativos están bajo el control de unos pocos genes principales y respaldados por numerosos genes con efectos menores. Los loci principales son los responsables de la mayor parte de la variación en los caracteres.

Un QTL principal o QTL mayor tiene un efecto lo suficientemente grande y consistente, aunque estos QTL pueden ser significativos en unos programas, pueden no serlo en otros. Es más sencillo encontrar QTL principales para caracteres morfológicos, fenológicos, tolerancia a estrés biótico y abiótico, pero es probable que estén ausentes para un carácter altamente seleccionado como el rendimiento en germoplasma elite (Bernardo, 2020).

Los QTL menores tienen una influencia sutil sobre el fenotipo, que se enmascara fácilmente con la epistasia, las interacciones gen-ambiente, la baja asociación con el fenotipo debido a un tamaño pequeño de muestra y las interacciones complejas con otros QTL. Los QTL menores son importantes porque juntos pueden producir de manera aditiva o sinérgica efectos igualmente dramáticos sobre el fenotipo que los QTL principales (Yang et al., 2013).

### **Peso de grano: un carácter complejo**

El incremento del rendimiento de grano es uno de los objetivos más desafiantes en el mejoramiento de trigo debido a su compleja arquitectura genética, baja heredabilidad y porque está influenciado por interacciones ambientales en todas las etapas de crecimiento de la planta (Li et al., 2019; Simmonds et al., 2014).

Dada la complejidad del rendimiento, este se puede disectar en componentes, los cuales son: número de espigas por unidad de superficie, número de granos por espiga y peso de mil granos. Estos caracteres tienen mayor heredabilidad que el rendimiento y son relativamente fáciles de seleccionar en los programas de mejoramiento (Li et al., 2019).

El peso de grano es una característica importante en el mejoramiento genético de trigo ya que es uno de los componentes del rendimiento más estables, tiene una alta heredabilidad y se correlaciona positivamente con el rendimiento en harina (Cooper et al., 2012; Guo et al., 2019; Li et al., 2019; Philipp et al., 2018).

El peso del grano se determina principalmente a través del llenado del grano entre la antesis y la madurez, durante el cual ocurren simultáneamente tres procesos fisiológicos: acumulación de materia seca del grano, acumulación de agua en el grano y posterior desecación y expansión

morfológica del grano. La acumulación de materia seca del grano es un proceso de depósito de almidón (~60-70 % del peso del grano maduro), proteínas (8-15 %) y otros nutrientes (por ejemplo, minerales, vitaminas y fibras) (Brinton y Uauy, 2019; Xie et al., 2015).

Los asimilados para el llenado de granos se originan principalmente a partir de la fotosíntesis del momento y por removilización de las reservas solubles acumuladas en los órganos vegetativos antes de la antesis. Los órganos senescentes también pueden suministrar algunos asimilados al transportar los nutrientes a los granos en desarrollo en la etapa tardía del crecimiento de las plantas (Distelfeld et al., 2014).

Varios estudios han desafiado la suposición general de que el peso del grano del trigo se determina exclusivamente durante el llenado del grano al mostrar que el período anterior a la antesis es una fase clave para establecer el potencial de peso del grano (Calderini y Reynolds, 2000; Hasan et al., 2011).

Otro dilema en el mejoramiento genético de trigo con mayor potencial de rendimiento es la asociación negativa entre peso de grano y número de granos, por lo que aumentos en el tamaño del grano puede tener poco impacto en el rendimiento (Calderini et al., 2021).

Si bien la mayoría de las fases del desarrollo del grano se han caracterizado fenotípicamente extensamente, la base genética de cómo se controlan estos procesos y cómo influyen en el peso final del grano no se comprenden bien en el trigo (Brinton y Uauy, 2019).

Por la importancia que tiene el peso de grano, muchos esfuerzos se han dirigido al mapeo QTL y GWAS para identificar regiones genómicas que controlan el peso de grano y que puedan ser objetivo de la selección asistida por marcadores moleculares, algunos de los genes con efectos importantes en peso de grano son: *TaGW2*, *TaGW2-6B*, *TaGS5*, *TaGW8*, *TaGS-D1*, *TaSus2*, *TaCKX-D1*, *TaTGW*, *TaPARG*, *TaTPP-6AL1*, *TaFlo2-A1*, *TaTPP-6AL1*, *TaGASR7*, *TaCwi-A1*, *BAL-A1*, *TaGS1a*, *TaSPL20*, *TaSAP1 -A1* (Rasheed y Xia, 2019).

El gen *TaGW2-6A* está involucrado en el desarrollo del grano, afectando principalmente ancho de grano y peso de mil granos. En trigo chinos se han estudiado dos haplotipos de este gen: *Hap-6A-A* (favorable) y *Hap-6A-G* (desfavorable). Un análisis de asociación en genotipos chinos antiguos no mostró diferencias significativas en peso de mil granos entre las accesiones *Hap-6A-G* y *Hap-6A-A*, sin embargo, hubo diferencias significativas para para ancho y peso de grano en un conjunto

de líneas modernas. Las diferencias de peso de grano entre *Hap-6A-A* y *Hap-6A-G* en las líneas modernas fueron de 3.2 y 3.0 g en 2002 y 2006 respectivamente. Las diferencias en grosor de grano fueron significativas en 2002 pero no lo fueron en 2006 (Su et al., 2011). Estudios para el mismo gen realizados en México no muestra una asociación consistente y significativa con el peso de grano en diferentes ambientes. Contrario a las líneas chinas, se encontró *Hap-6A-G* como el haplotipo favorable, aunque esto solo fue significativo en dos de diez ambientes analizados (Sukumaran et al., 2018).

El gen *TaGASR7-A1* fue reportado como controlador de la longitud y peso de grano en trigo harinero chino, este gen tiene dos haplotipos: H1c (favorable) y H1g (desfavorable) (Ling et al., 2013). Al cultivar 94 líneas de trigo chinas en nueve ambientes que incluían dos ubicaciones, dos ciclos y tres tratamientos (cultivo en temporal, fertilización nitrogenada reducida y fertilización reducida en fosforo), se encontró que H1c fue significativamente superior que H1g en longitud de grano, peso de mil granos y rendimiento en siete, seis y cuatro de los nueve ambientes, respectivamente (Dong et al., 2014).

Un QTL asociado a la enzima trehalosa 6-fosfato fosfatasa fue reportado por Zhang et al. (2017), el QTL fue nombrado como *TaTPP-6AL1* y se reportó como controlador del peso de grano. Fue validado mediante asociación utilizando líneas chinas cultivadas en cinco ambientes que incluían tres ubicaciones. El peso de mil granos de las líneas con el alelo *TaTPP-6AL1a* fue significativamente más alto que los cultivares con el alelo *TaTPP-6AL1b* en cuatro de cinco ambientes. El incremento promedio del peso de mil granos en los cinco ambientes fue de 1.7 g.

Después de la identificación de QTL que afectan el peso de grano, es necesaria una validación funcional. Esto podría lograrse utilizando múltiples mutantes independientes o líneas genéticamente modificadas (con la modificación adecuada para cada caso) (Brinton y Uauy, 2019). Con estas herramientas se puede demostrar de manera concluyente que los genes candidatos y los haplotipos favorables propuestos son directamente responsables de controlar el peso del grano, en lugar de estar asociados con un mayor peso del grano debido al desequilibrio de ligamiento con un gen cercano no relacionado. Esto es especialmente relevante en especies con recombinación desigual entre los cromosomas, como el trigo (Akhunov et al., 2003).

### **Caracteres morfométricos del grano**

El peso de grano puede ser diseccionado en caracteres morfométricos del grano (longitud, ancho y grosor del grano). Los caracteres morfométricos exhiben valores de heredabilidad mayores que el rendimiento general (Dong et al., 2014; Kuchel et al., 2007), por lo tanto, el aprovechamiento de la variación genética de estos caracteres es un enfoque prometedor para mejorar el rendimiento del trigo (Chen et al., 2020).

En trigos hexaploides sintéticos se han reportado correlaciones positivas significativas entre el peso del grano y caracteres morfométricos, como la longitud del grano ( $r=0.43$ ), el ancho y el grosor ( $r=0.64$ ) (Rasheed et al., 2014). En trigo harinero se ha encontrado que el ancho y el grosor del grano están bajo un control genético similar y que tiene una correlación positiva con el peso y el rendimiento de grano. La longitud del grano también se asocia con el rendimiento del grano, pero su control genético parece diferir del ancho del grano (Maphosa et al., 2014).

El gen *TaGS5-3A* se ha reportado como controlador del tamaño y el peso de grano en trigo, un estudio realizado por Ma et al., (2016) reportó los alelos: *TaGS5-3A-T* (favorable) y *TaGS5-3A-G* (desfavorable), mediante un estudio de asociación el alelo favorable confirió mayor longitud y peso de grano en una colección de líneas chinas antiguas, los incrementos en peso de grano fueron de 2.29, 7.81 y 6.65 g para los años 2002, 2005 y 2010 respectivamente, Los incrementos en longitud de grano fueron de 0.05 mm para los tres años de cultivo. En líneas chinas modernas el efecto no fue consistente.

Yang et al., (2012) reportó una inserción de 1 pb en el octavo exón de *TaGW2-6A*, que afectó el ancho y peso de grano. El gen mutado codificó una proteína incompleta con pérdida de función. El análisis de asociación indicó que las líneas con el alelo mutado *TaGW2* aumentó significativamente el ancho del grano y el peso de mil granos, posiblemente por un aumento en el número de células.

Los genes *TaSu1* y *TaSus2* tienen participan en la síntesis de almidón del endospermo por lo que tiene una influencia directa en el peso de grano (Hou et al., 2014), el gen *TaSus2-2B* tiene dos haplotipos: Hap-H y Hap-L, en un estudio utilizando líneas de trigo chinas el Hap-H tuvo un efecto consistente en peso de grano (Jiang et al., 2011), pero en un estudio asociación en trigos coreanos Hap-L fue el haplotipo favorable, mostrando superioridad en peso, largo, ancho y grosor de grano. Las diferencias promedio en cuatro años entre los dos haplotipos fue de 6.2 g, 0.33 mm, 0.13 mm

y 0.14 mm en peso de mil granos, largo, ancho y grosor de grano respectivamente (Kim et al., 2020).

### **Estudio de asociación del genoma completo**

GWAS es un método de mapeo de QTL que ayuda a comprender la variación de caracteres complejos y enfermedades, relacionando un gran número de marcadores (normalmente SNP) con los fenotipos observados y encontrando los SNPs que se asocian significativamente con la variación en el fenotipo (Bush y Moore, 2012).

GWAS ha tenido gran aplicación en el mejoramiento genético y se ha convertido en el enfoque *de facto* para identificar regiones candidatas asociadas con enfermedades, caracteres agronómicos, fisiológicos y de calidad. En muchos casos, GWAS ha identificado variantes que están indudablemente asociadas con variaciones en los caracteres y, en algunos casos, los estudios han culminado con el descubrimiento de las variantes causales y su mecanismo de acción. En otros casos, estos estudios han evidenciado que algunos caracteres son altamente poligénicos y, en lugar de identificar algunas regiones genómicas responsables, se encontró una dispersión de efectos en todo el genoma. Este último caso es un claro ejemplo de que algunos procesos biológicos son más complejos de lo quisiéramos (Gondro et al., 2013).

GWAS es una herramienta que puede mejorar la eficiencia, la precisión y el ritmo de desarrollo de nuevas líneas, por esta razón muchos investigadores han centrado sus esfuerzos en reportar nuevos QTL (Vanderzande et al., 2018).

### **Prueba de asociación**

Cuando se ha seleccionado un fenotipo para una población de estudio y los genotipos se recopilan utilizando técnicas sólidas, se puede comenzar el análisis estadístico de los datos genéticos. El análisis de facto de los datos de asociación de todo el genoma es una serie de pruebas estadísticas de un solo locus, examinando cada SNP de forma independiente para la asociación con el fenotipo (Bush y Moore, 2012).

Los caracteres cuantitativos generalmente se analizan utilizando enfoques de modelo lineal generalizado (GLM), más comúnmente el Análisis de varianza (ANOVA), que es similar a la regresión lineal con una variable predictora categórica, en este caso clases de genotipos. La hipótesis nula de un ANOVA que usa un solo SNP es que no hay diferencia entre las medias de



los caracteres de ningún grupo de genotipos (Bush y Moore, 2012). Entre las ventajas de GLM se encuentra, que proporciona cierto control sobre la estructura de la población, eficiencia computacional y no se ha limitado a los caracteres que se distribuyen normalmente (Mohammadi et al., 2020).

Sin embargo, en una población con una estructura compleja, los métodos GLM muestran niveles de significancia aumentados que conducen a un control inadecuado de los falsos positivos y pérdida de poder debido a la relación entre los individuos de la población. Como alternativa a estos problemas, se propusieron modelos lineales mixtos (MLM) los cuales proporcionan más poder, flexibilidad y un control más estricto de la estructura de la población (Mohammadi et al., 2020). Si bien es cierto que un modelo MLM es capaz de controlar los falsos positivos al incluir efectos fijos y efectos aleatorios, que incorporan la estructura de la población y el parentesco entre los individuos se corre el riesgo de comprometer los verdaderamente positivos (Liu et al, 2016).

Gracias al desarrollo de procedimientos de estimación penalizados y bayesianos, así como a los avances en el campo de las regresiones no paramétricas, se han propuesto y utilizado varios métodos como alternativa a la búsqueda de QLT (de los Campos et al., 2013; Xu, 2003). Una alternativa al análisis de marcador único es la regresión del genoma completo, comúnmente utilizado en las predicciones genómicas. Xu (2003) propuso detectar QTL utilizando un modelo lineal en el que cada marcador se trataba como un efecto aleatorio individual. Una ventaja extra de los métodos de regresión del genoma completo es que condicionan la evaluación estadística de cada SNP a todos los demás SNP, eliminando la necesidad de correcciones de pruebas múltiples (Mohammadi et al., 2020).

Otra alternativa al análisis de marcador único, son los modelos mixtos de múltiples locus como: MLMM (del inglés - Multi-Locus Mixed-Model) y FarmCPU (del inglés - Fixed and random model Circulating Probability Unification). Estos métodos prueban múltiples marcadores simultáneamente, son computacionalmente eficiente y tienen mejor control de falsos positivos en comparación con los análisis de marcador único (Segura et al., 2012; Liu et al. 2016).

Dentro de las opciones no paramétricas para GWAS se encuentra Random Forest, esta metodología es independiente de la distribución del carácter, no es necesario asumir la normalidad y es capaz de manejar rasgos categóricos (Mohammadi et al., 2020). En lugar de estimar el efecto de marcadores individuales sobre el fenotipo, Random Forest se basa en árboles de decisión,

pudiendo detectar asociaciones de marcadores que se basan en efectos aditivos, dominantes y epistáticos, y es capaz de capturar haplotipos de alta complejidad (Botta et al., 2014).

Existen muchas opciones para realizar las pruebas de asociación. Es tarea del investigador encontrar las herramientas adecuadas para lograr mejores resultados.

### **Estructura de la población**

Los genotipos analizados mediante GWAS suelen ser genotipos con relaciones genéticas desiguales entre grupos de individuos, a esto se le conoce como estructura de la población. En general, la estructura de la población puede surgir debido a diferencias en los orígenes geográficos, adaptaciones locales o el historial de mejoramiento de las líneas en el panel (Tibbs et al., 2021).

El mal análisis de la estructura de la población conduce al descubrimiento de falsos positivos, por lo tanto, es indispensable comprender la estructura de la población, desarrollar métodos para inferir con precisión la estructura y realizar los análisis de asociación con modelos adecuados. Estas consideraciones son fundamentales para reducir o eliminar asociaciones espúreas o falsos positivos (Sukumaran y Yu, 2014).

Se han propuesto varios métodos para abordar la estructura de la población, incluido el control genómico, la asociación estructurada, el análisis de componentes principales y los modelos lineales mixtos (Sukumaran y Yu, 2014).

### **Validación de QTL obtenido mediante GWAS para su uso en mejoramiento genético**

Después de identificar nuevas regiones genómicas que afectan un carácter mediante GWAS, es importante verificar que no se trata de asociaciones espurias, ya que existe el riesgo que las estimaciones significativas hayan sido sobreestimadas (Henshall, 2013).

Antes de describir los métodos de validación es necesario aclarar que un SNP identificado como significativo en GWAS, puede estar en asociación directa o indirecta. La asociación directa se da cuando se genotipó el SNP que influye en un sistema biológico y que finalmente conduce al fenotipo y también es llamado SNP funcional. En la asociación indirecta no se genotipó el SNP funcional, sino que se genotipó un SNP-etiqueta que se encuentra con alto grado de desequilibrio de ligamiento con la mutación que causa la variación fenotípica. Debido a estas dos posibilidades, un SNP con asociación significativa mediante GWAS, no debe asumirse como la variante causal

y se requieren estudios adicionales para mapear la ubicación precisa del SNP influyente (Schaid et al., 2018).

A continuación, se describen propuestas de validación de QTL identificados mediante GWAS, y se excluyendo los enfoques transgénicos, ARNi y CRISPR que son herramientas más apropiadas para la validación de genes causales.

### **Replicación estadística**

La replicación ayuda a garantizar que una asociación genotipo-fenotipo observada en GWAS represente una asociación creíble y no sea un hallazgo casual además de descartar asociaciones debidas a sesgos (Kraft et al., 2009). La estrategia general para un estudio de replicación es repetir el diseño del GWAS lo más cerca posible, pero examinar solo los efectos genéticos específicos que se encuentran significativos en el GWAS. Los efectos que son consistentes en los dos estudios se pueden etiquetar como efectos replicados (Bush y Moore, 2012).

A menudo, los efectos identificados en un GWAS inicial sufren “la maldición del ganador”, donde el efecto detectado es probablemente más importante en la muestra de GWAS que en la población general. Por lo tanto, se sugiere hacer la replicación en un conjunto de datos independiente extraído de la misma población que el GWAS, en un intento de confirmar el efecto en la población objetivo de GWAS. El número de individuos utilizados para la replicación debe ser, idealmente, más grande que el conjunto de individuos empleados en GWAS original, para tener en cuenta la sobreestimación del efecto (Bush y Moore, 2012).

En la replicación deben usarse criterios fenotípicos idénticos que en GWAS, porque la replicación de un resultado de GWAS debe considerarse como la replicación de un modelo estadístico específico: un SNP dado predice un efecto de fenotipo específico. Cambios ligeros al fenotipar puede enturbiar la interpretación del resultado final (Bush y Moore, 2012).

### **Validación en líneas casi isogénicas**

#### **líneas endogámicas retrocruzadas**

El desarrollo de líneas casi isogénicas (NIL) a través del retrocruzamiento sucesivo es bastante sencillo. Por lo general el padre donante transfiere una región genómica pequeña a un padre receptor (recurrente) (Kooke et al., 2012).

Los dos progenitores (donador y recurrente) se cruzan entre ellos, después la F1 se cruza con el progenitor recurrente, este último paso se repite las veces que sean necesarias hasta que se recupere la mayor parte del genotipo recurrente. La selección asistida con marcadores moleculares es una herramienta que facilita la obtención de líneas retrocruzadas, en especial para mantener alelos recesivos (Harlan y Pope, 1922; Kooke et al., 2012). La validación del QTL se puede hacer al comparar el fenotipo del progenitor recurrente contra las líneas retrocruzadas.

### **Familias endogámicas heterogéneas**

Las familias endogámicas heterogéneas (HIF - heterogeneous inbred families) son un tipo especial de NIL. Se obtienen al cruzar dos padres distintos y realizar autopolinizaciones durante varias generaciones. En este punto se han creado genotipos homocigotos casi completos, exceptos por algunas regiones pequeñas (<5% del tamaño del genoma) (Kooke et al., 2012).

En generaciones F<sub>6</sub> o posteriores es posible detectar líneas heterocigóticas para la región del QTL a verificar, estas líneas identificadas se autopolinizan una vez más para que la región de interés tenga una segregación mendeliana. En las líneas segregantes se puede comparar el efecto de cada región segregada (Tuinstra et al., 1997).

## **IV. MATERIALES Y MÉTODOS**

### **Diseño de marcadores KASP**

Se diseñaron primers de tipo KASP para los polimorfismos C/T y A/G (Cuadro 1) a partir de las secuencias GBS de los marcadores S6A\_481437887 y S6A\_481437894 que formaban el haplotipo del QTL *HAP-6A-13* reportado por Sehgal et al. (2019). El diseño de los KASP se realizó con PolyMarker (Ramirez et al., 2015).

Secuencia de los marcadores S6A\_481437887 y S6A\_481437894:  
CGGGATGCTCTGCTTCTCTCTGTTCTCAGCCCAACATGT[C/T]TACACA[A/G]CCCACC  
TGCA

### Cuadro 1. Primer KASP

SNP	Alelo favorable	Alelo HEX	Alelo FAM	Alelo común
C/T	T	ctgttctcagccaacatgctc	ctgttctcagccaacatggt	ggttaggtgggctgcagg
A/G	G	tctcagccaacatgtctacacaa	tctcagccaacatgtctacacag	gaggggttaggtgggctgc

### Genotipado con los marcadores KASP

Se genotiparon tres colecciones de trigo harinero que forman parte de ensayos internacionales de rendimiento (EYT - Elite Yield Trial) de CIMMYT. Los ensayos EYT2013-14, EYT2014-15 y EYT2015-16 estaban formados por 983, 942 y 829 genotipos respectivamente.

La extracción de ADN de las colecciones; EYT2013-14, EYT2014-15 y EYT2015-16, se realizó a partir de hojas jóvenes siguiendo el protocolo de extracción de ADN a pequeña escala en placas de 96 pocillos descrito por Dreisigacker et al. (2016). El ADN se cuantificó en NanoDrop 8000 (Thermo Scientific, Waltham, Massachusetts, Estados Unidos) y fue diluido a una concentración de 50 ng/μl.

El genotipado se realizó mediante PCR, con los primers del Cuadro 1. Los componentes para cada reacción fueron: 3 μl de ADN, 2 primers alelo específicos, 1 primer inverso común y master mix de genotipado PACE (3cr Bioscience, Essex, Reino Unido).

La amplificación por PCR se realizó en un termociclador GeneAmp PCR System 97 00 (Applied Biosystems, Alemania). Los ciclos de PCR se realizaron con una temperatura inicial de desnaturalización a 95 °C durante 15 min, seguido por diez ciclos con rampa de temperatura (95 °C durante 20 s; 65 °C durante 25 s inicialmente y disminuyendo en -1 °C en cada ciclo posterior), seguido de 30 ciclos de alineamiento (95 °C durante 10 s y 57 °C durante 60 s). No fue necesario un paso de extensión porque los amplicones esperados tienen 50 pares de bases de longitud.

Las placas de PCR fueron leídas en PHERAstar Plus (BMG LABTECH, Alemania), las lecturas de genotipado se analizaron en el software KlusterCaller TM (LGC Genomics, Reino Unido).

El genotipado KASP de los EYT se comparó con los resultados previos del genotipado reportado por Sehgal et al. (2019), y se realizó una prueba de  $X^2$  (Chi-cuadrada).

### **Validación de marcadores KASP mediante GWAS**

Se probó la asociación de los marcadores KASP diseñados con el peso de mil granos en cuatro colecciones de trigo harinero, mediante GWAS. Se incluyeron los EYT descritos anteriormente (EYT2013-14, EYT2014-15 y EYT2015-16) y el panel WAMI (wheat association mapping initiative). La asociación de los marcadores con el rendimiento se realizó con el panel WAMI, este es un set diverso de trigo harinero de primavera que consta de 287 líneas avanzadas liberadas por CIMMYT, tiene poca variación en fenología y altura de la planta (Lopes et al., 2015).

Los paneles de validación EYT2013-14, EYT2014-15 y EYT2015-16 anteriormente se genotiparon usando el método de genotipado por secuenciación (GBS - Genotyping by sequencing) realizado mediante un HiSeq2000 de Illumina, con lecturas de 1 x 100 pb, datos proporcionado por Sehgal et al. (2019). El panel WAMI fue genotipado con Illumina iSelect 90K SNP Assay, los datos fueron proporcionados por Sukumaran et al. (2015). A los resultados de secuenciación se adicionó el genotipado KASP con los dos marcadores de Cuadro 1. En el Cuadro 2 se muestra el total de individuos y marcadores utilizados en los GWAS.

### **Cuadro 2. Colecciones de trigo utilizadas en la validación de marcadores KASP**

Colecciones de trigo	Número de individuos	Número de marcadores
EYT2013-14	983	13,770
EYT2014-15	942	13,768
EYT2015-16	829	13,768
WAMI	287	15,740

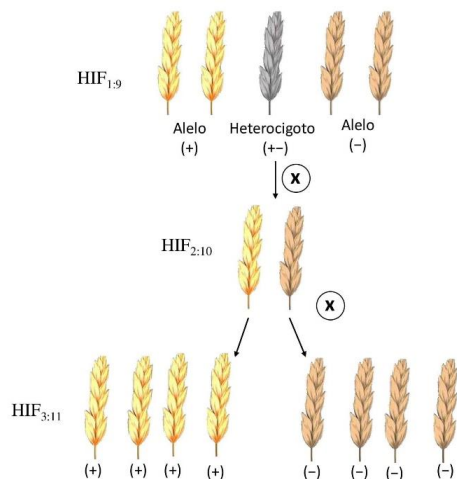
El fenotipado de los EYT y WAMI se realizó en el Campo Experimental Norman E. Borlaug (CENEB), Obregón, Sonora (noreste de México). El cultivo se realizó en condiciones de manejo estándar procurando riego, fertilización y manejo de plagas, enfermedades y malezas de acuerdo con lo requerido. EYT fueron sembrados cada uno en sus respectivos años (2013, 2014 y 2015) en un diseño experimental  $\alpha$ -lattice con tres repeticiones. WAMI fue sembrado en los años 2010,

2011 y 2013 en un diseño experimental  $\alpha$ -lattice con dos repeticiones. El peso de grano y rendimiento se midió como se describe en Pask et al. (2013).

El mapeo de asociación se realizó en el software TASSEL 5.0 (Bradbury et al., 2007) utilizando el modelo lineal generalizado (GLM) y análisis de componentes principales para la estructura de la población. Se excluyeron los marcadores con una frecuencia de alelos menores a 0.05.

### Validación del efecto del QTL en familias endogámicas heterogéneas

Se construyeron diez familias HIF (Cuadro 3) siguiendo el método de Tuinstra et al. (1997) (Figura 1). En la primera siembra se utilizaron semillas de generación F8, de diez cruzas de trigo harinero de primavera, desarrollado en CIMMYT y se identificaron como generación HIF<sub>1:9</sub>. La siembra se realizó en macetas bajo condiciones de invernadero y todas las plantas fueron genotipadas con los marcadores KASP diseñados, para identificar plantas segregantes en el QTL *HAP-6A-13*. Se seleccionó una planta heterocigota de cada craza y se autopolinizó. La siguiente generación (HIF<sub>2:10</sub>) fue sembrada en condiciones similares a la siembra anterior, y dentro de cada familia se seleccionaron plantas homocigotas para el QTL *HAP-6A-13*. En las dos generaciones posteriores se realizaron las evaluaciones fenotípicas en campo, HIF<sub>3:11</sub> fue evaluada en Batán, Estado de México (centro de México) y HIF<sub>4:12</sub> fue evaluada en Obregón, Sonora (noreste de México).



**Figura 1.** Construcción de familias endogámicas heterogéneas (HIF). Se realizó una craza y se continuó con autopolinizaciones, en la generación F<sub>1:9</sub> se detectó una planta segregante heterocigótica para la región del QTL *HAP-6A-13*, esta planta identificada, se autopolinizó una vez más y la región heterocigótica segregó de manera mendeliana, permitiendo la comparación del carácter en los genotipos segregados en un contexto isogénico.

**Cuadro 3. Pedigree de las familias endogámicas heterogéneas**

HIF	Haplotipo	Pedigree
HIF1	CA (-)	BAJ #1/3/TRCH/SRTU//KACHU
	TG (+)	BAJ #1/3/TRCH/SRTU//KACHU
HIF2	CA (-)	SAUAL/3/KAUZ/PASTOR//PBW343/4/TRCH/SRTU//KACHU/5/SAUAL//PI 610750/ATTILA/3/SAUAL
	TG (+)	SAUAL/3/KAUZ/PASTOR//PBW343/4/TRCH/SRTU//KACHU/5/SAUAL//PI 610750/ATTILA/3/SAUAL
HIF3	CA (-)	CROC_1/AE.SQUARROSA (205)//BORL95/3/PRL/SARA//TSI/VEE#5/4/FRET2/6/MTRWA92.161/PRINIA/ 5/SERI*3//RL6010/4*YR/3/PASTOR/4/BAV92
	TG (+)	CROC_1/AE.SQUARROSA (205)//BORL95/3/PRL/SARA//TSI/VEE#5/4/FRET2/6/MTRWA92.161/PRINIA/ 5/SERI*3//RL6010/4*YR/3/PASTOR/4/BAV92
HIF4	CA (-)	PRL/2*PASTOR//SUNSTATE/3/MUNAL #1/4/OTUS//PRL/2*PASTOR
	TG (+)	PRL/2*PASTOR//SUNSTATE/3/MUNAL #1/4/OTUS//PRL/2*PASTOR
HIF5	CA (-)	HEILO//MILAN/MUNIA/8/2*NG8201/KAUZ/4/SHA7//PRL/VEE#6/3/FASAN/5 /MILAN/KAUZ/6/ACHYUTA/7/PBW343*2/KUKUNA
	TG (+)	HEILO//MILAN/MUNIA/8/2*NG8201/KAUZ/4/SHA7//PRL/VEE#6/3/FASAN/5 /MILAN/KAUZ/6/ACHYUTA/7/PBW343*2/KUKUNA
HIF6	CA (-)	SOKOLL//W15.92/WBLL1/3/BAVIS
	TG (+)	SOKOLL//W15.92/WBLL1/3/BAVIS
HIF7	CA (-)	SAUAL/3/KAUZ/PASTOR//PBW343/4/TRCH/SRTU//KACHU
	TG (+)	SAUAL/3/KAUZ/PASTOR//PBW343/4/TRCH/SRTU//KACHU
HIF8	CA (-)	TRCH/SRTU//KACHU*2/3/KINGBIRD #1
	TG (+)	TRCH/SRTU//KACHU*2/3/KINGBIRD #1
HIF9	CA (-)	SUP152/KACHU/3/TRCH/SRTU//KACHU
	TG (+)	SUP152/KACHU/3/TRCH/SRTU//KACHU
HIF10	CA (-)	BL3264*2/PAURAQ
	TG (+)	BL3264*2/PAURAQ

La generación HIF<sub>3:11</sub> fue sembrada en condiciones de campo, para hacer las evaluaciones fenotípicas. La siembra se estableció el 3 de junio de 2020 en la Estación Experimental El Batán (CIMMYT), Texcoco, Estado de México, bajo un diseño experimental de bloques completos al azar con dos repeticiones. La unidad experimental fue un surco de 2 m sembrado a doble hilera. El manejo del cultivo se realizó de manera estándar procurando los riegos, fertilización, manejo de plagas, enfermedades y malezas de acuerdo con los requerimientos del cultivo. La cosecha se realizó de forma manual y la trilla mecánicamente.



La generación HIF<sub>4:12</sub> fue sembrada en el campo Experimental Norman E. Borlaug (CENEB), Obregón, Sonora. La siembra se realizó el 17 de diciembre de 2020, la unidad experimental fue un surco de 1.5 m sembrados a doble hilera. El diseño experimental y manejo del cultivo fue similar a la generación HIF<sub>3:11</sub>.

El grano cosechado fue secado en un horno con flujo de aire caliente (Thelco Precision Scientific, Estados Unidos) a una temperatura entre 37 y 40 °C hasta lograr una humedad del 11 %. La humedad del grano se midió mediante espectroscopia de infrarrojo cercano (DA 7200 NIR, Perten Instruments, Suecia). El grano cosechado de cada unidad experimental fue pesado en una báscula digital (Ohaus, Estados Unidos) para obtener el rendimiento. El peso de mil granos, longitud, ancho y grosor de grano, se midieron con el equipo SeedCount SC5000 (Next Instruments, Australia).

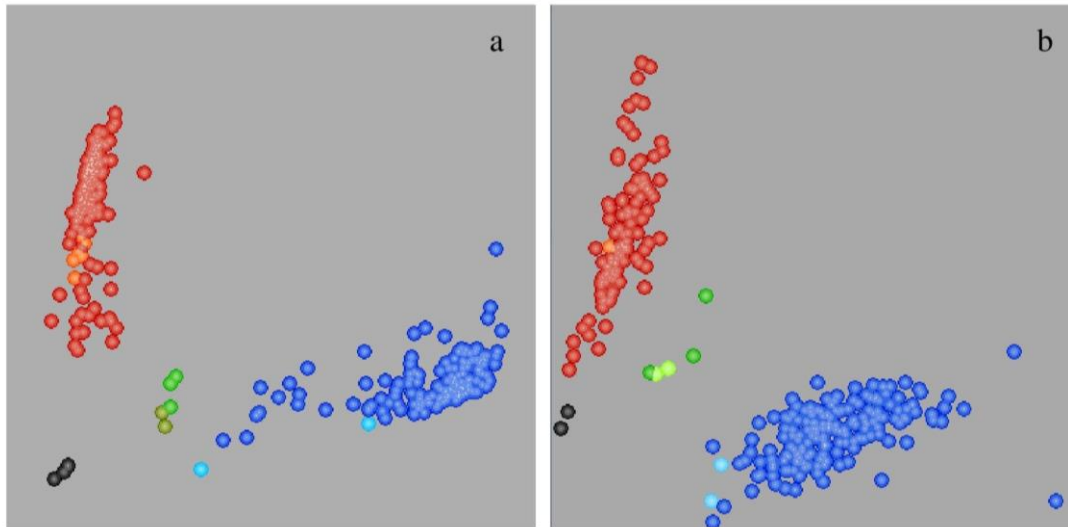
Para estudiar simultáneamente el efecto de las familias y los haplotipos sobre las variables fenotípicas: rendimiento, peso de mil granos, longitud, ancho y grosor de grano, se realizó un ANOVA de dos vías utilizando las librerías lme4 y lmerTest del lenguaje estadístico RStudio v4.0.3. La prueba Tukey ( $P \leq 0.05$ ) nos permitió discernir si las diferencias de los valores promedios de los haplotipos en cada familia fueron significativamente distintos.

## V. RESULTADOS Y DISCUSIÓN

### Diseño de marcadores moleculares

Se desarrollaron y verificaron dos marcadores KASP (Cuadro 1), a partir de secuencias GBS obtenidas en Sehgal et al. (2019). Los ensayos KASP pudieron diferenciar genotipos en grupos separados como se esperaba (Figura 2). Al comparar el genotipado GBS contra KASP, se obtuvo una coincidencia del 96 % para el SNP C/T y 95 % para el SNP A/G.

Las diferencias en el genotipado correspondieron principalmente a los resultados reportados como heterocigotos. La baja coincidencia en resultados heterocigotos puede deberse a limitantes de la tecnología GBS como la poca profundidad de secuenciación y la secuenciación multiplexada, dificultades similares en el genotipado de individuos heterocigotos mediante GBS han sido reportadas por Swarts et al. (2014), otra razón puede ser la heterocigosidad residual dentro y entre los lotes de semillas la cual no es tan infrecuente (Dreisigacker et al., 2020). En cuanto a datos faltantes el genotipado KASP tuvo un porcentaje de datos perdidos del 6 % mientras que para GBS fue del 59 %.



**Figura 2.** Gráficos de dispersión para ensayos KASP mostrando el agrupamiento de genotipos de trigo en los ejes X (FAM) e Y (HEX). **a)** Ensayo KASP para SNP C/T, **b)** Ensayo KASP para SNP A/G. Puntos de color rojo corresponde a genotipos con el alelo HEX, puntos en color azul corresponde a genotipos con el alelo FAM, puntos de color verde corresponde a genotipos heterocigotos y puntos de color negro corresponden a controles sin ADN.

Se realizó una prueba de  $X^2$  (Chi-cuadrado) ( $GL=2$ ,  $P \leq 0.05$ ) para comparar los resultados GBS contra cada marcador KASP (Cuadro 4). Dado que  $X^2$  calculada es menor que  $X^2$  de tablas no se rechaza  $H_0$  y se concluye que, los valores observados se ajustan a los esperados. Los dos métodos de genotipado, tienen resultados similares para los polimorfismos C/T y A/G, esto prueba que los marcadores KASP amplifican el mismo locus que GBS.

El buen desempeño de los ensayos KASP los convierte en herramientas útiles en la selección asistida y al ser una plataforma de genotipado de alto rendimiento los ensayos KASP pueden utilizarse para analizar un gran conjunto de accesiones de trigo.

Los polimorfismos C/T y A/G generaron los haplotipos TG, CA y CG que se encontraron en frecuencias de 30.3, 62.6 y 6.9 % respectivamente en el panel WAMI y frecuencias de 47.8, 32.4 y 19.7 % respectivamente en los EYT analizados. Considerando que WAMI está formado por líneas liberadas desde hace 35 años y los EYT contienen líneas más recientes, se puede demostrar una selección indirecta para el haplotipo favorable TG.

**Cuadro 4. Prueba de X<sup>2</sup> para comparar genotipado GBS y KASP**

	Valores observados		Valores calculados		X <sup>2</sup> Calculada	X <sup>2</sup> Tablas
	GBS	KASP	GBS	KASP		
SNP C/T					4.91	5.99
C	547	558	552.5	552.5		
T	433	434	433.5	433.5		
Y	21	9	15	15		
SNP A/G					0.53	5.99
A	379	369	374	374		
G	604	616	610	610		
R	8	6	7	7		

Y representa el genotipo heterocigoto CT; R representa el genotipo heterocigoto AG, de acuerdo con la nomenclatura IUPAC.

### Estudio de GWAS

En el mapeo de asociación, los dos marcadores KASP muestran asociación significativa con el peso de grano en las cuatro colecciones de trigo analizadas (Cuadro 5). El marcador KASP A/G tuvo valores superiores que el marcador KASP C/T, sin embargo, al analizar los dos marcadores como un haplotipo, se obtuvieron asociaciones más significativas que los análisis de marcador individual, esto fue así en seis de los siete análisis. La variación fenotípica explicada por el marcador (R<sup>2</sup>) osciló entre 4.6 y 5.9 % en las cuatro colecciones de trigo analizadas.

Debido al bajo R<sup>2</sup> del marcador (menos del 10 %) se considera como un QTL menor para peso de grano y dado que el QTL tuvo significancia estadística en las cuatro colecciones de trigo analizadas se sugiere que el QTL *HAP-6A-13* es estable, por lo tanto, ofrece el potencial de ser usado para la piramidación de alelos favorables a través de la selección asistida por marcadores (Zhang et al., 2016).

Estudios anteriores de GWAS han reportado QTL asociados a peso de mil granos con valores de R<sup>2</sup> del marcador iguales o menores al 10 %. Sukumaran et al. (2015) reportaron dos QTL asociados a peso de mil granos en los cromosomas 5A y 6A con R<sup>2</sup> de 9 y 8 % respectivamente. Jamil et al. (2019) reportaron un QTL en el cromosoma 6A el cual se mantuvo estable en condiciones de cultivo estándar y en estrés por calor, teniendo valores de R<sup>2</sup> de 10 %.

**Cuadro 5. Resumen de resultados de GWAS con marcadores individuales y haplotipos para validar la asociación de los marcadores KASP con el peso de mil granos en cuatro colecciones de trigo**

	<b>Colección de trigo</b>	<b>P-valor</b>	<b>Marcador_R<sup>2</sup></b>
<b>Marcador KASP C/T</b>	EYT2013-14	2.70E-08	0.0367
	EYT2014-15	5.91E-04	0.01512
	EYT2015-16	0.0055	0.00927
	WAMI 2010	1.92E-04	0.03413
	WAMI 2011	6.94E-04	0.03188
	WAMI 2013	0.00929	0.02013
<b>Marcador KASP A/G</b>	EYT2013-14	7.64E-09	0.03873
	EYT2014-15	3.95E-09	0.03506
	EYT2015-16	1.07E-08	0.03923
	WAMI 2010	7.39E-06	0.05922
	WAMI 2011	6.57E-05	0.0545
	WAMI 2013	4.17E-04	0.04708
<b>Haplotipo</b>	EYT2013-14	3.95E-10	0.05077
	EYT2014-15	1.90E-10	0.04661
	EYT2015-16	4.87E-09	0.04841
	WAMI 2010	8.95E-06	0.05902
	WAMI 2011	1.18E-04	0.05177
	WAMI 2013	3.51E-04	0.0485

La asociación de los marcadores KASP con el rendimiento solo se analizó en WAMI, el análisis basado en haplotipos mostró significancia estadística en los años 2010 y 2011 pero no fue así en 2013, el valor promedio de R<sup>2</sup> del marcador para estos tres años fue del 2 % (Cuadro 6). Los resultados obtenidos muestran un comportamiento propio de los QTL menores, los cuales tienen un efecto pequeño en el fenotipo y son afectados por las interacciones ambientales y epistáticas al

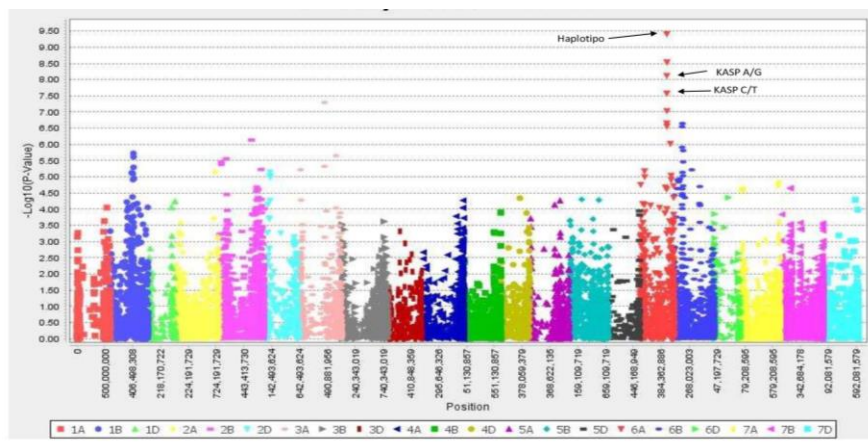
punto de enmascarar su efecto (Yang et al., 2013; Zhang et al., 2016). A pesar de las limitaciones de los QTL menores estos pueden usarse en conjunto con otros QTL para producir de manera aditiva o sinérgica incrementos en peso de grano y rendimiento (Zhang et al., 2016).

La superioridad de los análisis basados en haplotipos comparados con los marcadores individuales puede deberse a que los haplotipos tienen el potencial de estar en mayor desequilibrio de ligamiento con las variantes causales que los SNP individuales (Stram, 2004; Dupuis et al., 2011; Sehgal y Dreisigacker, 2019).

**Cuadro 6. Resumen de resultados de GWAS con marcadores individuales y haplotipos para validar la asociación de los marcadores KASP con rendimiento en WAMI**

	Año de cultivo	P-valor	Marcador_R <sup>2</sup>
<b>Marcador KASP C/T</b>	2010	3.57E-04	0.04059
	2011	0.00301	0.02652
	2013	0.24788	0.00466
<b>Marcador KASP A/G</b>	2010	1.79E-04	0.05491
	2011	0.01197	0.0271
	2013	0.07567	0.01831
<b>Haplotipo</b>	2010	2.07E-04	0.05396
	2011	0.00734	0.03038
	2013	0.30965	0.00846

Se trazaron los gráficos de Manhattan de las posiciones de los marcadores y los valores de  $-\log_{10}(\text{P-Value})$  (Figura 3).



**Figura 3** Gráficos de Manhattan de los resultados de GWAS para peso de mil granos en el ensayo de rendimiento EYT2013-14

### **Validación en familias endogámicas heterogéneas**

El ANOVA para los caracteres fenotípicos no mostró un efecto significativo de algún haplotipo para rendimiento, peso de mil granos, ancho y grosor de grano, en ningún ambiente, pero hubo significancia estadística para longitud de grano (Cuadro 7).

El efecto de los haplotipos en el rendimiento, peso de mil granos, ancho y grosor de grano fue inconsistente. El haplotipo TG confirió superioridad en peso de mil granos en los dos ambientes solo en las familias HIF 3, HIF 4 y HIF 9, el mayor incremento se registró en la familia HIF 3 cuando se cultivó en el centro de México con una diferencia del 13 %. El efecto de los haplotipos en rendimiento fue también variable, el haplotipo TG confirió mayor rendimiento en los dos ambientes solo en la familia HIF 8. El haplotipo TG confirió mayor ancho de grano en las familias HIF 3, HIF 4, HIF 6 y HIF 9 en los dos ambientes, el mayor incremento se registró en la familia HIF 3 cuando se cultivó en el centro de México con una diferencia del 3.8 %. El haplotipo TG confirió superioridad en grosor de grano en los dos ambientes solo en las familias HIF 3 y HIF 9.

Los genotipos portadores del haplotipo TG mostraron un efecto superior en longitud del grano en las diez familias cuando se cultivaron en el centro de México y en ocho familias cuando se cultivaron en el noreste de México, aunque solo hubo significancia estadística en la familia HIF 5 cuando se cultivó en el centro de México, el haplotipo favorable TG le confirió un incremento del 3 % en longitud de grano. La longitud promedio del grano de las diez familias en los dos ambientes fue de 6.47 mm para los genotipos con el haplotipo CA y de 6.55 mm para los genotipos con el haplotipo TG, lo que representa un incremento del 1.2 %.

Estudios anteriores de QTL que controlan longitud del grano, han reportado incrementos similares a los reportados en este estudio para longitud de grano. Por ejemplo, Guan et al. (2019) construyeron NIL para el QTL *QTgw.cau-4A.3* el cual se había reportado como controlador del tamaño del grano, las NIL fueron evaluadas en tres ambientes durante dos años, el alelo favorable de este gen confirió incrementos en longitud de grano desde 0.74 % hasta 3.87 %. Una investigación que se encuentra en preimpresión realizó una validación del QTL *QTgw.caas-5B* mediante NIL, el alelo favorable (5B+) confirió incrementos desde 2 % hasta 10 % en longitud de grano (Zhao et al., 2021).

**Cuadro 7. Comparación de medias de rendimiento, peso, longitud, ancho y grosor de grano en diez HIF portadoras de los haplotipos CA y TG del QTL HAP-6A-13 evaluadas en los ambientes Batán y Obregón**

Familia	Haplotipo	Rendimiento		Peso de mil granos		Longitud de grano		Ancho de grano		Grosor de grano	
		(g)	(g)	(g)	(g)	(mm)	(mm)	(mm)	(mm)	(mm)	(mm)
		Batán	Obregón	Batán	Obregón	Batán	Obregón	Batán	Obregón	Batán	Obregón
HIF 1	CA (-)	849.00	2070.00	43.75	42.80	6.40	6.54	3.52	3.35	3.38	3.01
	TG (+)	918.50	2034.00	44.90	42.30	6.54	6.50	3.51	3.30	3.38	3.01
HIF 2	CA (-)	798.50	2015.00	41.40	38.25	6.36	6.29	3.45	3.28	3.35	2.96
	TG (+)	742.50	1932.00	40.60	40.35	6.49	6.41	3.43	3.31	3.28	2.97
HIF 3	CA (-)	957.50	2130.00	39.40	35.80	6.26	6.38	3.39	3.09	3.24	2.82
	TG (+)	921.00	2054.00	44.80	38.60	6.42	6.53	3.52	3.15	3.34	2.87
HIF 4	CA (-)	987.50	1883.00	49.75	43.75	6.69	6.76	3.65	3.32	3.38	3.00
	TG (+)	890.50	1932.00	51.55	45.65	6.72	6.77	3.70	3.40	3.43	3.00
HIF 5	CA (-)	1037.50	2030.00	44.50	40.05	6.54**	6.67	3.47	3.15	3.24	2.92
	TG (+)	886.50	2029.00	45.15	40.05	6.74	6.73	3.47	3.08	3.24	2.84
HIF 6	CA (-)	977.00	1889.00	35.00	33.50	6.48	6.54	3.20	2.89	3.09	2.63
	TG (+)	830.00	1926.00	36.35	32.65	6.62	6.67	3.28	2.97	3.08	2.57
HIF 7	CA (-)	1037.00	2022.00	42.35	35.50	6.51	6.43	3.43	3.10	3.27	2.87
	TG (+)	927.00	1850.00	39.55	34.30	6.53	6.47	3.36	3.06	3.20	2.82
HIF 8	CA (-)	821.50	1933.00	44.40	39.90	6.41	6.63	3.42	3.21	3.38	2.87
	TG (+)	865.00	2022.00	43.20	38.95	6.53	6.52	3.39	3.13	3.30	2.89

Familia	Haplotipo	Rendimiento		Peso de mil granos		Longitud de grano		Ancho de grano		Grosor de grano	
		(g)		(g)		(mm)		(mm)		(mm)	
		Batán	Obregón	Batán	Obregón	Batán	Obregón	Batán	Obregón	Batán	Obregón
HIF 9	CA (-)	915.50	2017.00	38.15	33.25	6.22	6.13	3.33	3.06	3.15	2.78
	TG (+)	1064.00	1920.00	39.80	35.10	6.28	6.17	3.37	3.15	3.16	2.88
HIF 10	CA (-)	978.00	2140.00	48.35	41.95	6.68	6.71	3.60	3.30	3.37	2.94
	TG (+)	824.50	1903.00	47.10	41.75	6.79	6.75	3.58	3.32	3.29	2.96

Los asteriscos indican las diferencias significativas entre haplotipos dentro de una misma familia (\*P<0.05, \*\*P<0.01). El rendimiento se expresa como gramos obtenidos de la unidad experimental la cual corresponde a 2 m en Batán y 1.5 m en Obregón.



Los caracteres analizados mostraron grandes diferencias entre los dos ambientes cultivados, el rendimiento promedio fue superior en el noreste de México mientras que el promedio de peso de mil granos, ancho y grosor de granos fueron superiores en el centro de México, lo que indica que la diferencia ambiental tuvo un gran efecto sobre los caracteres estudiados.

Previo a este estudio sólo se había analizado la asociación del QTL *HAP-6A-13* con el peso de mil granos, con los nuevos datos y análisis realizados en esta investigación, se añade que el QTL también controla la longitud de grano. Estudios posteriores para el QTL *HAP-6A-13* pueden incluir GWAS de caracteres morfométricos para validar el efecto en longitud de grano.

La asociación significativa encontrada mediante GWAS para peso de grano y rendimiento, pero con falta de significancia al comparar líneas isogénicas demuestra que el estudio de QTL menores que afectan caracteres complejos sigue siendo un reto para la genética cuantitativa. El desempeño inestable de los QTL menores son resultado de las interacciones del QTL con el ambiente y del fondo genético, el cual propicia interacciones complejas con otros QTL.

## VI. CONCLUSIONES

El diseño de marcadores KASP a partir de secuencias GBS fue exitoso, los resultados de ambos métodos de genotipado fueron comparados y se obtuvieron altas similitudes. Mediante GWAS se verificó la asociación de los marcadores KASP con el peso de mil granos y rendimiento, obteniendo mejores resultados en el análisis basado en haplotipos. Los dos KASP diseñados son una herramienta para la selección asistida por marcadores para mejorar peso de grano y rendimiento.

En el análisis de HIF, el haplotipo TG mostró superioridad en longitud de grano respecto al haplotipo CA, en las diez HIF evaluadas en Batán y en ocho HIF evaluadas en Obregón, aunque solo se detectó diferencias significativas en una familia en un ambiente. El haplotipo favorable TG confirió un incremento promedio en longitud de grano del 1.2 % respecto al haplotipo CA en los dos ambientes evaluados.

Para el aprovechamiento del QTL *HAP-6A-13* en un programa de mejoramiento genético, se deben evaluar las combinaciones posibles con otros alelos favorables para peso de grano y rendimiento y seleccionar solo las combinaciones que actúan de manera sinérgica para la piramidación de alelos favorables.

## VII. LITERATURA CITADA

- Akhunov, E. D., Goodyear, A. W., Geng, S., Qi, L. L., Echaliier, B., Gill, B. S., Miftahudin, A., Gustafson, J. P., Lazo, G., Chao, S., Anderson, O. D., Linkiewicz, A. M., Dubcovsky, J., La Rota, M., Sorrells, M. E., Zhang, D., Nguyen, H. T., Kalavacharla, V., Hossain, K., ... Dvorak, J. (2003). The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosomes arms. *Genome Research*, *13*(5), 753–763. <https://doi.org/10.1101/gr.808603>
- Barton, N. H., Etheridge, A. M., & Véber, A. (2017). The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology*, *118*, 50–73. <https://doi.org/10.1016/j.tpb.2017.06.001>
- Bernardo, R. (2020). Reinventing quantitative genetics for plant breeding: something old, something new, something borrowed, something BLUE. In *Heredity* (Vol. 125, Issue 6, pp. 375–385). Springer Nature. <https://doi.org/10.1038/s41437-020-0312-1>
- Botta, V., Louppe, G., Geurts, P., & Wehenkel, L. (2014). Exploiting SNP correlations within random forest for genome-wide association studies. *PLoS ONE*, *9*(4). <https://doi.org/10.1371/journal.pone.0093379>
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. <https://doi.org/10.1093/bioinformatics/btm308>
- Brinton, J., & Uauy, C. (2019). A reductionist approach to dissecting grain weight and yield in wheat. *Journal of Integrative Plant Biology*, *61*(3), 337–358. <https://doi.org/10.1111/jipb.12741>
- Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLoS Computational Biology*, *8*(12). <https://doi.org/10.1371/journal.pcbi.1002822>
- Calderini, D. F., Castillo, F. M., Arenas-M, A., Molero, G., Reynolds, M. P., Craze, M., Bowden, S., Milner, M. J., Wallington, E. J., Dowle, A., Gomez, L. D., & McQueen-Mason, S. J. (2021). Overcoming the trade-off between grain weight and number in wheat by the ectopic expression of expansin in developing seeds leads to increased yield potential. *New Phytologist*, *230*(2), 629–640. <https://doi.org/10.1111/nph.17048>
- Calderini, D. F., & Reynolds, M. P. (2000). Changes in grain weight as a consequence of de-graining treatments at pre- and post-anthesis in synthetic hexaploid lines of wheat (*Triticum durum* x *T. tauschii*). *Australian Journal of Plant Physiology*, *27*(3), 183–191. <https://doi.org/10.1071/pp99066>
- Chen, Z., Cheng, X., Chai, L., Wang, Z., Bian, R., Li, J., Zhao, A., Xin, M., Guo, W., Hu, Z., Peng, H., Yao, Y., Sun, Q., & Ni, Z. (2020). Dissection of genetic factors underlying grain size and fine mapping of QTgw.cau-7D in common wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, *133*(1), 149–162. <https://doi.org/10.1007/s00122-019-03447-5>
- Cooper, J., Ibrahim, A., Rudd, J., Malla, S., Hays, D., & Baker, J. (2012). Increasing Hard Winter Wheat Yield Potential via Synthetic Wheat: I. Path-Coefficient Analysis of Yield and Its

- Components. *Crop Science*, 52. <https://doi.org/10.2135/cropsci2011.07.0382>
- de los Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D., & Calus, M. P. L. (2013). Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics*, 193(2), 327–345. <https://doi.org/10.1534/genetics.112.143313>
- Distelfeld, A., Avni, R., & Fischer, A. M. (2014). Senescence, nutrient remobilization, and yield in wheat and barley. *Journal of Experimental Botany*, 65(14), 3783–3798. <https://doi.org/10.1093/jxb/ert477>
- Dong, L., Wang, F., Liu, T., Dong, Z., Li, A., Jing, R., Mao, L., Li, Y., Liu, X., Zhang, K., & Wang, D. (2014). Natural variation of TaGASR7-A1 affects grain length in common wheat under multiple cultivation conditions. *Molecular Breeding*, 34(3), 937–947. <https://doi.org/10.1007/s11032-014-0087-2>
- Dreisigacker S., D. Sehgal, A. E. Reyes-Jaimez, B. Luna-Garrido, S. Muñoz-Zavala, C. Núñez-Rios, ... S. Mall (2016). CIMMYT Wheat Molecular Genetics: laboratory protocols and applications to wheat Breeding. CIMMYT, México. 142 p.
- Dreisigacker, Susanne, Xiao, Y., Sehgal, D., Guzman, C., He, Z., Xia, X., & Peña, R. J. (2020). SNP markers for low molecular 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>
- Duan, P., Rao, Y., Zeng, D., Yang, Y., Xu, R., Zhang, B., Dong, G., Qian, Q., & Li, Y. (2014). SMALL GRAIN 1, which encodes a mitogen-activated protein kinase kinase 4, influences grain size in rice. *Plant Journal*, 77(4), 547–557. <https://doi.org/10.1111/tpj.12405>
- Dupuis, M. C., Zhang, Z., Druet, T., Denoix, J. M., Charlier, C., Lekeux, P., & Georges, M. (2011). Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. *Mammalian Genome*, 22(9–10), 613–620. <https://doi.org/10.1007/s00335-011-9337-3>
- Falconer, D. S; Mackay, T. F. C. 2006. Introducción a la genética cuantitativa. ACRIBA, Zaragoza, España. 500 p.
- Gondro, C., Werf, J., & Hayes, B. (2013). Genome-Wide Association Studies and Genomic Prediction. Humana Press. USA. 566 p.
- Guan, P., Di, N., Mu, Q., Shen, X., Wang, Y., Wang, X., Yu, K., Song, W., Chen, Y., Xin, M., Hu, Z., Guo, W., Yao, Y., Ni, Z., Sun, Q., & Peng, H. (2019). Use of near-isogenic lines to precisely map and validate a major QTL for grain weight on chromosome 4AL in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 132(8), 2367–2379. <https://doi.org/10.1007/s00122-019-03359-4>
- Guo, H., Xiong, H., Xie, Y., Zhao, L., Gu, J., Zhao, S., Ding, Y., & Liu, L. (2019). Functional mutation allele mining of plant architecture and yield-related agronomic traits and characterization of their effects in wheat. *BMC Genetics*, 20(1), 1–15. <https://doi.org/10.1186/s12863-019-0804-2>
- Harlan, H. V, & Pope, M. N. (1922). The use and value of back-crosses in small-grain breeding. *Journal of Heredity*, 13(7), 319–322. <https://doi.org/10.1093/oxfordjournals.jhered.a102237>

- Hasan, A. K., Herrera, J., Lizana, C., & Calderini, D. F. (2011). Carpel weight, grain length and stabilized grain water content are physiological drivers of grain weight determination of wheat. *Field Crops Research*, *123*(3), 241–247. <https://doi.org/10.1016/j.fcr.2011.05.019>
- Henshall J. M. (2013). Validation of Genome-Wide Association Studies (GWAS) Results. In: Gondro C, van der Werf J and Hayes B (2013). *Genome-Wide Association Studies and Genomic Prediction. Methods in Molecular Biology*, Humana Press, Springer Science + Business Media. USA. 411 p.
- Hou, J., Jiang, Q., Hao, C., Wang, Y., Zhang, H., & Zhang, X. (2014). Global Selection on Sucrose Synthase Haplotypes during a Century of Wheat Breeding. *Plant Physiology*, *164*(4), 1918 LP – 1929. <https://doi.org/10.1104/pp.113.232454>
- Jiang, Q., Hou, J., Hao, C., Wang, L., Ge, H., Dong, Y., & Zhang, X. (2011). The wheat (*T. aestivum*) sucrose synthase 2 gene (TaSus2) active in endosperm development is associated with yield traits. *Functional and Integrative Genomics*, *11*(1), 49–61. <https://doi.org/10.1007/s10142-010-0188-x>
- Kim, K.-M., Cho, S.-W., Choe, R., Ryu, N. K., Hwang, J.-Y., Kim, J.-W., Kim, K.-H., Kang, C.-S., & Park, C. S. (2020). Allelic Variation at Loci Associated with Grain Properties in Korean Wheat Cultivars. *Korean Journal of Breeding Science*, *52*(4), 362–373. <https://doi.org/10.9787/kjbs.2020.52.4.362>
- Kooke R., Wijnker E., Keurentjes J.J.B. (2012) Backcross Populations and Near Isogenic Lines. In: Rifkin S. (eds) *Quantitative Trait Loci (QTL). Methods in Molecular Biology (Methods and Protocols)*, vol 871. Humana Press. 329 p. [https://doi.org/10.1007/978-1-61779-785-9\\_1](https://doi.org/10.1007/978-1-61779-785-9_1)
- Kraft, P., Zeggini, E., & Ioannidis, J. P. A. (2009). Replication in genome-wide association studies. *Statistical Science*, *24*(4), 561–573. <https://doi.org/10.1214/09-STS290>
- Kuchel, H., Williams, K. J., Langridge, P., Eagles, H. A., & Jefferies, S. P. (2007). Genetic dissection of grain yield in bread wheat. I. QTL analysis. *Theoretical and Applied Genetics*, *115*(8), 1029–1041. <https://doi.org/10.1007/s00122-007-0629-7>
- Li, F., Wen, W., Liu, J., Zhang, Y., Cao, S., He, Z., Rasheed, A., Jin, H., Zhang, C., Yan, J., Zhang, P., Wan, Y., & Xia, X. (2019). Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biology*, *19*(1), 1–19. <https://doi.org/10.1186/s12870-019-1781-3>
- Ling, H. Q., Wang, J., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., Fan, H., Li, D., Dong, L., Tao, Y., Gao, C., Wu, H., Li, Y., Cui, Y., Guo, X., Zheng, S., Wang, B., Yu, K., ... Zhang, A. (2013). Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature*, *496*(7443), 87–90. <https://doi.org/10.1038/nature11997>
- Liu, H., Zhang, X., Xu, Y., Ma, F., Zhang, J., Cao, Y., Li, L., & An, D. (2020). Identification and validation of quantitative trait loci for kernel traits in common wheat (*Triticum aestivum* L.). *BMC Plant Biology*, *20*(1), 529. <https://doi.org/10.1186/s12870-020-02661-4>
- Liu X, M. Huang, B. Fan, E. S. Buckler, Z. Zhang (2016). Iterative Usage of Fixed and Random

Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS Genetics* 12. <https://doi.org/10.1371/journal.pgen.1005767>

- Lopes, M. S., Dreisigacker, S., Peña, R. J., Sukumaran, S., & Reynolds, M. P. (2015). Genetic characterization of the wheat association mapping initiative (WAMI) panel for dissection of complex traits in spring wheat. *Theoretical and Applied Genetics*, 128(3), 453–464. <https://doi.org/10.1007/s00122-014-2444-2>
- Ma, L., Li, T., Hao, C., Wang, Y., Chen, X., & Zhang, X. (2016). TaGS5-3A, a grain size gene selected during wheat improvement for larger kernel and yield. *Plant Biotechnology Journal*, 14(5), 1269–1280. <https://doi.org/10.1111/pbi.12492>
- Mackay, T. F. C. (2001). The Genetic Architecture of Quantitative Traits. *Annual Review of Genetics*, 35(1), 303–339. <https://doi.org/10.1146/annurev.genet.35.102401.090633>
- Maphosa, L., Langridge, P., Taylor, H., Parent, B., Emebiri, L. C., Kuchel, H., Reynolds, M. P., Chalmers, K. J., Okada, A., Edwards, J., & Mather, D. E. (2014). Genetic control of grain yield and grain physical characteristics in a bread wheat population grown under a range of environmental conditions. *Theoretical and Applied Genetics*, 127(7), 1607–1624. <https://doi.org/10.1007/s00122-014-2322-y>
- Miao, L., Mao, X., Wang, J., Liu, Z., Zhang, B., Li, W., Chang, X., Reynolds, M., Wang, Z., & Jing, R. (2017). Elite haplotypes of a protein kinase gene TaSnRK2.3 associated with important agronomic traits in common wheat. *Frontiers in Plant Science*, 8:1–11. <https://doi.org/10.3389/fpls.2017.00368>
- Mohammadi, M., Xavier, A., Beckett, T., Beyer, S., Chen, L., Chikssa, H., Cross, V., Freitas Moreira, F., French, E., Gaire, R., Griebel, S., Lopez, M. A., Prather, S., Russell, B., & Wang, W. (2020). Identification, deployment, and transferability of quantitative trait loci from genome-wide association studies in plants. *Current Plant Biology*, 24, 100145. <https://doi.org/10.1016/j.cpb.2020.100145>
- Pask A., J. Pietragalla, D. M. Mullan, P. N. Chavez-Dulanto, y M. P. Reynolds (2013). Fitomejoramiento fisiológico II: una guía de campo para la caracterización fenotípica de trigo. CIMMYT. México. 132 p.
- Philipp, N., Weichert, H., Bohra, U., Weschke, W., Schulthess, A. W., & Weber, H. (2018). Grain number and grain yield distribution along the spike remain stable despite breeding for high yield in winter wheat. *PLOS ONE*, 13(10), e0205452. <https://doi.org/10.1371/journal.pone.0205452>
- Pumphrey, M. O., Bernardo, R., & Anderson, J. A. (2007). Validating the Fhb1 QTL for fusarium head blight resistance in near-isogenic wheat lines developed from breeding populations. *Crop Science*, 47(1), 200–206. <https://doi.org/10.2135/cropsci2006.03.0206>
- Ramirez-Gonzalez, R. H., Uauy, C., & Caccamo, M. (2015). PolyMarker: A fast polyploid primer design pipeline. *Bioinformatics*, 31(12), 2038–2039. <https://doi.org/10.1093/bioinformatics/btv069>

- Rasheed, A., & Xia, X. (2019). From markers to genome-based breeding in wheat. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *132*(3), 767–784. <https://doi.org/10.1007/s00122-019-03286-4>
- Rasheed, A., Xia, X., Ogonnaya, F., Mahmood, T., Zhang, Z., Mujeeb-Kazi, A., & He, Z. (2014). Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. *BMC Plant Biology*, *14*(1), 1–21. <https://doi.org/10.1186/1471-2229-14-128>
- Schaid, D. J., Chen, W., & Larson, N. B. (2018). From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature Reviews Genetics*, *19*(8), 491–504. <https://doi.org/10.1038/s41576-018-0016-z>
- Segura, V., Vilhjálmsson, B. J., Platt, A., Korte, A., Seren, Ü., Long, Q., & Nordborg, M. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature Genetics*, *44*(7), 825–830. <https://doi.org/10.1038/ng.2314>
- Sehgal, D.; Dreisigacker, S. (2019) Haplotypes-based genetic analysis: Benefits and challenges. *Vavilovskii Zhurnal Genet. Seleksii* 23:803–808.
- Sehgal, D., Mondal, S., Guzman, C., Garcia Barrios, G., Franco, C., Singh, R., & Dreisigacker, S. (2019). Validation of Candidate Gene-Based Markers and Identification of Novel Loci for Thousand-Grain Weight in Spring Bread Wheat. *Frontiers in Plant Science*, *10*(September), 1–12. <https://doi.org/10.3389/fpls.2019.01189>
- Shrimpton, A. E., & Robertson, A. (1988). The Isolation of Polygenic Factors Controlling Bristle Score in *Drosophila Melanogaster*. II. Distribution of Third Chromosome Bristle Effects within Chromosome Sections. *Genetics*, *118*(3), 445–459. <https://pubmed.ncbi.nlm.nih.gov/17246417>
- Simmonds, J., Scott, P., Leverington-Waite, M., Turner, A. S., Brinton, J., Korzun, V., Snape, J., & Uauy, C. (2014). Identification and independent validation of a stable yield and thousand grain weight QTL on chromosome 6A of hexaploid wheat (*Triticum aestivum* L.). *BMC Plant Biology*, *14*(1), 1–13. <https://doi.org/10.1186/s12870-014-0191-9>
- Stram, D. O. (2004). Tag SNP selection for association studies. *Genetic Epidemiology*, *27*(4), 365–374. <https://doi.org/10.1002/gepi.20028>
- Su, Z., Hao, C., Wang, L., Dong, Y., & Zhang, X. (2011). Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, *122*(1), 211–223. <https://doi.org/10.1007/s00122-010-1437-z>
- Sukumaran, S., Dreisigacker, S., Lopes, M., Chavez, P., & Reynolds, M. P. (2015). Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theoretical and Applied Genetics*, *128*(2), 353–363. <https://doi.org/10.1007/s00122-014-2435-3>
- Sukumaran, S., Lopes, M., Dreisigacker, S., & Reynolds, M. (2018). Genetic analysis of multi-

- environmental spring wheat trials identifies genomic regions for locus-specific trade-offs for grain weight and grain number. *Theoretical and Applied Genetics*, 131(4), 985–998. <https://doi.org/10.1007/s00122-017-3037-7>
- Sukumaran S., Yu J. (2014) Association mapping of genetic resources: achievements and future perspectives. In: Genomics of plant genetic resources, eds R. Tuberosa, A. Garner, and E. Frison. Springer, Netherlands, pp 207–235.
- Swarts, K., Li, H., Romero Navarro, J. A., An, D., Romay, M. C., Hearne, S., Acharya, C., Glaubitz, J. C., Mitchell, S., Elshire, R. J., Buckler, E. S., & Bradbury, P. J. (2014). Novel Methods to Optimize Genotypic Imputation for Low-Coverage, Next-Generation Sequence Data in Crop Plants. *The Plant Genome*, 7(3). <https://doi.org/https://doi.org/10.3835/plantgenome2014.05.0023>
- Tibbs Cortes, L., Zhang, Z., & Yu, J. (2021). Status and prospects of genome-wide association studies in plants. *Plant Genome*, 14(1), 1–17. <https://doi.org/10.1002/tpg2.20077>
- Tuinstra, M. R., Ejeta, G., & Goldsbrough, P. B. (1997). Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. *Theoretical and Applied Genetics*, 95(5), 1005–1011. <https://doi.org/10.1007/s001220050654>
- Tyagi, S., Mir, R. R., Kaur, H., Chhuneja, P., Ramesh, B., Balyan, H. S., & Gupta, P. K. (2014). Marker-assisted pyramiding of eight QTLs/genes for seven different traits in common wheat (*Triticum aestivum* L.). *Molecular Breeding*, 34(1), 167–175. <https://doi.org/10.1007/s11032-014-0027-1>
- Vanderzande, S., Piaskowski, J. L., Luo, F., Edge Garza, D. A., Klipfel, J., Schaller, A., Martin, S., & Peace, C. (2018). Crossing the Finish Line: How to Develop Diagnostic DNA Tests as Breeding Tools after QTL Discovery. *Journal of Horticulture*, 05(01), 1–6. <https://doi.org/10.4172/2376-0354.1000228>
- Xie, Q., Mayes, S., & Sparkes, D. L. (2015). Carpel size, grain filling, and morphology determine individual grain weight in wheat. *Journal of Experimental Botany*, 66(21), 6715–6730. <https://doi.org/10.1093/jxb/erv378>
- Xu, S. (2003). Estimating polygenic effects using markers of the entire genome. *Genetics*, 163(2), 789–801. <https://doi.org/10.1093/genetics/163.2.789>
- Yang, Y., Foulquié-Moreno, M. R., Clement, L., Erdei, É., Tanghe, A., Schaerlaekens, K., Dumortier, F., & Thevelein, J. M. (2013). QTL Analysis of High Thermotolerance with Superior and Downgraded Parental Yeast Strains Reveals New Minor QTLs and Converges on Novel Causative Alleles Involved in RNA Processing. *PLoS Genetics*, 9(8). <https://doi.org/10.1371/journal.pgen.1003693>
- Yang, Z., Bai, Z., Li, X., Wang, P., Wu, Q., Yang, L., Li, L., & Li, X. (2012). SNP identification and allelic-specific PCR markers development for TaGW2, a gene linked to wheat kernel weight. *Theoretical and Applied Genetics*, 125(5), 1057–1068. <https://doi.org/10.1007/s00122-012-1895-6>

- Zhang, H. W., Fan, Y. Y., Zhu, Y. J., Chen, J. Y., Yu, S. Bin, & Zhuang, J. Y. (2016). Dissection of the qTGW1.1 region into two tightly-linked minor QTLs having stable effects for grain weight in rice. *BMC Genetics*, *17*(1), 1–10. <https://doi.org/10.1186/s12863-016-0410-5>
- Zhang, P., He, Z., Tian, X., Gao, F., Xu, D., Liu, J., Wen, W., Fu, L., Li, G., Sui, X., Xia, X., Wang, C., & Cao, S. (2017). Cloning of TaTPP-6AL1 associated with grain weight in bread wheat and development of functional marker. *Molecular Breeding*, *37*(6), 1–8. <https://doi.org/10.1007/s11032-017-0676-y>
- Zhao, D., & Liu, D. (2021). Fine Mapping and Validation of a Major QTL for Grain Weight on Chromosome 5B in Bread Wheat. *Theoretical and Applied Genetics*. 1–15. <https://doi.org/10.21203/rs.3.rs-495159/v1>



## ANEXOS

**Anexo 1.** Pedigree de los *Elite Yield Trial* (EYT2013-14, EYT2014-15, EYT2015-16) archivos Excel en CD.